

NCHRP Web Document 12

(Project 3-47)

Contractor's Final Report

CAPACITY ANALYSIS OF INTERCHANGE RAMP TERMINALS

FINAL REPORT

**Prepared for
National Cooperative Highway Research Program
Transportation Research Board
National Research Council**

**Carroll J. Messer
Texas Transportation Institute**

and

**James A. Bonneson
University of Nebraska-Lincoln**

for

**Texas A & M Research Foundation
TAMURF 7241**

April 1997

ANS-

c. 2

ACKNOWLEDGMENT

This work was sponsored by the American Association of State Highway and Transportation Officials (AASHTO), in cooperation with the Federal Highway Administration, and was conducted in the National Cooperative Highway Research Program (NCHRP), which is administered by the Transportation Research Board (TRB) of the National Research Council.

DISCLAIMER

The opinion and conclusions expressed or implied in the report are those of the research agency. They are not necessarily those of the TRB, the National Research Council, AASHTO, or the U.S. Government.

This report has not been edited by TRB.

ACKNOWLEDGMENTS	iv
ABSTRACT	v
SUMMARY	vi
GLOSSARY	x
CHAPTER 1 Introduction and Research Approach	1-1
Operational Problem	1-1
Research Objective	1-2
Research Approach	1-3
CHAPTER 2 Findings	2-1
Summary of Existing Conditions	2-1
Survey of Existing Traffic Models	2-6
Field Studies	2-13
Capacity Characteristics	2-24
Traffic Control, Spillback and Performance	2-42
Arterial Weaving Speed	2-57
Ramp Weaving Capacity Model	2-61
CHAPTER 3 Interpretation, Appraisal, Applications	3-1
Capacity Characteristics	3-1
Interchange Capacity Estimation Methodology	3-11
Recommended Changes to the Highway Capacity Manual	3-20
Description of Software Needed for Implementation	3-31
Signal Timing Improvements for Closely Spaced Intersections	3-33
Arterial Weaving Speed Analysis Methodology	3-35
Ramp Capacity Analysis Methodology	3-39
CHAPTER 4 Conclusions and Recommendations	4-1
Conclusions	4-1
Recommendations	4-3
REFERENCES	4-5
APPENDIX A State-Of-The-Art	A-1
APPENDIX B Field Study Data Collection and Reduction	B-1
APPENDIX C Capacity Characteristics for Interchanges and Closely-Spaced Intersections	C-1
APPENDIX D Closely-Spaced Intersection Flow Models	D-1
APPENDIX E Arterial Weaving Speed Models	E-1
APPENDIX F Development of INTERCHANGE Analysis Software	F-1

ACKNOWLEDGMENTS

This research was performed under NCHRP Project 3-47 by the Texas Transportation Institute (TTI), Texas A&M University and the Department of Civil Engineering, University of Nebraska-Lincoln (UNL). A large number of researchers from both organizations contributed significantly to this research. Dr. Carroll J. Messer, P.E., of TTI served as Principal Investigator. Several other TTI staff made notable contributions to this research. The evaluation of existing computer models described in Chapter 2 was provided by Mr. Torsten Lienau. Mr. Michael D. Lloyd and Mr. Sriram Natarajan made major contributions to the experimental design and modeling of ramp/arterial weaving. Mr. Venugopal R. Neerudu conducted many of the statistical analyses presented and developed numerous figures shown. Ms. Sreelanta Nanduri developed the first operational computer code for TTI's PDX Model and Mr. Yang Ouyang further enhanced and tested it, as described in Appendix D. Mr. Joseph Van Arendonk developed the graphic designs and software code for INTERCHANGE, the interchange database software presented in Appendix F.

Dr. James A. Bonneson, P.E., of UNL served as Co-Principal investigator. The UNL staff members who made significant contributions to the research include Mr. Joel W. Fitts and Mr. James A. Kollbaum. The University of Nebraska's Center for Infrastructure Research provided funding for a state-of-the-art video data collection system that was essential to the field study activity associated with this project.

The following individuals and highway agencies were also very helpful in providing site locations and support during the field data collection task:

- Mr. Richard Berry, Traffic Engineering Division, City of Mesquite, Texas
- Mr. John Black, Traffic and Transportation Department, City of Richardson, Texas
- Mr. Jack A. Burgess, City of Newark, California
- Mr. Michael J. Cynecki, Street Transportation Department, City of Phoenix, Arizona
- Mr. Steven Dana, Department of Transportation, State of Arizona, Phoenix, Arizona
- Ms. Lynne B. Filson, Public Works Department, City of Antioch, California
- Mr. Glenn Hansen, Department of Public Works, City of Omaha, Nebraska
- Mr. Garry W. Metcalf, Traffic Engineering Department, City of Overland Park, Kansas
- Mr. Lynn J. Miller, Department of Transportation, State of California, Oakland, California
- Mr. Bob Simard, Department of Roads, State of Nebraska, Lincoln, Nebraska
- Mr. James W. Sparks, Street Transportation Department, City of Phoenix, Arizona
- Mr. Robert G. Temmermand, Department of Public Works, City of Sunnyvale, California
- Mr. James E. Tobaben, Department of Transportation, State of Kansas, Topeka, Kansas
- Mr. John H. Warner, Department of Transportation, State of Arizona, Phoenix, Arizona

These individuals assisted in the identification of candidate field study sites and also provided geometric and traffic volume data for these sites. In addition, many of these agencies graciously provided traffic control support at the start and end of each field study. The research team is grateful to each of these individuals and agencies.

ABSTRACT

This research study was conducted to develop and validate an appropriate methodology for determining capacity and level of service at signalized (service) interchanges for possible inclusion in a future edition of the *Highway Capacity Manual*. The research was sponsored by the National Cooperative Highway Research Program (NCHRP) and programmed as NCHRP 3-47, FY '94 under the title of "Capacity Analysis of Interchange Ramp Terminals." The research agencies conducting the work were the Texas Transportation Institute of Texas A&M University and the University of Nebraska-Lincoln. The study considered a wide variety of field characteristics that might affect saturation flow, including turning radius, traffic volume, lane use and downstream queue spillback. A large program of field studies was conducted at 12 interchanges located in 5 states. A sizeable data base of over 51,000 queued vehicles covering 3,800 signal cycles of operations was electronically recorded. These data were used in some cases to directly develop analytic models of a statistical nature; whereas in other cases, the field data were used to calibrate, test, and verify various analytic models having been developed from a more theoretical basis. As an example, the Prosser-Dunne model was enhanced and tested extensively using a NETSIM-based experimental process. Other prototype software (INTERCHANGE) was developed to illustrate how a large group of interchanges configurations might be systematically analyzed for capacity and level of service.

The research found that mobility at signalized interchanges is dependent on many factors that influence capacity. The importance of the prevailing factors may change depending on the general volume level and degree of existing congestion. Traditional prevailing factors affecting capacity include the interchange geometry, traffic mix, and signal green splits. The effect of queue spillback on upstream saturation flow was modeled and verified with field data. This model could prove very useful for providing reliable operational analysis of traffic conditions near capacity-flow levels. During periods of oversaturation where the arterial links are filled with queues, additional non-traditional factors come into play. During oversaturation, the upstream input capacity becomes highly dependent on the downstream signal timing and capacity to keep the output link flowing. Upstream input flow cannot exceed the total downstream capacity, and may be even less if demand starvation occurs. Signal offset during oversaturation was found to be the most important factor in determining which upstream signal phases can use the downstream capacity. This offset is not the offset that provides optimal traffic progression during uncongested flow periods. This report will be published as a NCHRP report at a later date.

SUMMARY

Chapters 9 and 11 of the *1994 Highway Capacity Manual* (HCM) cover signalized intersections and arterial streets, respectively, but no chapter addresses three important related considerations: (1) the types and operational features of signalized (service) interchanges, (2) signal coordination needs and progression characteristics of the signalized ramp terminals and other closely-spaced signalized intersections, and (3) the effects of queue spillback on the output flow of the crossing arterial and exit ramps at service interchanges.

The 1994 HCM capacity analysis methodology was judged to be inadequate for accurately evaluating closely-spaced intersections during near-capacity and oversaturated traffic conditions, primarily due to its lack of sensitivity to queue spillback approaching or impeding the subject intersection. Saturation flow estimates would not change in the HCM, as it has no methodology, outside of direct field observation, to deal with spillback. Moreover, the proposed 1997 HCM delay estimation update for intersections and arterials may over estimate its primary performance measure (delay) by a sizeable amount during oversaturation for closely-spaced arterial signals, even when the downstream output saturation flow is not impeded. In addition, the HCM definition of "effective green" needs to be revised to: (1) improve delay calculations during undersaturated conditions, and (2) cover periods of spillback and oversaturation.

NCHRP 3-47 was conducted to develop and validate an appropriate methodology for determining capacity and level of service at signalized (service) interchanges, including their ramp terminals and connecting closely-spaced arterial intersections. The research found that mobility at a signalized (service) interchange is dependent on many factors that influence capacity. The importance of the prevailing factors may change depending on the general volume level and degree of existing congestion. Traditional prevailing factors affecting capacity include the interchange geometry, traffic mix, and signal green splits of the serving phases.

During periods of oversaturation where the storage links are filled, additional non-traditional factors come into play. During oversaturation, the upstream input capacity becomes highly dependent on downstream signal timings and capacity to keep the output link clear. Upstream input capacity cannot exceed the total downstream service capacity available, and may be less than this capacity if demand starvation on short links occurs. In addition to phase capacity, signal offset during oversaturation is the most important factor in the allocation of downstream capacity to upstream phases. Also, this offset is not the offset that provides optimal progression during uncongested flow periods.

Signal timing and coordination together with signal spacing (link length) are important variables in determining the ability of the cross arterial to move traffic. Moreover, in the mesosaturation (near capacity) range of traffic conditions, the ultimate determination of the question of "oversaturation?" depends on whether potential flow impediments due to queue spillback can be sufficiently mitigated to the extent that oversaturation can be avoided.

Capacity analysis methods of signalized interchanges should be able to determine whether oversaturation will occur. However, this determination is a complex task in the mesosaturation range near capacity. These conditions not only depend on nominal volumes and phase capacities, but capacity also depends on the interactions of the variables with queue storage, link storage capacity and signal timing. Criteria have been provided that indicate the likelihood of queue spillback affecting nominal output capacity. These criteria can address probably 80 percent of the operational problems. However, close-calls can only be solved now using computer simulation models, like NETSIM, or possibly using models like the PDX Model developed in this research.

Phase capacity is dependent on the prevailing saturation flow rate and effective green time when motion can occur at the stopline. Based on this research, it was found that the distance to the downstream queue, the radius of the turn path, and traffic pressure have a significant effect on the saturation flow rate of a traffic movement. Specifically, saturation flow rate decreases when the distance to the downstream queue decreases and is relatively short. This effect is amplified when the signal timing relationship between the two intersections allows queue spillback to occur.

Turn radius has a significant effect on the saturation flow rate of a turn-related traffic movement. Saturation flow rates are lower for turn movements with small radii than they are for turn movements with large radii. The saturation flow rate for the left-turn movements at single-point urban interchanges are more nearly equal to those of through movements because of the large turn radii associated with this interchange type.

Traffic pressure, as quantified by traffic volume per cycle per lane, has a significant effect on saturation flow rate. Traffic pressure relates to the presence of aggressive, commuter drivers in the traffic stream. Traffic volume is used as a surrogate measure of the number of these aggressive drivers in the traffic stream. Saturation flow rates of low-volume movements are much lower than those of high-volume movements because the low-volume movements have less traffic pressure.

Other factors were examined for their potential effect on saturation flow rate. These factors include: g/C ratio, junction type, downstream signal indication at the start of the upstream phase, and dual versus single left-turn lane. Of these factors, only g/C ratio was found to be correlated with saturation flow rate in a statistically significant manner. Specifically, the saturation flow rate for left-turn movements with low g/C ratios was found to be higher than the rates of similar movements with larger g/C ratios. This effect was also found in the through movements studied; however, it was much smaller in magnitude and not statistically significant. Therefore, it was determined that more research is needed to verify the significance of this trend and its magnitude before an adjustment factor for g/C effect can be recommended.

The definition of effective green time should be changed to be only the time when saturation flow can occur at the stopline for existing conditions. This definition is more robust and can be used in all operating conditions, including periods of oversaturation. Moreover, delay estimates will be improved using this new definition of effective green even during undersaturated conditions because

the queue profile will be estimated more accurately than using the current HCM methodology.

The estimation of phase lost time should be improved. Start-up lost time is not a constant value; rather, it is statistically dependent on the prevailing saturation flow rate. Specifically, start-up lost time increases with increasing saturation flow rate. This paradoxical increase is due to the increased acceleration time the discharging queue requires to attain the higher speed associated with a higher saturation flow rate. The term "lost" is a bit of a misnomer in this case, as more lost time occurs as saturation flow gets better. Start-up lost times typically range from 0.61 to 3.18 seconds for prevailing saturation flow rates of 1,400 to 2,100 pcphgpl, respectively.

The average yellow warning interval used by drivers clearing the intersection at the end of the phase is termed "green extension, or end use." In the context of phase capacity, end use is equivalent to an extension of the effective green period into the yellow. The study of end use indicates that it is a relatively constant value for intersections and interchanges and that it averages about 2.5 seconds for most undersaturated conditions. Thus, this quantity can be subtracted from the signal change interval duration to estimate the lost time at the end of a signal phase.

Lane use is almost always uneven (or unbalanced) in intersection lane groups. The degree of this imbalance is expressed in terms of the lane utilization factor. The lane utilization factor varies depending on the nature of drivers' lane-choice decisions. Lane utilization factors based on travel time minimization tend to be subject to randomness in the lane-choice decision process. The factors stemming from this process range from 1.1 to 2.0, depending on the number of lanes in the lane group and its corresponding traffic volume. Lane utilization factors based on driver desire to preposition can vary widely, depending on the volume of traffic that is prepositioning in the subject lane group.

Neither the signal capacity of various interchange types nor their relative capacity per lane was specifically determined within this research. Some examples of this form of analysis are illustrated in a related NCHRP publication. However, examination of parclo as compared to diamonds reveals obviously different traffic volume input patterns that may result in one design being more efficient than another for a given case. The software INTERCHANGE described in Appendix F can readily examine the patterns provided by each interchange type. Parclo versus diamonds also have more right-turn capacity per input lane due to their normal signal overlaps, but this feature may tend to overflow downstream closely-spaced links more than diamonds. Moreover, single-point diamonds are known to have more arterial right-turn capacity per input lane than on its exit ramps because of unbalanced right-turn signal overlaps using three-phase signal operations.

The ideal saturation flow rate recommended for signalized (service) interchanges is 2,000 pcphgpl. In the context of the factors studied for this research, this ideal flow rate applies to through traffic movements that have an infinite distance to the back of downstream queues, operate under non-spillback conditions, discharge along tangent, level alignments, and have traffic volumes that are relatively high, reflecting those found during peak demand periods.

The equations provided in Chapter 3 should be used to estimate the saturation flow rate, start-up lost time, green extension, end lost time, and lane utilization factor. In recognition of the relationship between saturation flow rate and start-up lost time, the equations provided in Chapter 3 should be used to estimate *all* necessary phase capacity characteristics.

The recommended green extension value is 2.5 seconds for most undersaturated conditions. Other values are possible if the approach speed is outside the range of 64 to 76 km/h or when the volume-to-capacity ratio for the analysis period is above 0.88. An equation is provided in Chapter 3 for these situations.

The definition of effective green should be changed slightly from that used in the 1994 HCM. The new definition should be "effective green is that time during the subject phase when saturation flow at the stopline can occur under prevailing conditions." All "lost" times should be removed from the phase, including: start-up, opposing queue blockage, output blockage due to spillback, and phase clearance lost times. This definition is very robust and covers all operating conditions, movements, and phases, including protected-plus-permitted. The PDX Model should be considered for estimating the output Clear Period and effective green time of the subject phase when oversaturation is likely. High-volume links, which are nominally oversaturated or less than 200 meters long, should be analyzed for queue spillback blockages using the features provided in the PDX Model.

Ramp weaving speeds and crossing capacity can be estimated using the methodology presented in Chapter 3 and Appendix E. Adequate travel distance to the back of the downstream receiving queue must be available for this capacity to be attained.

The *Highway Capacity Manual* should contain a chapter on Interchanges which emphasizes the unique forms and features of interchanges together with the special challenges associated with urban interchange traffic operations in general. Two-level signalized interchanges operating within a crossing arterial system should be presented together with freeway system integration issues associated with freeway traffic management. Special design and operational issues dealing with continuous one-way frontage roads should be presented. Unsignalized rotary interchanges could be identified as an alternative design concept. Moreover, the selective application of signalized interchanges to upgrade the capacity of a major urban arterial corridor should be noted, as illustrated in Appendix F of this report.

A major development of computer software should be funded in support of the *Highway Capacity Manual* effort. However, this software should not be limited to being just a processor to the HCM Interchange chapter for conducting only "operational analysis of existing conditions." This new software should promote "options analysis" as needed to expediently conduct operational impact analyses for preliminary planning and design activities. The new program INTERCHANGE described in Appendix F illustrates recommended analysis concepts and software features. However, significant funding is still needed to complete a professional-level software package for signalized service interchanges.

GLOSSARY

- Δ = difference between the actual and ideal signal offset, sec;
- σ_{act}^2 = variance obtained from a regression analysis (i.e., = root mean square error²).
- σ_r^2 = variance in data due to random sources;
- σ_x^2 = variance in the independent variable;
- b_0, b_1, b_2, b_3 = calibration coefficients;
- C = cycle length, sec;
- c_i' = capacity of lane i (estimated as $c_i' = G / H_i$), vpcpl;
- CP = clear period during cycle/phase when subject flow is unblocked, sec;
- c' = capacity of the lane group, vpc;
- CV_{il} = critical lane volume of lane group i , vphpl;
- D = effective distance to the back of downstream queue (or stop line if no queue) at start of subject (or upstream) phase, m;
- $D_{m,w}$ = average maneuver distance for weaving vehicles ($= L_w - L_{q,w}$), m;
- $d_{m,i}$ = maneuver distance for vehicle i , m;
- $f(v'_{max})$ = distribution of the maximum demand flow rate in any of N lanes;
- f_{bb} = adjustment factor for bus blockage;
- f_D = adjustment factor for distance to downstream queue at green onset;
- f_g = adjustment factor for approach grade;
- $f_{g/C}$ = adjustment factor for signal timing;
- f_{HV} = adjustment factor for heavy vehicles;
- f_{LT} = adjustment factor for left-turns in lane group;
- f_p = adjustment factor for parking;
- f_{pp} = adjustment factor on delay depending on platoon arrivals;
- f_R = adjustment factor for radius of travel path;
- f_{RT} = adjustment factor for right-turns in lane group;
- f_u = proportion of drivers that do not attempt to evenly distribute themselves;
- f_v = adjustment factor for volume level (i.e., traffic pressure);
- f_w = adjustment factor for lane width;
- G = green signal interval, sec;
- g = effective green time where platoon saturation flow can occur at stopline, sec;
- G_{max} = maximum green signal interval duration for the subject (or upstream) phase that is allowable without spillback during saturated flows, sec;
- g_x = maximum g/C ratio (larger g/C ratios have no additional effect on headway);
- g_Y = effective green extension into yellow warning interval, sec;
- H = minimum discharge headway, sec/veh;
- h_c = clearance headway between last queued vehicle and first arriving vehicle, sec;
- H_i = minimum discharge headway in lane i , sec/veh;
- h_i = headway of the vehicle in the i th queue position, sec;
- H_j = minimum discharge headway based on specification of the j th queue position as the first to achieve minimum discharge headway, sec/veh;
- H_{li} = left-turn movement minimum discharge headway, sec/veh;

H_{th} = through movement minimum discharge headway, sec/veh;
 I_g = indicator variable (1.0 if $g/C < g_x$, 0.0 otherwise);
 I_L = indicator variable (1.0 if $D_{m,w} > 90 (N_i - 1)$, 0.0 otherwise);
 I_p = indicator variable (1.0 if $Max(v'_{dl}, v'_{dr})/v' > 1/N$, 0.0 otherwise);
 I_s = indicator variable (1.0 if spillback occurs during phase, 0.0 otherwise);
 I_X = indicator variable (1.0 if $X_i > b_3$, 0.0 otherwise);
 J = last queue position to discharge;
 j = "specified" first queue position to discharge at minimum discharge headway;
 L = distance between subject and downstream intersection stoplines (i.e., link length), m;
 l_e = clearance lost time at end of phase, sec;
 L_{HV} = lane length occupied by a queued heavy vehicle (≈ 13 m/veh), m/veh;
 L_{pc} = lane length occupied by a queued passenger car ($= 7.0$ m/pc), m/pc;
 $L_{q,w}$ = average length of queue joined by weaving vehicles, m;
 $l_{q,i}$ = length of queue joined by vehicle i , m;
 $l_{s(j)}$ = start-up lost time based on H_j , sec;
 l_s = start-up lost time to attain saturation flow, sec;
 L_v = average lane length occupied by queued vehicle, m/veh;
 L_w = distance between off-ramp entry point and stop line of downstream intersection (i.e., length of weaving section), m;
 $Max(v'_{db}, v'_{dr})$ = larger of v'_{dl} and v'_{dr} .
 N = number of lanes in lane group;
 N_d = number of through lanes on downstream segment, lanes;
 n_i = number of vehicles in queue on downstream street segment at end of phase, veh;
 n_s = number of vehicles on downstream street segment (moving or queued) at start of the subject phase, veh;
 N_t = number of arterial through lanes in the subject direction, lanes;
 off_a = actual offset between subject phase and downstream through movement (phase start time downstream minus phase start time upstream), sec;
 off_i = ideal offset to ensure progression without speed disruption, sec;
 P_{HV} = portion of heavy vehicles in the traffic stream;
 P_{LT} = portion of left-turns in the lane group;
 P_{RT} = portion of right-turns in the lane group;
 P_U = probability of a weaving vehicle being unblocked;
 P_{il} = average phase capacity at interchange I per lane, vphpl;
 R = radius of curvature of the left-turn travel path (at center of path), m;
 Φ = total interchange overlap at four-phase signalized interchange, sec;
 R^2_{act} = R^2 value obtained from a regression analysis;
 R^2_{best} = largest possible R^2 value obtainable (i.e., all systematic error explained);
 R_c = red clearance interval, sec;
 s = saturation flow rate for the lane group under prevailing conditions, vphg;
 s_0 = saturation flow rate per lane under ideal conditions, pcphgpl;
 s_j = saturation flow rate for the subject lane based on specification of the j th queue position as the first to achieve the minimum discharge headway, vphgpl;
 SL = approach speed limit, km/h;
 s_l = saturation flow rate per lane under prevailing conditions; vphgpl;

- s_n = saturation flow rate for the subject lane under prevailing conditions assuming the "no-spillback" condition, vphgpl;
 s_w = saturation flow rate for the subject lane under prevailing conditions assuming the "with-spillback" condition, vphgpl;
 T = duration of HCM delay analysis period, hours;
 $T_{d(j)}$ = discharge time of the J th queued vehicle, sec;
 $T_{d(i)}$ = discharge time of the i th queued vehicle ($i = j-1$ to J), sec;
 t_g = signalization variable ($0.0 < t_g < g_x$);
 t_j = time of event j for vehicle i (j = entry, stop, exit), sec;
 $t_{m,i}$ = maneuver time for vehicle i , sec;
 θ_{ij} = relative signal offset from upstream signal i to downstream signal j , sec;
 U = lane utilization factor for the lane group;
 U_a = average arterial speed entering the weaving section, m/s;
 $U_{m,a}$ = average maneuver speed for arterial through vehicles, m/s;
 $U_{m,w}$ = average maneuver speed for weaving vehicles, m/s;
 $u_{m,i}$ = maneuver speed for vehicle i , m/s;
 U_p = lane utilization factor for prepositioning;
 U_r = lane utilization factor for random lane-choice decisions;
 u_s = speed at saturation flow, m/s;
 V_a = average arterial flow rate entering the weaving section, vph;
 v_g' = unadjusted demand flow rate for the lane group, vpc;
 v_i' = demand flow rate in lane i , $i = 1, 2, \dots, N$, vpcpl;
 v_l' = demand flow rate per lane (i.e., traffic pressure), vpcpl;
 v_l' = demand flow rate per lane, vpcpl;
 V_w = average weaving flow rate, vph;
 v' = demand flow rate for the lane group, vpc;
 v'_{dl} = flow rate in lane group turning left at the downstream intersection, vpc;
 v'_{dr} = flow rate in lane group turning right at the downstream intersection, vpc;
 v'_{max} = maximum demand flow rate in any of N lanes, vpcpl;
 X = volume-to-capacity ratio for the lane group;
 X_i = volume-to-capacity ratio in lane i , $i = 1, 2, \dots, N$; and
 Y = yellow warning interval, sec.

CHAPTER 1

INTRODUCTION AND RESEARCH APPROACH

1.1 OPERATIONAL PROBLEM

Traffic congestion is a growing problem on all forms of traffic facilities, particularly at grade-separated, signalized interchanges in urban areas. Congestion reduces the efficiency and safety of these major traffic facilities, primarily due to problems associated with queue spillback and overall capacity deficiencies. Queue spillback can be caused by poor signal timing or by traffic demand exceeding the basic capacity of the facility to serve it. Queue spillback may surprise motorists who are not familiar with the area, causing serious merging and weaving problems from the off-ramp terminals to the cross street arterial (1,2). In addition, spillback can also reduce the output capacity of the signal when signalized intersections are closely spaced, say less than 200 meters apart.

The signalized service interchange, typically either a diamond or partial cloverleaf, serves a critical function in the urban highway system. Within freeway/arterial and arterial/arterial interchanges, signalized ramp terminals connecting to the arterial cross street are often the key operational geometric element. Unfortunately, the interchange facility, which is a very costly link in the highway network, often performs poorly, having numerous conflict points, high traffic demands, large changes in speeds required by motorists, and high-volume turning movements often exceeding four times the normal average observed at conventional intersections.

Traffic conflicts between the major through movements at interchanges are typically grade-separated, while the other traffic movements are served at signalized intersections. To improve major road operations, many maneuvers that tend to generate conflicts and delays (such as stopping, turning, and weaving) are designed to occur on the minor cross street. However, the cross street may also have large traffic volumes and the close spacing of the ramp terminals combined with the high volume of interchanging traffic often cause significant operational problems. These problems include long delays, poor minor-street signal progression, long queues and, as noted above, queue spillback between adjacent ramp intersections in some restricted cases.

Resolution of the operational problems noted above is critical to the safety and efficiency of the traffic corridor. Moreover, Intelligent Transportation Systems (ITS) technologies are dependent on the ability to divert traffic around congested freeway areas, often through congested signalized service interchanges. If the corridor interchanges are operating inefficiently, the potential diversion is diminished, and the benefits of ITS may be significantly reduced.

The *Highway Capacity Manual* (HCM) of 1994 (3) is the national guide (if not defacto standard) for conducting highway capacity analysis and level of service evaluations. Chapters 9 and 11 of the HCM cover signalized intersections and arterial streets, respectively. Neither of these chapters, nor others in the HCM, address two important considerations: (1) signal coordination needs

and progression characteristics of signalized ramp terminals (at interchanges) and other closely spaced signalized intersections, and (2) queue spillback on the cross street and onto the ramps. Moreover, operational problems resulting from oversaturation, when the existing traffic demand exceeds signal capacity, are only minimally addressed in the HCM.

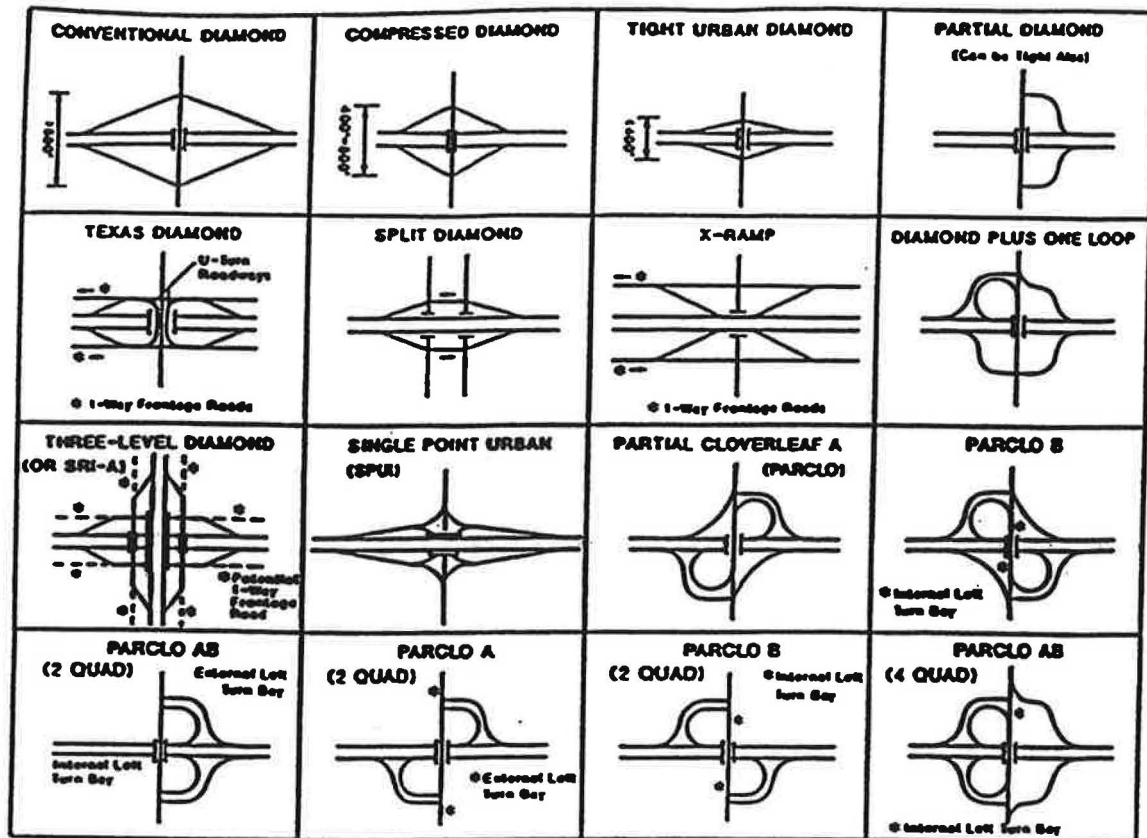
Thus, a methodology is needed to analyze signalized ramp terminals and adjacent intersections, simultaneously. The methodology should provide the procedures, guidelines and analytic tools needed to efficiently conduct pertinent capacity analyses. This methodology should be capable of addressing both coordinated pretimed and coordinated actuated signal systems through signalized interchanges. Undersaturated as well as oversaturated capacity conditions should be rationally addressed.

1.2 RESEARCH OBJECTIVE

The objective of this research was to develop and validate an appropriate methodology for determining capacity and level of service (LOS) at signalized ramp terminals of two-level grade-separated interchanges, including any adjacent closely spaced arterial intersections. The methodology should describe and quantify appropriate measures of traffic performance (measures of effectiveness, MOEs) at the signalized interchanges.

A wide range of service interchange configurations that include one or more traffic signals at the ramp terminals should be specifically addressed. This statement of scope includes conventional diamonds, compressed diamonds, tight urban diamonds, single-point interchanges, and three partial cloverleafs (parclos): the traditional names are the parclo A, parclo B, and the parclo AB. The parclo A is noted herein as a parclo AA because the major road's approach traffic sees both loop ramps in advance of the crossing structure; likewise, the parclo B is noted as the parclo BB because both loop ramps are seen beyond the cross street grade-separation structure. The parclo AB has one loop ramp in advance and one beyond the cross street bridge structure. Two basic varieties of parclos exist: the two-quad and four-quad designs. The four-quad separates two conflicting right turning movements from the signalized intersection area by placing each of them on direct connecting ramps. Thus, the four-quad parclo is more expensive than the two-quad design, but provides more capacity. Figure 1 illustrates the types of signalized service interchanges that were considered in this project. Unsignalized at-grade movements, such as right turns primarily from the off-ramps onto the cross street and resulting arterial weaving between the intersections, were also observed in some cases.

The research objectives of this project were to specifically (1) evaluate the state-of-the-art in operational analysis of signalized ramp terminals, including adjacent closely spaced intersections, (2) develop and test a new methodology for analyzing the capacity and LOS for these interchanges, and (3) validate the proposed methodology. The research results support the development of a new chapter in the HCM that is compatible with existing chapters. No operational computer software packages were to be developed as end products of this research effort.



NOTE: ADAPTED FROM MICHIGAN DOT & AASHTO

Figure 1. Examples of Signalized Service Interchanges in urban/suburban areas.

1.3 RESEARCH APPROACH

1.3.1 Field Survey

To insure that the scope and nature of the research were well defined, a two-stage survey of practicing traffic engineers was conducted during the initial task of this research. An objective of the survey was to gain insight into the current practices and concerns of engineers who evaluate interchange traffic operations. To achieve broad-based unbiased results, both public and private sectors were initially surveyed. The first-stage survey was mailed to more than 2,400 engineers across the U.S.A. and abroad during February, 1994. A total of 350 returned questionnaires were deemed completely responsive and valid. The first-stage survey results revealed that engineers analyze traffic operations of interchanges quite often, and that most service interchanges are diamonds (either compressed or tight urban). Most engineers responded that inadequate capacity and queue spillback are the most common operational problems experienced at the interchanges under their responsibility. All respondents cited operational problems existing at interchanges, and unsatisfactory methods of analysis. The findings indicated a need to further examine the reported traffic problems and the way they are analyzed.

The purpose of the second-stage survey was to further inquire about specific operational problems at interchanges, the methods that are presently used to analyze these problems, actions that are taken or recommended to resolve these problems, and identification of interchanges that could potentially serve as study sites. The second-stage survey was mailed during the last week of March 1994. This survey was sent to 190 individuals who indicated a willingness to respond to it from the first survey. The results from the second-stage survey helped in guiding the direction of research described in the following chapters. The second-stage survey was then used to determine what specific problems are encountered in interchange analysis. The results of the second-stage survey indicated that queue spillback, lane-changing, and arterial weaving between the ramp terminal and downstream intersection all deserved further investigation. Predicting the effect that queue spillback has on the output (saturation flow) of the traffic signal appears to be the most significant need for development. In addition, thirty-five interchanges located around the country were suggested as candidate study sites for subsequent model development, testing and validation.

1.3.2 Traffic Models

Following a detailed literature review of intersection and interchange capacity models, analysis methods and recent research, the development of a targeted set of traffic models was initiated toward satisfying the specific operational needs identified in the field survey and project objectives. Specified areas included: (1) discharge headways, (2) queue storage and spillback, (3) lane utilization, and (4) arterial weaving. The general effects of cross street (arterial) progression on queuing and delay were also desired together with a formal casting of the interchange capacity analysis methodology.

1.3.3 Field Studies

An extensive program of field studies was conducted of traffic operations at interchanges identified by the field survey as experiencing operational problems of congestion and queue spillback. A total of eighteen special studies were conducted at twelve interchanges located in five states. Data collection combined inputs from video cameras, tape switches and remote sensing of controller operations into a computer-based data collection system. Subsequent data processing by special video/traffic fusion software provided a chronographic record of observed traffic operations and control. Further model building and statistical analysis used standard statistical packages (4) and special software developed by the research team on a prior NCHRP project (2).

Traffic simulation was also used to develop and test analytic models of the targeted study areas. TRAF-NETSIM (5) was calibrated with field study results and used to develop analytic models of arterial weaving. Extensive testing and validation of other theoretical models of queue spillback and signal capacity were likewise conducted.

1.3.4 Study Results

The following chapter provides an overview of the research results followed by an appraisal of the significant findings, followed by our conclusions and recommendations. No software was specifically developed as deliverables for this project, although development of spreadsheets and some coding was necessary to test and evaluate some of the models and methods recommended. In particular, the PDX Model required extensive software coding to get the related complex algorithms described to perform efficiently for a wide range of traffic conditions. A prototype software design for service interchange database management, named INTERCHANGE, was also developed to demonstrate an efficient database architecture for processing all types of two-level signalized service interchanges for subsequent capacity and operational analyses.

The details of the user surveys, field studies, and traffic models are provided in subsequent appendices. Some data bases may be available on request through NCHRP to the subcontractor—University of Nebraska-Lincoln.

CHAPTER 2

FINDINGS

2.1 SUMMARY OF EXISTING CONDITIONS

A comprehensive evaluation of the state-of-the-art in areas related to interchange design and traffic operations was conducted as part of this research. This evaluation consisted of a survey of practitioners and a review of existing traffic models. The focus of this evaluation was on issues underlying the design and operation of interchanges in urban or suburban areas. More specifically, the focus was on issues related to the signal-controlled ramp terminals and traffic flow along the cross street through these terminals. Consideration was also given to the relationship between the interchange ramp terminals and any adjacent, closely-spaced signalized intersections.

2.1.1 Survey of Current Practice

The intent of this survey was to gain insight into the current practices and concerns of engineers who are responsible for interchange traffic operations. The survey was conducted in two stages. The first-stage survey was intended to obtain basic types of interchange-related information such as common interchange types, traffic flow problems, and operational analysis techniques.

The second-stage survey was designed to obtain more detailed information about interchange operations. This survey asked the respondent to select one interchange that they were familiar with and then respond to detailed questions about its operation and any steps taken to alleviate flow problems at this interchange. The respondent was also asked to describe the analysis techniques (or computer models) that they had successfully used to evaluate interchange operations. The findings from these two surveys are summarized in this section. A more detailed discussion of the survey findings is provided in Appendix A.

Distribution. The first-stage survey was sent to more than 2,400 transportation engineers in the U.S. and abroad. The members of the American Association of State Highway and Transportation Officials' (AASHTO) subcommittees on traffic engineering, on design, and on transportation systems operation were specifically targeted. A large number of the Institute of Transportation Engineers' (ITE) Urban Traffic Engineers Council and its Consultants Council were also included in the survey. In addition, several hundred surveys were sent to other selected members of ITE.

After a review of each returned questionnaire, a total of 350 first-stage questionnaires were deemed completely responsive and valid for further processing. Overall, there were 146 responses from the public sector which included state, city, and county highway agencies. Seventeen responses were received from outside of the United States. Responses were also received from 187 consultants in 23 states.

The second-stage survey was sent to 190 individuals who responded to the first survey. A total of 31 completed surveys were returned, representing a 16 percent response rate. Of these surveys, 29 were determined to be valid responses in the context that they addressed the interchange types and issues described in the survey. Overall, 21 states are represented among the 29 valid returned surveys.

Findings. The first-stage questionnaire consisted of six questions that were primarily of the multiple-choice type. The second-stage questionnaire consisted of eleven questions, several of which had follow-up questions. In general, these questions inquired about the kinds of interchanges being used or constructed, the type of signal control used, the types of operational problems found at existing interchanges, and the methods used to evaluate and mitigate these problems. The responses to the questions on both questionnaires are summarized in the following paragraphs.

The diamond interchange was found to be the most commonly used interchange configuration. This trend is likely due to the reduced right-of-way and construction costs associated with diamond interchanges relative to other configurations (e.g., partial cloverleaf). The distance between the diamond interchange ramp terminals can vary from 60 meters in densely-developed urban areas to 240 meters in suburban areas. In contrast, the distance between ramp terminals associated with a partial cloverleaf interchange generally range from 180 to 280 meters.

Regardless of configuration, the interchanges that tend to experience operational problems are those with relatively short distances between the ramp terminals or between one terminal and an adjacent signalized intersection. These close spacings often lead to problems such as queue spillback, flow turbulence due to weaving, and left-turn bay overflow. Queue spillback represents the blockage of an upstream intersection by a traffic queue from a downstream intersection. The interchanges described by the survey respondents as having operational problems had ramp terminal distances in the range of 61 to 410 meters. The distance to the adjacent intersection for these same interchanges was in the range of 46 to 436 meters.

The survey indicated that most interchanges have two semiactuated signal controllers, one controller for each ramp terminal. The two controllers are typically coordinated to facilitate progressed traffic flow along the arterial and minimum queuing on the street segment between the two terminals. Some interchanges have pretimed control with either one or two controllers. The few diamond interchanges that were pretimed and had one controller used four-phase-with-overlap phasing. Only a few interchanges had full-actuated, uncoordinated control.

The distribution of operational problems found in interchange areas is shown in Figure 2. As this figure indicates, the operational problem that occurred most frequently is queue spillback at some junction on the cross street. This problem was generally related to the spilling back of a queue from a downstream ramp terminal or intersection into an upstream terminal or intersection. This spillback tended to significantly reduce the capacity of the upstream junction. Also included in this category is spillback stemming from a left-turn bay overflow.

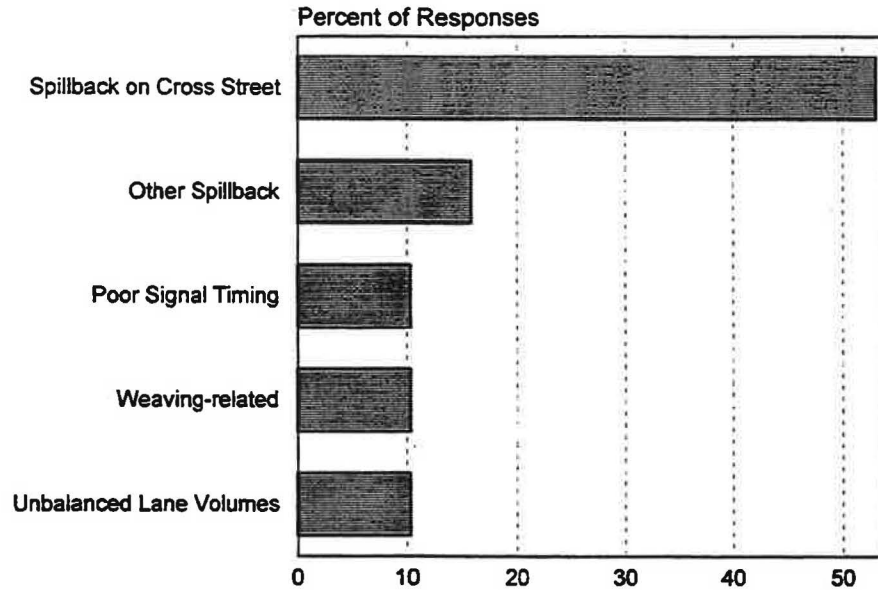


Figure 2. Distribution of operational problems in urban interchange areas.

The reported flow problems related to queue spillback between the ramp terminals were generally associated with tight or compressed diamond interchanges. Flow problems related to queue spillback between a ramp terminal and adjacent intersection were more commonly associated with conventional (wide) diamond interchanges and partial cloverleaf interchanges. The wide spacing between ramp terminals for these interchanges tends to be associated with shorter distances between these terminals and the adjacent intersections. By design, the single point diamond configuration does not experience spillback between its terminals; however, it can experience spillback between it and the adjacent intersection during high-volume conditions.

Other frequently cited problems at interchanges include unbalanced lane volumes on the ramp terminal approaches, flow turbulence due to weaving, and a lack of effective signal coordination between the ramp terminals. The unbalanced lane volume problem stems from frequent driver prepositioning for downstream turns in interchange areas. Drivers desiring to turn left (right) at a downstream intersection tend to move into the inside (outside) lane of a multilane lane group at the upstream intersection. This prepositioning effectively reduces the capacity of the lane group by leaving some traffic lanes underutilized, even during high volume conditions.

The weaving maneuver that is predominate in interchange areas is the off-ramp right-turn movement that weaves across the arterial to make a left-turn at the next downstream signalized intersection. This maneuver typically has a high volume associated with it such that considerable turbulence is created on the cross street. This turbulence results in significant speed reductions to the nonweaving traffic movements.

The lack of effective signal coordination along the cross street in interchange areas occurs for a variety of reasons. These reasons generally include incompatibility of interchange phasing with the cross street system coordination plan and institutional barriers (i.e., the state operates the interchange and the city operates the adjacent intersection). The lack of efficient signal coordination can lead to increased delays and stops and precipitate the occurrence of spillback when the ramp terminals or intersections are closely spaced.

A wide range of methods were described by the respondents for alleviating the aforementioned operational problems. Geometric improvements were most commonly cited. These improvements included adding a second left-turn lane or an additional through lane to the cross street. Many respondents indicated that improved or updated signal timing and coordination helped mitigate some operational problems. These latter improvements were often obtained through the use of existing software-based traffic analysis models.

In general, software programs are more frequently used than manual methods for evaluating interchange traffic operations. The most commonly used software program is the signalized intersection analysis procedure included in the Highway Capacity Software (HCS). In general, this procedure was used to evaluate the individual ramp terminals after appropriate calibration of the progression adjustment factors to account for nearby intersections. The popularity of this program may be due to its widespread acceptance by transportation engineers, its consistency with the methods described in Chapter 9 of the *Highway Capacity Manual* (HCM) (3), and the relative ease with which it can be used. The most frequently cited strength of this program is that it is easier to use than multiple-intersection software programs (e.g., PASSER II, TRANSYT-7F, NETSIM, etc.).

Of the various software programs available, TRANSYT-7F was cited by nearly half of all the respondents as being useful for analyzing interchange operations. This finding may be due to the fact that TRANSYT-7F is sensitive to the proximity of adjacent ramp terminals or signalized intersections in its signal timing optimization routine. Another software model, PASSER-II was also cited by many of the respondents as being a useful tool to analyze arterial traffic flow through interchange ramp terminals. In the case of this latter model, the large response may be due to the fact that PASSER-II optimizes signal phasing based on progression analysis. NETSIM was used by some of the respondents. This program was noted to be the only one that modeled queue spillback and congested flow conditions.

The respondents also noted that the existing software programs had some weaknesses that limited their ability to accurately model interchange traffic operations. The weaknesses cited for the HCS program (i.e., the HCM Chapter 9 procedure) were that it did not accurately model the effect of closely-spaced upstream intersections and that it did not yield queue length estimates. The weaknesses cited for PASSER II were that it did not provide progression solutions for left-turn movements, did not consider queues when determining progression, did not allow the user to enter some types of interchange phasing, and did not fully consider right-turn demand. NETSIM was noted to be very time consuming to use due to its microscopic simulation formulation. A couple of respondents noted that none of the programs dealt explicitly with the coordination of a downstream

ramp meter with the ramp terminal. Further evaluations of these computer-based models, primarily for research applications in this project, will be presented later in this chapter

The survey found that the most commonly selected measure of effectiveness (MOE) for evaluating interchange traffic operations was traffic signal delay, followed by spillback frequency, and volume-to-capacity ratio. Delay was likely chosen by the practitioners because it represents the most tangible measure of effectiveness that is also comprehensible by the motoring public. After delay, queue spillback frequency was the next most frequently cited MOE by the respondents.

2.1.2 Field Survey of Interchange Operations

The research team studied over a dozen service interchanges during the field studies and spent many hours observing traffic operations at the sites. Comparisons could be rapidly made among interchange types, types of operational problems observed, and the hypothesized probable cause of these problems. Our summary of these field sites having congested operations are noted below:

1. Design life of interchange probably exceeded; overall traffic demand exceeded interchange capacity during rush hours.
2. Due to growth in suburban areas, older four-lane crossing arterials now need to be six lanes. The average daily traffic on many of the four-lane crossing arterials exceeded 30,000 ADT.
3. Many "next" downstream signalized intersections along the crossing arterial experience high access demands to/from the freeway (interchange) and are routinely too closely spaced to provide good operating conditions. Better access management, intersection spacing and design policies are needed.
4. Traffic management of queuing and spillback is difficult at interchanges due to high volumes and high percentages of turning traffic having typical lane distribution problems. Some approaches along the crossing arterial and within the interchange can have almost constant demand within the cycle, so queuing can not be mitigated using traditional signal coordination techniques.

Four-quad parclo would seem to be more susceptible to constant demand conditions within the interchange because of their free flowing loop ramps. All parclo interchanges, including the four-quad AB that exits both left and right turns from the same side of a cross arterial approach, may experience high lane imbalances of arrival flow on that side of the street, even at intersections along the crossing arterial upstream of the interchange.

5. Many of the congested interchanges noted above had a predominant number of single-lane left turn bays within the interchange and/or have single lanes assigned on approach ramps

at ramp terminals to serve left and/or right turning movements. Many approach ramps were single lane with only a modest flare to a two-lane approach at the ramp terminal.

6. Traffic actuated operations on high-volume, single-lane movements appear to result in excessively long cycles that reduce the overall input capacity of the interchange. Protected-permissive left-turn operations, while reducing delays during moderate traffic, loses capacity during rush-hour conditions and, consequently, cannot be depended upon to provide significant capacity increases during these critical times.
7. Most traffic control strategies employed appear to be based on undersaturated flow conditions and may lose efficiency when oversaturated conditions arise. Management of queue spillback to mitigate the onset of congestion is needed together with the need to transition to downstream bottleneck control strategies once oversaturation has occurred.

2.2 SURVEY OF EXISTING TRAFFIC MODELS

The first-round survey inquired about the types of analysis methods used to evaluate (not optimize timing) signalized interchange traffic operations. In general, software models were more frequently used than manual methods. The most commonly used software method is the Highway Capacity Software (HCS). PASSER II and TRANSYT-7F were also found to be frequently used in practical engineering applications. However, research applications usually require more complex computer simulation models than application-specific models like HCS and PASSER II.

Computer simulation is a viable method with which to analyze situations which may occur at signalized interchanges, but for whatever reason are difficult to witness or collect data from field studies. This investigation was primarily based on literature and manuals for each model, and discussions with individuals familiar with the models. Experience with each model is arguably the most informative method of discovering what a program can and cannot do. Time constraints always limit the depth with which each of these models can be investigated. A list of the simulation models investigated is included in Table 1.

Simulation models can be described by their analysis approach, basis, objective and outcome. A model's analysis approach is either macroscopic or microscopic. A macroscopic simulation model is one in which the traffic stream is moved as one homogenous aggregate group, whereas a microscopic simulation model is vehicle specific in which each vehicle moves as its own identifiable entity. A simulation model's analysis basis is either empirical or analytical. The analysis basis refers to the algorithm on which the model is based. An empirical model is based on field observations or data and/or previous experience. Analytical models use mathematical formulas based on theoretical relationships. The analysis objective refers to the purpose of the simulation model. Models simulate traffic given certain geometric constraints, and/or optimize some specific traffic parameter. Lastly, a simulation model is described by its analysis outcome, which is either stochastic or deterministic. A stochastic model attempts to model human behavior by providing a

degree of randomness to its methodology. In this way the output is never the same given a set of inputs. Given the same inputs, a deterministic model would have the same output every time the same data is input. Each model's analysis is also given in Table 1.

Table 1. Simulation Models Examined

Model	Description	Analysis			
		Approach	Basis	Objective	Outcome
FREFLO	Freeway Simulation	Macroscopic	Analytical	Simulation	Deterministic
INTRAS 1980	Freeway and Surface Street Network Model	Microscopic	Analytical	Simulation or Optimization	Stochastic
CORFLO	Freeway and Surface Street Network Model	Macroscopic	Analytical	Simulation	Deterministic
INTEGRATION Version 1.5	Freeway and Surface Street Network Model	Microscopic	Analytical	Simulation or Optimization	Unknown
NETSIM	Urban Street Network Model	Microscopic	Analytical	Simulation	Stochastic
PASSER III 1990	Signalized Diamond Interchanges	Macroscopic	Analytical	Optimization	Deterministic
PASSER II 1990	Signalized Multi-Intersections	Macroscopic	Analytical	Optimization	Deterministic
TRANSYT-7F	Signalized Multi-Intersections	Macroscopic	Analytical	Simulation or Optimization	Deterministic
TEXAS Version 3.11	Isolated Intersection Analysis	Microscopic	Analytical	Simulation	Stochastic
HCS	Complete Implementation of 1985 HCM	Macroscopic	Empirical	Simulation	Deterministic
FREWEV Version 1.1	Freeway Weaving Analysis	Macroscopic	Empirical	Simulation	Deterministic

2.2.1 Input and Output

Obviously, each model has a required amount of input. Many models have options that may or may not be important to this project, and therefore the input for some data is optional. An abbreviated list of model inputs is included in Table 2. The table indicates the inputs (both required and optional) by each model. The list is not all-inclusive. Model names were abbreviated in Table 2, but they are presented in the same order as they are listed in Table 1.

Table 2. Model Inputs

Input	Model										
	FRE	INT	COR	ITG	NET	PIII	PII	T-7F	TX	HCS	WEV
Bus Stop Delay								X			
Capacity	X		X			X	X			X	X
Driver & Vehicle Characteristics		X							X		
Grades		X									X
Horiz. Curve Data		X									
Incident Data				X							
Intersection Spacing						X	X				
Link Lengths	X	X	X	X	X			X			X
Load Factors					X						
Number of Approaches						X	X	X	X		
Number of Lanes	X	X	X	X	X	X	X	X	X	X	X
O-D Travel Patterns				X	X						
Pedestrian Actuation					X					X	
Percentage of Vehicle Types		X								X	
Ramp Metering Rate		X		X							
Rte. Detouring Data				X							
Saturation Flows				X			X	X			
Signal & Sign Control Parameters		X		X	X	X	X	X	X	X	
Simulation Parameters	X	X	X	X	X			X	X		
Speed average free flow	X	X	X	X	X	X	X	X		X	
Through Volumes	X	X	X	X	X	X	X	X	X	X	X
Turning Percentages	X		X								
Turning Volumes		X		X	X	X	X	X	X	X	X
Vertical Curve Data		X									

The output available for each model investigated is included in Table 3. The table does not include all output for every model. FREFLO, INTRAS, CORFLO, INTEGRATION, NETSIM, and TRANSYT-7F display most of its output on a link specific basis. TEXAS Model provides output by lane, approach, and for the intersection as a whole.

Table 3. Model Outputs

Output	Model										
	FRE	INT	COR	ITG	NET	PIII	PII	T-7F	TX	HCS	WEV
Degree of Saturation						X	X	X			
Delay											
delay		X			X	X	X	X	X	X	
queue					X				X		
stopped					X				X		
Density	X	X	X								
Fuel Consumption		X				X	X				
Graphical Simulation				X					X		
Level of Service										X	X
Lane Changes											X
O-D Chart		X		X							
Optimal Timing						X	X	X			
Person ...											
trips	X		X								
miles	X		X		X						
minutes			X								
Queue Length		?		X	X		X	X			
Saturation Flow								X		X	
Speed											
mean	X	X	X	X	X				X		
time mean									X		
space mean									X		
Time Space Diag.						X	X				
Travel Time											
total	X	X		X	X			X	X		
average				X					X		
mean/vehm					X						
Vehicle...											
trips	X		X								
miles	X	X	X		X				X		
minutes		X	X								
in		X									
out		X									
stops				X	X	X	X	X			
% Ratio				X						X	
Volume			X	X				X			

2.2.2 Summary of Model Capabilities

An important decision in this project is what models should be used and how should they be used for research purposes. All the models investigated have some link to interchange and arterial operations. However, they may be used to develop relationships for situations where it would be difficult to collect field data. As a result, a list of geometric and operational characteristics, as well as other concerns, typical of interchange operations has been compiled in Table 4. Each model was then investigated as to its capability to model the stipulated geometric or operational characteristic. The results were shown in Table 4, and a brief discussion of the results follows.

An interchange ramp terminal/frontage road operates differently from an arterial street due to the effect of the freeway and its ramps. For this reason, a model capable of simulating traffic on both arterial streets and freeways would be advantageous. INTRAS and CORFLO are the only two models investigated in this initial study capable of interacting freeway vehicles and arterial street vehicles. Because INTRAS is a microscopic model, a greater level of detail can be both input and extrapolated from INTRAS than from CORFLO.

Weaving is another important factor. A level of service can be assumed from FREFLO output (and CORFLO) for weaving areas such as an entrance ramp closely followed by an exit ramp. For INTRAS, entrance/exit ramp weaving is not specifically addressed in the manual; however, TTI has used INTRAS for freeway weaving analysis and has found the model to operate adequately. However, it is improbable that the logic used in FREFLO and INTRAS for a freeway weaving analysis can be applied to a interchange ramp terminal weaving sections.

Other weaving scenarios involve the interaction of vehicles exiting the freeway and requiring a right turn at the ramp terminal intersection or vehicles turning out of a driveway and requiring a left turn at the first downstream intersection. These scenarios cannot be specifically modeled in INTRAS; however, INTRAS output does contain O-D charts which can quantify those maneuvers, and the output also quantifies the number of missed maneuvers. In other words, if a vehicle was destined to exit the freeway and turn right at the next intersection on the frontage road, but could not complete the maneuver, INTRAS includes this information in its output. NETSIM, on the other hand, is capable of traffic assignment parameters which could require a certain percentage of freeway exiting vehicles to turn right at the frontage road intersection. This process is, however, very complex and careful attention must be made to keep percentages of vehicle movements at each link equal to 100 percent. PASSER III deals specifically with diamond interchanges at which such a weaving maneuver would take place, however, simulation of weaving in the vicinity of the intersection is beyond its scope.

With interchanges being an integral part of freeway traffic management systems in some states, and with ramp metering becoming more prevalent, the issue of queue length could play an important role in freeway corridor operations. Queue length would aid in determining an adequate distance between a ramp exit or entrance and the interchange. Therefore, it would be desirable

Table 4. Computer Simulation Model's Capabilities

Model Constraints	Computer Model							
	FREFLO	INTRAS	CORFLO	NETSIM	PASSER III	PASSER II	TRANSY T-7F	TEXAS
Freeway Simulation	Yes	Yes	Yes	n/a	n/a	n/a	n/a	n/a
Frontage Road Simulation	n/a	Yes	Yes	Yes	No	No	No	No
Interchange Simulation	No	Yes	Unknown	Yes	Yes	Yes	Yes	Yes
Driveways	n/a	Yes	Yes	Yes	No	No	No	No
Type of Traffic Control	No Ramp Metering	Stop, Yield, Fixed, Actuated Control, 3 types of Ramp Metering, Merge and Diverge	Stop, Yield, Pretimed Signal Control, Some Actuated Control	Stop, Yield, Fixed Control, Actuated Control	Pretimed or Traffic-Responsive Fixed Sequence Signals	Signals	Pre-timed Signals or Unsignalized	No Control, Stop, Yield, Pretimed, Semi-actuated, or Full-Actuated
Freeway Weaving Analysis	LOS Provided	Yes	LOS Provided	n/a	n/a	n/a	n/a	n/a
Arterial Weaving Analysis Caused by Two Closely Spaced Ramps	n/a	Yes	No	Unknown	n/a	n/a	n/a	n/a
Varying Distance of Weaving Area	Yes	Yes	Yes	Yes	n/a	n/a	n/a	n/a
Arterial Weaving Analysis Caused by Either Vehicle Exiting Freeway and Turning Right at Interchange, or Vehicle Turning From Driveway and Turning Left at Intersection	n/a	O-D output	No	Yes	No	No	No	Unknown
Varying Distance Between Exit Ramp Terminal and Downstream Arterial Intersection	n/a	Yes	Yes	Yes	Yes	No	No	Yes
U-Turn Area at Interchanges	n/a	Yes	Unknown	Yes	Yes	No	Unknown	Yes
Exit Ramp Vehicles Able to Yield to Cross Arterial Traffic	n/a	Yes. All lanes yield	Unknown	Unknown	No	No	No	No
Prediction of Queue Length at Intersections	n/a	Unknown	Unknown	Yes	No	Yes	Yes	No
Realistic Output At or Near Capacity Levels	No	Yes	Unknown	Unknown	No*	No*	Unknown	Unknown

* at $v/c > 0.95$

to have queue length as an output from a model. Sources indicate a discrepancy regarding queue length output for INTRAS. The PASSER III model does not estimate queue length. However, PASSER II, although not specifically designed to handle diamond interchanges, can be used to evaluate a diamond interchange and produces similar results to PASSER III, with the added advantage of queue analysis output. In addition, NETSIM and TRANSYT-7F are both capable of producing queue lengths, whereas TEXAS is not.

High volume operations are critical conditions to be studied in this project. The effects of queue spillback on impeding and blocking output flow at an upstream intersection must be assessed. None of the models except INTRAS and NETSIM appear to perform as desired at oversaturated volume levels. PASSER II and PASSER III were not designed to operate at or near capacity, and may not produce realistic results when the v/c ratio is greater than 0.95. The literature does not say whether TEXAS produces realistic results when the v/c ratio nears one.

2.2.3 Other Considerations

Other considerations not included in Table 4, yet found in the literature and/or manuals that may be important in determining which model(s) to use in this project, are summarized in this section. One important factor associated with urban freeway ramp terminals that cannot be simulated by FREFLO is ramp metering. Also, the model deals with operations on freeways, whereas ramp terminals/arterial systems are important in this project. The main reason FREFLO was investigated is for its role in CORFLO and as a possible weaving analysis tool.

Other considerations regarding INTRAS not being included in Table 4 are that the manual is difficult to understand, and modeling a complex freeway/interchange systems could be very time consuming. Many inputs are required for any simulation run, which may require significant time for data preparation and processing.

It is interesting to note that one study that compared results of an analysis of a single point diamond interchange using several different models found that PASSER III and the TEXAS Model resulted in data most similar to the field data. PASSER II is somewhat more limited in its abilities than PASSER III with regard to frontage roads in that it cannot simulate U-turn lanes. Some additional disadvantages of the PASSER II program is that it does not allow for separate right turning lanes and the user cannot input the clearance interval of each phase. In addition, the TEXAS Model can only simulate one interchange/intersection at a time. Therefore a network evaluation of intersections would not be possible using TEXAS. Both PASSER models, the TEXAS Model, and TRANSYT-7F would be advantageous for this project to assess existing conditions at signalized interchanges; however, they do not take the whole interchange/arterial "experience" (driveways, weaving, etc.) into consideration.

2.2.4 Findings

All the models investigated are notable, respectable models that have their specific purposes. Because the geometric and operational characteristics of the interchange/arterial system are complex, each model investigated has its advantages and disadvantages for use. However, this investigation indicated that, in ranked order: TRAF-NETSIM, TEXAS, and TRANSYT 7-F were the more attractive operational models for potential research use in this project.

NETSIM was selected because it can simulate almost all aspects of interchange/arterial traffic operations desired. Its capability to view the simulation process gives the analyst an added sense of the fidelity of the simulation in progress. NETSIM offers the traffic assignment option which may also be helpful in analyzing arterial weaving caused by the interaction of turning vehicles and ramp-to-intersection spacing.

2.3 FIELD STUDIES

Several analytic models were developed during this study to facilitate the evaluation of interchange ramp terminal capacity and level of service. This section provides a description of the traffic flow problems for which models were developed, a description of the field study sites, and some summary statistics from the field study database. A more detailed discussion of the data collection and reduction activities is provided in Appendix B.

2.3.1 Traffic Flow Problems Associated With Interchange Ramp Terminals

This section describes traffic flow problems commonly found in interchange areas as related to the objectives of this research. Problems of primary interest were those occurring on the cross street at or between the interchange ramp terminals and any adjacent, closely-spaced intersections.

The findings from the survey of practitioners indicated that there were several types of traffic flow problems associated with signalized interchange ramp terminals. These flow problems were broadly categorized as: (1) midblock turbulence (i.e., weaving) and unbalanced lane volumes that stem from high-volume turn movements in the interchange vicinity; and (2) flow restriction or impediment to discharging queues due to a relatively near downstream traffic queue. Four models were developed to facilitate the evaluation of these flow problems. The variables included in these models were used to identify the data needed for model calibration. These four models are described in the remainder of this section.

Capacity Model. This model quantifies the effect of downstream traffic conditions on the traffic characteristics used to estimate the capacity of left-turn and through movements at interchange ramp terminals and adjacent intersections. These characteristics include start-up lost time, saturation flow rate, and clearance lost time. The capacity of an upstream signal phase has been found to be adversely affected by the close proximity of a downstream queue, particularly when the queue spills back into the upstream intersection.

Approach Lane Utilization Model. This model quantifies the extent of unbalanced lane use in multi-lane lane groups. On a cycle-by-cycle basis, many drivers in the interchange area tend to use one lane of a multi-lane lane group more than the others; they rarely choose the lane with the fewest vehicles in it. One possible reason for this unbalanced lane use in interchange areas may be driver desire to "preposition" for a downstream turn. As a result of this behavior, some lanes in a lane group are underutilized which effectively translates into a reduced lane group capacity.

Queue Length Model. This model can be used to convert a predicted queue length from the number of queued vehicles into units of distance (e.g., meters). This queue length conversion model was found to be an important component of the capacity model.

Arterial Weaving Model. This model quantifies the effect of weaving activity on the efficiency of arterial traffic flow. The weaving maneuver that is predominate in interchange areas is the off-ramp right-turn movement that weaves across the arterial to make a left-turn at the downstream signalized intersection. This maneuver has been observed to cause significant turbulence in the arterial traffic flow resulting in significant increases in travel time and, in some cases, lengthy queues on the off-ramp.

Figure 3 illustrates the extent of queuing that is commonly found during peak hours at many interchanges in urban areas. The queue shown in this figure extends back from the downstream intersection to the ramp terminal in the foreground. The proximity of this queue to the ramp junction was observed to significantly slow the discharge of arterial through traffic at the ramp terminal. It also caused the off-ramp drivers that desired to make a downstream left turn to cross the arterial at nearly right-angles in order to join the back of queue.



Figure 3. Queue growth between an interchange terminal and a downstream intersection.

Figure 3 also illustrates the field of view obtained from one of the two trailer-mounted video recording systems used during the field study. Figure 4 shows one of these systems. It was deployed at this location to record arterial weaving activity. The video camera for this system is mounted atop a 30-foot telescoping mast built into the two-wheeled trailer shown.

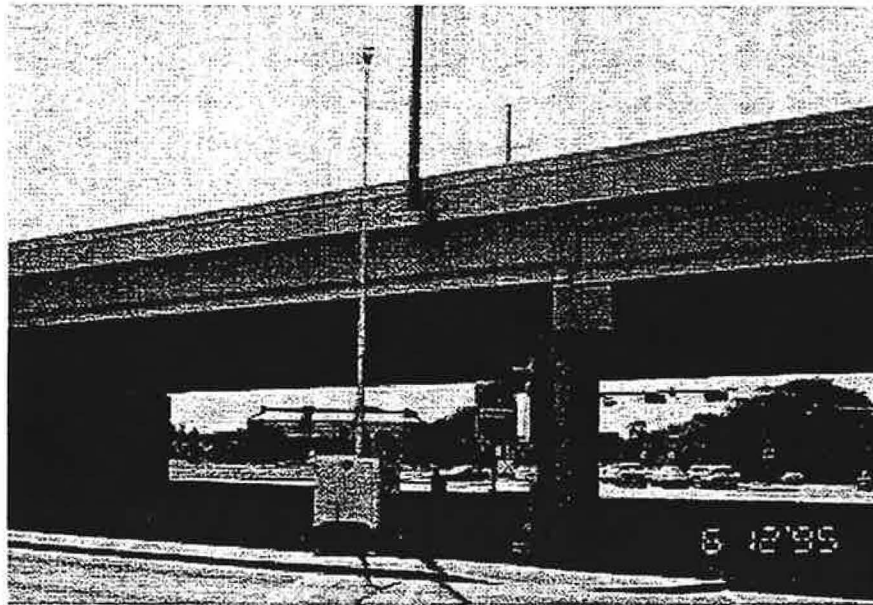


Figure 4. Video recording system used during the field study of queue interaction and weaving.

2.3.2 Field Study Site Description

The data collection plan was developed to obtain calibration data for the four models described in the preceding section. An initial step in developing this plan was the identification of sites that exhibited one or more of the four flow problems. It was also desired that the study sites collectively offer some diversity in their geometric design and geographic location. This section describes the characteristics used during the study site selection process and provides a brief description of the traffic and geometric characteristics of the twelve interchanges ultimately selected.

Study Site Characteristics. The study sites were selected to collectively include the two basic forms of service interchange commonly used in suburban and urban areas: the diamond and the partial cloverleaf (or parclo) interchanges. Variations of these two interchange forms stem from variations in the distance between the ramp terminals and from the routing of the traffic movements making the equivalent of a left or right-turn movement at the interchange. An assessment of the correlation between interchange type, the extent of its operational problems, and its frequency of application in urban areas led to the following six interchange types being identified as the most appropriate candidates for the field studies:

Diamond Interchange

1. Compressed Diamond
2. Tight Urban Diamond (without frontage roads)
3. Tight Urban Diamond (with frontage roads)
4. Single Point Urban Diamond

Partial Cloverleaf (Parclo)

5. Parclo B (2-quad)
6. Parclo AB (2-quad)

In addition to having one of the six interchange forms listed above, the interchange study sites were selected to have characteristics that would promote the four operational problems to be studied. Thus, sites were selected that had ramp terminal spacings of 275 meters or less; ramp-to-intersection spacings of 275 meters or less, average daily traffic demands in excess of 20,000 vpd, and generally unconstrained geometrics (i.e., 3.6-m lanes, no curvature, minimal grade, etc.). In addition to these characteristics, the study sites were selected to have frequent and recurring traffic queues on the arterial during the peak traffic periods.

Study Site Locations. In addition to the aforementioned characteristics, there was a need for geographic diversity in the collective list of study sites. In this regard, study sites were identified in six geographic regions of the U.S. Within these regions, highway agencies in the states with large metropolitan areas were contacted and inquiry was made as to potential study locations. Interchanges that had the desired characteristics were identified as candidates for a preliminary site visit.

Based on the results of the preliminary visit to the candidate sites, twelve interchanges were identified as being most suitable for field study. Table 5 describes the traffic and geometric characteristics of these twelve interchange study sites. Details of the traffic signalization at each of these sites is provided in Appendix B.

2.3.3 Data Collection

The data collection activities focused on the collection of the basic characteristics describing traffic flow at and between signalized ramp terminals and adjacent intersections. The data collected at the terminals and intersections included discharge headways, speeds, driver use of the yellow interval, and lane utilization. The data collected between the ramp terminals and the adjacent intersections included the speed and volume of weaving and non-weaving vehicles.

The equipment used to collect the field data included video cameras and computer-monitored tape switch sensors placed in the traffic lanes. The equipment deployment followed one of two study types (i.e., a capacity or weaving study). All data were collected during weekday, daytime periods between the hours of 7:00 a.m. and 7:00 p.m.

The typical data collection setup for a capacity study is shown in Figure 5. As indicated in this figure, a trailer-mounted video camera was located at the upstream end of each of two street segments. The data collection setup for a weaving study was similar to that for the capacity study; however, the video cameras were located at both ends of the arterial segment between the adjacent intersection and nearest interchange off-ramp.

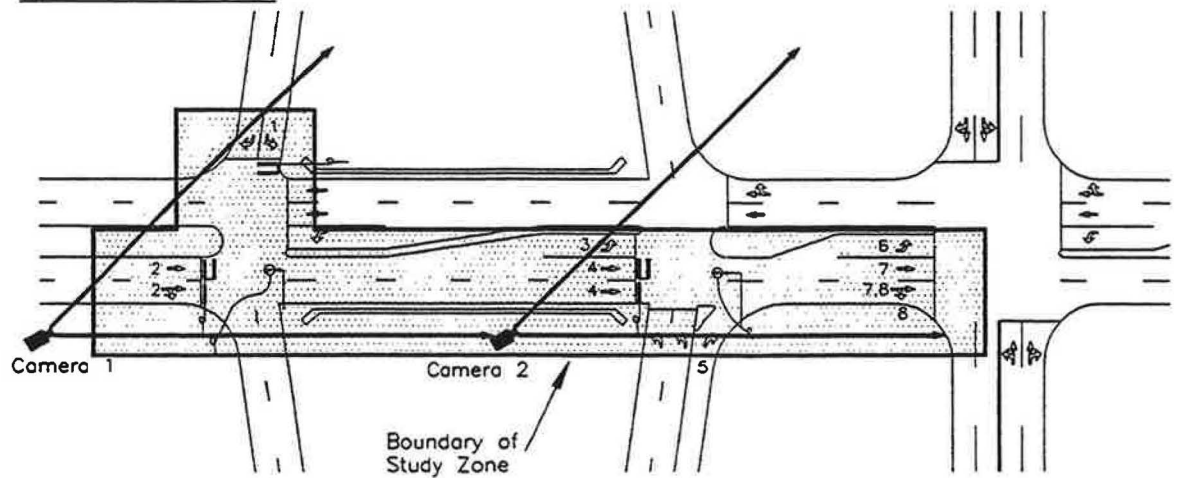
Table 5. Traffic and geometric characteristics of the study sites

Interchange Type	Arterial	City, State	Arterial AADT	Arterial Thru Lanes	Ramp to Ramp Distance (meters)¹	Ramp to Intersection Distance (meters)¹	Speed Limit (km/h)
Compressed Diamond	Metcalfe Ave 110th to I-435	Overland Park, Kansas	58,600	6	200	204	72
	75th Street I-35 to Frontage	Overland Park, Kansas	32,000	4	174	155	56
	Maple Street 102nd to I-680	Omaha, Nebraska	34,200	4	268	198	72
Tight Urban Diamond	Peoria Road 25th Ave. to I-17	Phoenix, Arizona	34,400	6	107	276	64
	Mathilda Ave SR-237 to Ross	Sunnyvale, California	34,540	6	87	110	72
Texas Diamond	Arapaho Road US75 to Greenville	Richardson, Texas	39,000	6	99	265	64
	Towneast Blvd Emporium to I-635	Mesquite, Texas	35,000	6	137	223	56
Parclo AB (2 quad)	60th Street I-80 to Grover	Omaha, Nebraska	31,800	4	259	216	64
Parclo B (2 quad)	Somersville Rd Delta Fair to SR-4	Antioch, California	39,700	4	265	119	56
	Stevenson Blvd Balentine to I-880	Newark, California	55,600	4	264	157	56
Single Point Urban Diamond (SPUI)	7th Street I-10 to McDowell	Phoenix, Arizona	42,000	6	78	331	56
	Indian School Rd 16th St. to SR-51	Phoenix, Arizona	54,500	6	91	316	56

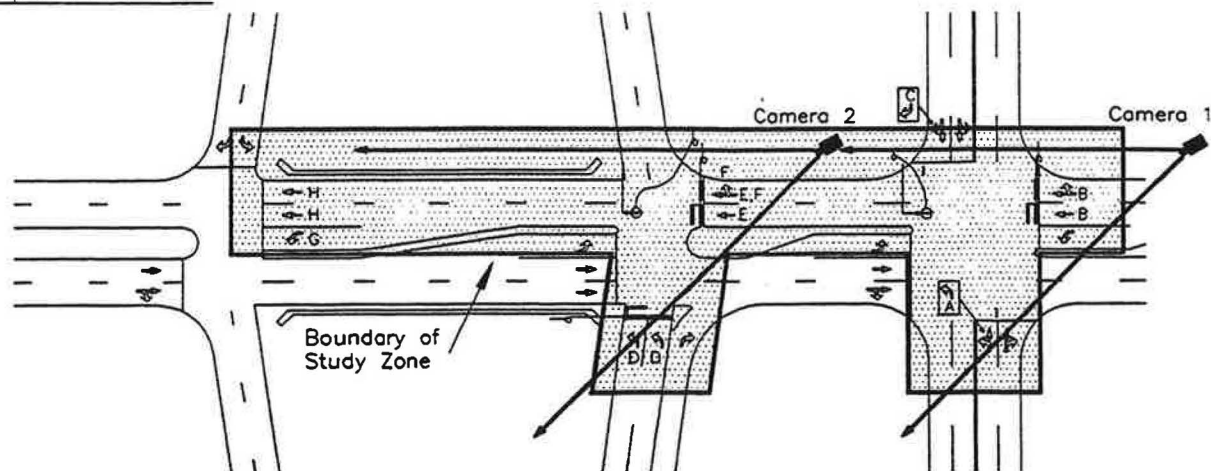
Notes:

- 1 - Distance measured from stop line to stop line in the same direction, except at SPUI's. At SPUI's, the "same direction" concept is also applied but the opposing direction through stop line is used as the reference point at the second ramp terminal (since the through stop line at the second ramp terminal does not exist at the SPUI).

Downstream Case



Upstream Case



LEGEND

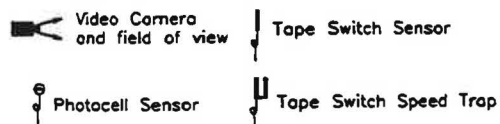
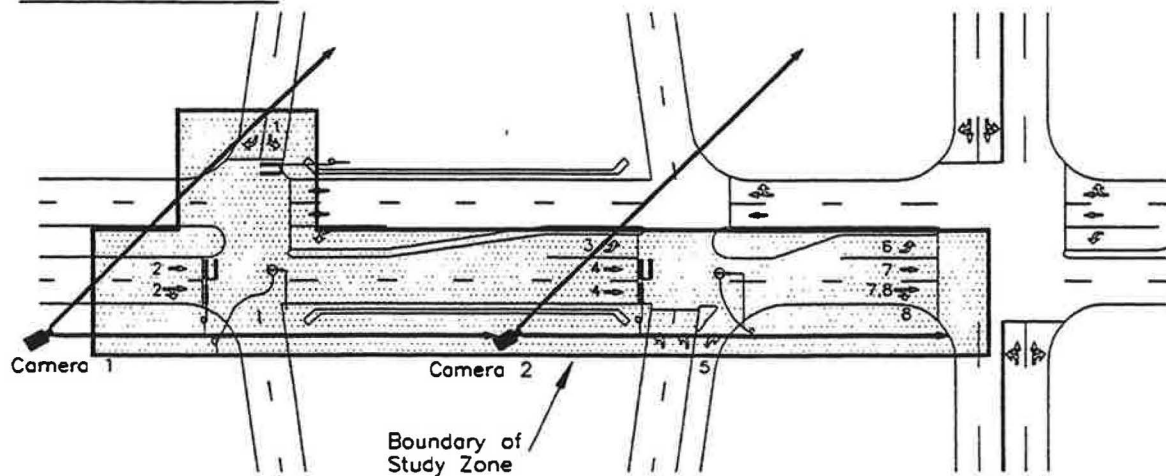
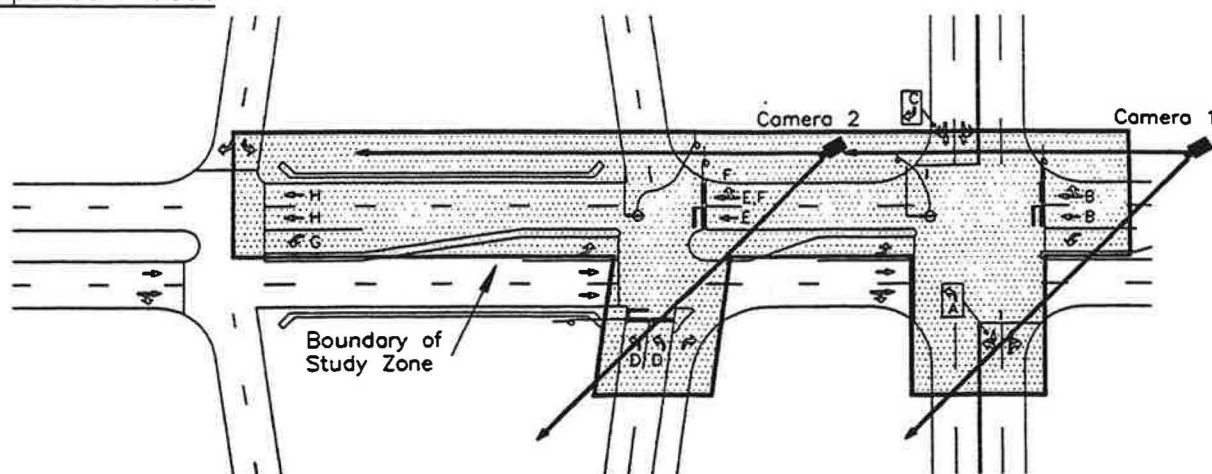


Figure 5. Capacity study data collection setup for a diamond interchange.

Downstream Case



Upstream Case



LEGEND

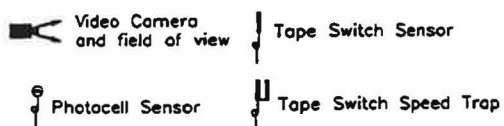


Figure 5. Capacity study data collection setup for a diamond interchange.

During each of the capacity studies, the video and tape-switch equipment were deployed in a manner consistent with that shown in Figure 5. Data collected during the capacity studies were used to calibrate the capacity and lane utilization models. The tape switches were used to record traffic flow behavior in two traffic lanes on three intersection or ramp terminal approaches. The video recorders were positioned to provide a visual record of traffic crossing the tape switches as well as information about queuing conditions on the downstream street segment.

2.3.4 Database Summary Statistics

The data reduction effort proceeded on a model-by-model basis. For all models, the data reduction procedures were defined, documented, and tested on a sample portion of the collected data prior to their full-scale implementation. Following data reduction, and prior to model calibration, relevant summary statistics for each of the data bases were computed and reviewed. This review included the computation of sample statistics for selected traffic characteristics and performance measures. These statistics are categorized by junction type and traffic movement where appropriate. The findings of this review are summarized in this section; a more detailed examination is provided in Appendix B.

Capacity. The collected data were used to create a database of traffic characteristics and performance measures for use in calibrating the capacity model. This database included vehicle-related data (i.e., discharge headway, speed, acceleration) and phase-related data (i.e., phase duration, cycle length, distance to downstream queue). It includes the discharge characteristics of more than 51,000 queued vehicles. These vehicles were observed at twelve sites for 33 traffic movements in 63 instrumented lanes. All of the sites had extensive traffic queues during some or all of the six-hour study period and at least eight sites had some degree of queue spillback. There were more than 3,800 signal cycles observed during the study periods.

The saturation flow rate, start-up lost time, and clearance lost time are summarized in Table 6 for the two junction and movement types studied. As the columns in this table indicate, the saturation flow rate and start-up lost time data were segregated into "with" and "without" spillback categories. The "with" spillback category relates to the vehicles observed during signal phases that experienced queue spillback from the downstream intersection. The data included in this category represent only those vehicles able to discharge *before* the onset of spillback. Vehicles that discharge prior to spillback were found to have low saturation flow rates; they had little incentive to discharge at higher rates because they were essentially discharging into the back of the downstream queue.

The saturation flow rate for each traffic lane studied was computed as the reciprocal of the minimum discharge headway measured for that lane. This latter quantity was computed as the average of all headways observed for the fifth and higher queue positions in each traffic lane studied. This technique for computing the saturation flow rate is consistent with the procedure described in the 1994 HCM (3, Chapter 9).

Table 6. Capacity database summary statistics

Junction Type	Movement Type ¹	Without Spillback			With Spillback ³			Clearance Lost Time (sec)
		No. Cycles	Sat. Flow Rate ² (pcphgpl)	Start-Up Lost Time (sec)	No. Cycles	Sat. Flow Rate ² (pcphgpl)	Start-Up Lost Time (sec)	
Interchange	Left-Turn	1,564	1,957	2.80	--	--	--	2.77
	Through	2,057	1,925	2.65	52	1,659	1.69	2.60
Intersection	Left-Turn	15	1,967	4.40	6	1,622	3.04	2.55
	Through	1,474	1,915	2.46	108	1,667	1.87	3.11
Average or Total:		5,110	1,935	3.08	166	1,651	2.20	2.77

Notes:

1 - Left-turn movements from both the off-ramp and the arterial. Through movements along the arterial.

2 - Based on the average headway of the fifth through last queued passenger car.

3 - Based on the average headway of the fifth through last queued passenger car able to discharge prior to queue spillback from the downstream intersection.

-- no data available.

In general, the saturation flow rate is very similar among the interchanges and intersections studied. An examination of the saturation flow rates categorized by interchange type (e.g., compressed diamond, parclo B, etc.) indicated that there were no significant differences in flow rate among types. On the other hand, the data in Table 6 indicate that the left-turn movements may be discharging more efficiently than the through movements at the study sites, however, the difference is relatively small.

As with the saturation flow rates, the start-up lost times in Table 6 varied among the "with" and "without" spillback categories. In general, start-up lost time in the "without" spillback category tends to be higher as a consequence of the extra time lost by the discharging traffic queue as it accelerates to the higher speeds associated with the higher saturation flow rates. Typical values of start-up lost time for the "without" category range from 2.46 to 2.80 seconds (excluding the data in the "intersection/left-turn" category) whereas values for the "with" category range from 1.69 to 1.87 seconds. The difference between these two ranges suggests that the more normal, "without" spillback condition is associated with about 1.0 seconds more start-up lost time than the "with" spillback condition.

Table 6 also summarizes the lost time at the end of the phase for the junction and movement types studied. The clearance lost time reported in this table was computed as being equal to the yellow-plus-red-clearance interval less the initial portion of the yellow interval used by the average driver (i.e., green extension). In general, it was found that drivers entered the intersection after the yellow was presented in about 27 percent of the phases studied; although, the frequency of this behavior varied widely among the study sites. The average amount of green extension was found to be relatively constant at 2.5 seconds across the twelve study sites during uncongested conditions.

The data in Table 6 indicate that the average clearance lost time is about 2.77 seconds. This value is within the range of 1.2 to 2.8 seconds recognized by the 1994 HCM (3, *Chapter 2*), although it is very near the upper limit of this range. This trend is likely due to the longer change intervals used at some of the interchanges and intersections studied.

Lane Utilization. Traffic events recorded on video tape during the capacity studies were used to create a database of traffic characteristics and performance measures for calibrating a lane utilization model. The data collected included the lane volume per cycle, number of approach traffic lanes, distribution of traffic volumes to downstream turns, and the type of interchange. The assembled lane utilization model database includes the entry and exit time and location for 8,198 vehicles observed at twelve sites for 32 traffic movements. Of these vehicles, about 65 percent represent through movements; the balance were left-turn vehicles in multi-lane lane groups.

The analysis of the lane utilization database focused on the computation of a lane utilization factor for each of the left-turn and through movement lane groups studied. This utilization factor was computed using the following equation:

$$U = \frac{v'_{max} N}{\sum v'_i} \quad (1)$$

where:

- U = lane utilization factor for the lane group;
- v'_{max} = maximum demand flow rate in any of N lanes, vpcpl;
- v'_i = demand flow rate in lane i , $i = 1, 2, \dots, N$, vpcpl; and
- N = number of lanes in the lane group.

The lane utilization factors computed for the through movement lane groups at the twelve study sites are shown in Table 7. The factors recommended in the 1994 HCM (3, *Chapter 9*) for application at isolated intersections are also shown in this table. These recommended values are consistent with the computed values in that larger factors are associated with lane groups with a larger number of lanes. In contrast, the recommended values tend to be smaller than the computed values. This trend suggests that lane utilization in interchange areas tends to be more unbalanced than at isolated intersections. This result was anticipated because of the significant turning activity in interchange areas and the resultant need for drivers to preposition themselves in the left-most (or right-most) lane on the street segment prior to the segment from which the turn will be made.

Queue Length. The data reduction procedure for the queue length database required a camera view of the front and back of the through movement traffic queue on an intersection approach. The front view was used to measure the distance-to-stop-line and starting-reaction time of the first queued driver. The back view was used to measure the same statistics for the last queued vehicle. All distance measurements were made at the start of the phase; all reaction time measurements were made relative to the start of the phase. Queues with trucks or motorcycles were not considered. Left-turn queues were not studied.

Table 7. Lane utilization database summary statistics

Movement Type ¹	Number of Lanes in the Lane Group			
	2 Lanes	3 Lanes	4 Lanes	5 Lanes
Left-Turn	1.17	1.28	--	--
Through	1.12	1.26	1.32	1.72
1994 HCM ²	1.05	1.10	1.10	1.10

Notes:

1 - Left-turn movements from both the off-ramp and the arterial. Through movements along the arterial.

2 - Recommended value in the 1994 HCM (3, p. 9-13) for through movements.

--" no data available.

The assembled queue length database contains queue length and reaction time measurement for 122 first-in-queue passenger cars and 1,053 last-in-queue passenger cars. This data was obtained at eight of the twelve study sites. Studies were not conducted at four sites because of the lack of an adequate view of the traffic queue.

The analysis of the queue length model database focused on both the computation of the average lane length occupied by a queued passenger car and the average queued driver starting reaction time. Each characteristic was quantified for the first vehicle in queue and for the "second and subsequent" vehicles in queue. This approach was undertaken because it was evident that the lane length and reaction time differed significantly among the two categories.

Based on an analysis of the queue length data, the average lane length occupied by the first queued passenger car was found to be 5.0 meters. This length includes four to five meters for the actual vehicle and up to one meter between the average vehicle's front bumper and the stop line. The average lane length occupied by the second and subsequent queued vehicles was found to be 7.0 meters per passenger car. This length includes four to five meters for the actual vehicle length and an average inter-vehicle "buffer" distance of two to three meters.

The average reaction time for the first-in-queue drivers was found to be 1.52 seconds. In contrast, the average reaction time for the subsequent queued drivers was found to be 1.06 seconds. The reaction time of the first driver was measured as the time from the start of green to observed initiation of motion. The reaction time of all subsequent drivers was measured as the time from the start of motion of the preceding vehicle to the observed start of motion of the subject vehicle. This trend was expected because the first driver has more of a "surprise" situation (i.e., the signal indication changing from red to green) than the subsequent queued drivers who can look ahead, see that the indication is green, and anticipate their time of departure. As a result, the first drivers should require slightly more reaction time than the subsequent queued drivers.

Weaving. Weaving data reduction required the use of both camera views to track vehicles through the weaving section. The weaving maneuver that was examined in this study was the off-ramp right-turn movement that weaves across the arterial to make a left-turn at the downstream signalized intersection. The two camera recordings were synchronized in time and played back simultaneously to obtain the travel time and stopping location of weaving and non-weaving vehicles. A sampling technique was used to select the tracked vehicles as the lengthy tracking time for each vehicle precluded the collection of a 100-percent sample.

The weaving model database contains entry times for 17,939 vehicles. Of these vehicles, 980 were tracked through the study segment. About one-half of the tracked vehicles (i.e., 421 of 980) were observed to complete a weaving maneuver.

The analysis of the weaving model database focused on the volume and speed of the weaving and non-weaving traffic streams. These data were collected because it was hypothesized that the volume of the two conflicting traffic streams would affect their individual running speeds through the weaving section. It was theorized that these speeds would decrease with increasing volume. The average volumes and speeds through the weaving section for the six study sites are listed in Table 8.

Table 8. Weaving database summary statistics

Variable	Average
Volume	
Total arterial volume entering weaving section	1,409 vph
Arterial lane volume entering weaving section	575 vphpl
Weaving volume (off-ramp right to downstream left)	151 vph
Speed	
Arterial vehicle spot speed at entry to weaving section	14.1 m/s
Arterial vehicle running speed through weaving section	10.6 m/s
Arterial vehicle speed reduction due to weaving activity	3.4 m/s
Weaving vehicle running speed through weaving section	8.0 m/s

As the volumes in Table 8 indicate, the six study sites had relatively high weaving volumes. On average, the weaving vehicles accounted for about one-half of the off-ramp right-turn volume at any one site. The arterial lane volumes were also relatively high such that weaving opportunities were limited during a significant portion of the signal cycle. It should be noted that the off-ramp right-turn movement at three of the sites was signalized (with right-turn on red allowed); the other three were yield-controlled.

Two types of speed statistic were reported for the arterial vehicles. One statistic is the spot speed of the arterial vehicles at a point just upstream of the off-ramp. The second statistic is the running speed of the same arterial vehicles. This latter speed related the distance traveled through the weaving section to the corresponding travel time. The distance and time were measured from the point of entry to the weaving section to the downstream intersection stop line or to the first point of joining the stopped queue associated with the downstream signal, whichever was reached first.

The difference between the arterial spot speed and the running speed is an indicator of a speed reduction in the weaving area due to weaving activity. The average speed reduction at the study sites was 3.4 m/s. This statistic is more useful than the spot or running speeds alone because it eliminates the effect of differing speed limits among the sites. A preliminary examination of this speed difference indicates a strong correlation between it and the total arterial and weaving volumes. Increases in either volume level tended to increase the speed reduction.

As shown in Table 8, the average weaving vehicle speed is 8.0 m/s. This speed tends to be lower than that of the arterial vehicles because the weaving vehicle enters the weaving section at a relatively slow speed due to the ramp control (i.e., signal or yield sign). Some preliminary analysis of this speed indicates that it decreases with increasing arterial lane volume.

2.4 CAPACITY CHARACTERISTICS

This section summarizes the models that can collectively be used to predict the capacity of traffic movements at signalized interchange ramp terminals and other closely-spaced intersections. Specifically, these models predict three important capacity characteristics, they are: saturation flow rate, start-up lost time, and clearance lost time. Details of the development and calibration of these models are provided in Appendix C. It should be noted that the traffic characteristics described in this section reflect passenger car performance only as all heavy vehicles were excluded from the database.

The analysis of the traffic data collected at the twelve study sites followed a two-step process. First, analysis of variance (ANOVA) techniques were used to identify factors influencing the traffic characteristic under examination, to control for differences in sample size, and to account for extraneous differences among otherwise similar sites. The ANOVA was implemented with the Statistical Analysis System's (SAS) (4) general linear model (GLM). All significance tests were conducted at a 95 percent confidence level (i.e., $\alpha = 0.05$). Then, once the influential factors were identified from the ANOVA, both linear and non-linear regression techniques were used to calibrate the data (via these factors) to the proposed model.

2.4.1 Saturation Flow Rate for Through Movements

The saturation flow rate model for through movements was developed from a model of the minimum discharge headway of a stopped queue. Specifically, the saturation flow rate model was derived as the inverse of the minimum discharge headway model. Headways were explicitly

modeled because they represent the most fundamental characteristic describing the efficiency of the discharge process. The minimum discharge headway is defined as the average headway of all headways observed for the fifth and higher queue positions; its reciprocal is saturation flow rate. This method of computing the saturation flow rate is consistent with the procedure described in the 1994 HCM (3, Chapter 9). The following discussion describes the calibration of a minimum discharge headway model for through movements, its algebraic transformation into a saturation flow rate model, and finally, a sensitivity analysis of the transformed model.

Factors Affecting Discharge Headway. A review of the literature on the topic of through movement headways suggests that several site-specific factors exist that can have an effect on the discharge process. For example, Bonneson (6) examined data from a previous study of single-point urban interchanges by Messer *et al* (2) and found that the number of vehicles served per cycle had an effect on the minimum discharge headway. Specifically, he found that the headways observed for each queue position were lower when there were more vehicles queued behind that position. He called this headway compression effect being due to “traffic pressure.”

In this context, traffic pressure is believed to result from the presence of aggressive drivers (e.g., commuters) that are anxious to minimize their travel time in otherwise high-volume conditions. As these drivers are typically traveling during the morning and evening peak traffic periods, they are typically found to be concentrated in the large queues associated with these periods. It should be noted that Stokes, Messer, and Stover (7) found a similar effect of traffic queues on headways; they termed this effect “headway compression.”

Bonneson (6) recommended the following equation for predicting the minimum discharge headway of a single-point urban interchange through movement as a function of traffic pressure:

$$H_{th} = 1.57 + \frac{7.70}{u_s} - 0.0086 v_l' \quad (2)$$

where:

H_{th} = through movement minimum discharge headway, sec/veh;

u_s = speed at saturation flow, m/s; and

v_l' = demand flow rate per lane (i.e., traffic pressure), vpcpl.

The speed in Equation 2 represents the maximum speed drivers tend to reach as they discharge from a traffic queue. In theory, it represents the speed associated with a traffic stream flowing at its saturation flow rate. This speed was found to vary between 12 and 15 m/s in the sites studied by Bonneson (6). One reason offered for this variation was the proximity of some sites to adjacent intersections. Specifically, Bonneson noted that lower speeds were associated with those sites where the distance to the downstream intersection (and its associated queue) was relatively short. This suggests that discharge headways may be lower because of lower discharge speeds that result from the impending downstream stop faced by the discharging drivers.

The HCM (3) describes many additional factors that can affect discharge headway. These factors include: lane width, vehicle classification, local bus frequency, parking activity, approach grade, and area type. To avoid confounding the effect of these factors with those specifically being considered in this study (e.g., distance to back of queue), several steps were taken to avoid or remove the aforementioned factors from the data collected for this project. Specifically, the study sites all had lane widths of about 3.6 meters, approach grades of less than ± 2.5 percent, no local busses, and no parking activity. In addition, all heavy vehicles (i.e., vehicles with more than two axles) and all queued vehicles that followed heavy vehicles were removed from the data base.

Model Calibration. The calibrated minimum discharge headway model for through movements is shown in Equation 3.

$$H_{th} = 1.94 \left(1 + \frac{8.13 (1 - I_s) + 21.8 I_s}{D} \right) (1 - 0.00453 v_l') \quad (3)$$

where:

H_{th} = through movement minimum discharge headway, sec/veh;

D = effective distance to the back of downstream queue (or stop line if no queue) at the start of the subject (or upstream) phase, m;

I_s = indicator variable (1.0 if spillback occurs during phase, 0.0 otherwise); and

v_l' = demand flow rate per lane (i.e., traffic pressure), vpcpl.

The statistics in Table 9 indicate that the calibrated model explains only about four percent of the variability in the headway data. The remaining variability is primarily due to the random (or unexplainable) variability inherent in headway data. Some of the variability is also due to differences among the traffic lanes and sites studied. Nevertheless, the statistics in Table 9 indicate that there is a statistically significant relationship between minimum discharge headway, traffic pressure, and distance to the back of queue. The root mean square error and number of observations can be used to estimate the minimum standard deviation (or precision) of the predicted average minimum discharge headway as ± 0.006 sec/veh.

Table 9. Calibrated through movement minimum discharge headway model

Model Statistics		Value		
R^2 :		0.04		
Root Mean Square Error:		0.56 sec/veh		
Observations:		7,704		
Range of Model Variables				
Variable	Variable Name	Units	Minimum	Maximum
H_{th}	Through movement min. discharge headway	sec/veh	0.61	6.8
v_l'	Demand flow rate per lane (traffic pressure)	vpcpl	5	37
D	Distance to back of downstream queue	meters	35	315

As the coefficient values in Table 9 show, the magnitude of the effect of distance-to-queue is dependent on whether queue spillback occurred during the phase. Phases without spillback had a smaller regression coefficient indicating less sensitivity to distance. In general, the coefficients predict a larger minimum headway for those queues discharging prior to the occurrence of spillback than for those that discharge without spillback ever occurring. Spillback conditions are characterized by the backward progression of a downstream queue into the upstream intersection such that the subject (or upstream) intersection movement is effectively blocked from discharging during some or all of the signal phase.

The "distance to queue" variable D is defined as the distance to the back of the downstream queue at the start of the subject phase. It is measured from the subject movement stop line to the "effective" back of queue. The effective back of queue represents the location of the back of queue if all vehicles on the downstream street segment (moving or stopped) at the start of the phase were joined into a stopped queue. If there are no moving vehicles at the start of the phase, then the effective and actual distance to queue are the same. If there are no vehicles on the downstream segment at the start of the phase, then the effective distance to queue would equal the distance to the through movement stop line at the downstream intersection. The calibrated model indicates that the minimum discharge headway decreases with increasing distance to downstream queue.

Several additional effects were also evaluated during the model calibration process. Specifically, the effect of junction type, phase duration, and downstream signal indication were also evaluated. This latter factor was considered because it was reasoned that drivers might discharge at a more efficient rate if the downstream signal indication was green (as opposed to red), particularly if there was no downstream queue. Based on this additional analysis, it was concluded that these factors did not significantly affect discharge headway after the effects of distance to queue and traffic pressure were removed.

Interpretation of Model Statistics. Three statistics are provided in Table 9 to indicate the quality of fit of the calibrated model. First, the "t-statistic" is provided for each independent variable to test the hypothesis that its regression coefficient equals to zero. When the t-statistic exceeds 1.96, the hypothesis is rejected and it can be concluded that the corresponding variable has a significant effect on the dependent variable. In this situation, there is a 5 percent (or less) chance of this conclusion being in error. In all cases, a graphical examination of the relationship between the dependent and independent variables was used to confirm the significance of the effect.

The second measure of quality of fit is the root mean square error. This statistic represents the standard deviation of the dependent variable. Presumably, the error represented by this statistic is from random sources; however, there could also be some variation due to systematic effects. Knowledge of typical values of the root mean square error for the dependent variable can be a useful gage to assess whether additional systematic error exists in the data. For example, the standard deviation of vehicle headways is rarely reported in the literature to be less than 0.45 sec. Therefore, as the root mean square error of 0.56 reported in Table 9 exceeds 0.45, it is possible that there is some additional systematic error in the data that additional model variables could explain.

The third measure of quality of fit is the R^2 statistic. This statistic represents the portion of the variability explained by the model relative to the total variability in the data. As such, it can range in value from 0.0 to 1.0. In general, larger values of R^2 indicate a good fit; however, the value (or range of values) used to denote a "good" fit is dependent on the amount of random variability in the data. For example, the only way that an R^2 of 1.0 can be achieved is when all of the variability in the data is due systematic sources (i.e., there is no random error) *and* the model properly includes an independent variable for each systematic effect.

The equation for computing R^2 was examined to determine the factors that would influence its magnitude (see Appendix C). This analysis indicated that three factors could have a significant influence on the R^2 value: (1) the amount of variability in the data due to random sources, (2) the variability in the independent variable, and (3) the magnitude of the regression coefficient associated with the independent variable. Based on this analysis, it was concluded that the largest R^2 value possible for the headway data is about 0.38 (as opposed to the 0.8 and above that is traditionally expected). Of course, the regression model would have to include enough variables to account for *all* of the systematic variability in the data in order to obtain this value. It was also concluded that the relatively subtle effect of the independent variables considered in this analysis (as represented by the magnitude of their regression coefficients) limit the R^2 to about the value obtained (i.e., 0.04).

Saturation Flow Rate Model. The calibrated through movement minimum discharge headway model was converted into an equivalent saturation flow rate model. The form of this model was patterned after that used in Chapter 9 of the HCM (3). Specifically, the saturation flow rate for a particular location is estimated as the product of the ideal saturation flow rate and the various site-specific adjustment factors. In this context, the adjustment factors found in this research relate to the effect of distance to the downstream queue at the start of green, spillback occurrence, and traffic pressure. The basic form of the model is:

$$s_l = s_0 \times f_D \times f_v \quad (4)$$

where:

- s_l = saturation flow rate per lane under prevailing conditions, vphgpl;
- s_0 = saturation flow rate per lane under ideal conditions, pcphgpl;
- f_D = adjustment factor for distance to downstream queue at green onset; and
- f_v = adjustment factor for volume level (i.e., traffic pressure).

The ideal saturation flow rate represents the saturation flow rate when not affected by any external environmental factors (i.e., grade), atypical vehicles (i.e., trucks), and constrained geometrics (e.g., less than 3.6-meter lane widths, curved travel path). In this regard, the saturation flow rate would be equal to the ideal rate when all factor effects are optimum for efficient traffic flow and the corresponding adjustment factors are equal to 1.0. Based on this definition, it was determined that an infinite distance-to-queue under non-spillback conditions and a traffic pressure of 15.0 vpcpl were representative of ideal conditions for through movements.

Using the aforementioned definition of ideal conditions and associated parametric values, the resulting ideal saturation flow rate s_0 was derived from Equation 3 as 1,990 pcphgpl. As the precision of this estimate (i.e., about ± 12 pcphgpl) was found to include 2,000 pcphgpl, this latter value is recommended as the ideal saturation flow rate for through movements. The definition of ideal conditions was also used to derive the following adjustment factors:

$$f_D = \begin{cases} \frac{1}{1 + \frac{8.13}{D}} & : \text{ no spillback} \\ \frac{1}{1 + \frac{21.8}{D}} & : \text{ with spillback} \end{cases} \quad (5)$$

$$f_v = \frac{1}{1.07 - 0.00486 v_l'} \quad (6)$$

where:

D = effective distance to the back of downstream queue (or stop line if no queue) at the start of the subject (or upstream) phase, m; and

v_l' = demand flow rate per lane (i.e., traffic pressure), vpcpl.

Sensitivity Analysis. The relationship between distance-to-queue, spillback, traffic pressure, and saturation flow rate are shown in Figure 6. The trends shown indicate that saturation flow rate increases as the distance to the back of queue becomes longer. They also indicate that phases that incur spillback have a lower saturation flow rate, for the same distance to queue, than phases that do not incur spillback. Finally, saturation flow rate is shown to increase with increasing traffic pressure.

2.4.2 Saturation Flow Rate for Left-Turn Movements

The saturation flow rate model for left-turn movements was developed from a model of the minimum discharge headway of left-turn vehicles. The following discussion describes the calibration of a minimum discharge headway model, its algebraic transformation into a saturation flow rate model, and finally, a sensitivity analysis of the transformed model.

With one exception, the left-turn movements included in this study represent left-turns at interchange ramp terminals. The one exception was a left-turn movement at an adjacent signalized intersection. Of the two types of left-turn movements at ramp terminals (i.e., off-ramp and arterial), the majority of the data were collected for the off-ramp left-turn movement. Nevertheless, it is believed that the factors identified in this section are sufficiently general that they are applicable to off-ramp and arterial left-turn movements at interchanges and to left-turn movements at adjacent intersections.

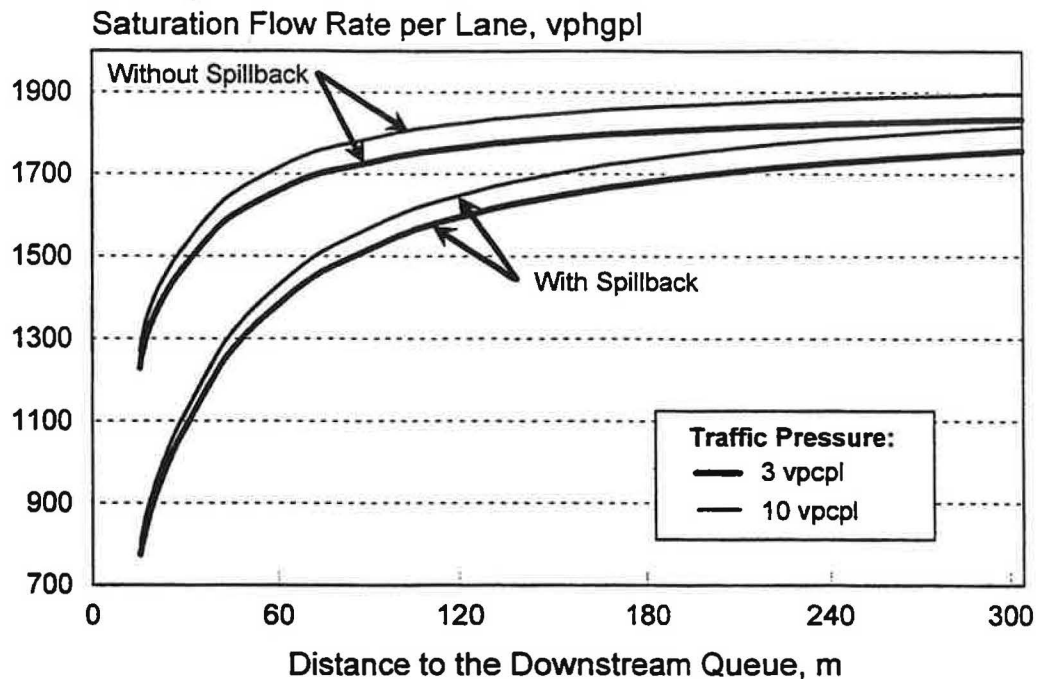


Figure 6. Effect of distance-to-queue, spillback occurrence, and traffic pressure on through movement saturation flow rate.

The left-turn movements studied rarely, if ever, experienced queue spillback during the study periods due to the nature of the signal phase coordination between the two interchange ramp terminals. Hence, in contrast to the through movements studied, the variability in left-turn headways among sites cannot be explained by differences in the distance to the downstream queue. This restriction is a characteristic of the twelve sites studied; certainly, left-turn movements can be affected by downstream queuing conditions at other sites. In fact, it is likely that the effect will be very similar to that found for the through movements.

Factors Affecting Discharge Headway. A review of the literature on the topic of left-turn headways suggests that several site-specific factors exist that can have an effect on the left-turn discharge process. For example, Kimber *et al* (8) measured saturation flows on curves with radii ranging from 6 to 35 meters and developed an equation for predicting saturation flow rate as a function of turn radius. An equivalent relationship, as it relates to minimum discharge headway, is:

$$H_u = 1.73 + \frac{2.60}{R} \quad (7)$$

where:

- H_u = left-turn movement minimum discharge headway, sec/veh; and
- R = radius of curvature of the left-turn travel path (at center of path), m.

In a previous study of headways at intersections and single-point urban interchanges, Bonneson (6), found an effect of radius on headway consistent with that found by Kimber (8). The range of radii included in this study was 18 to 84 meters. Bonneson also found that the number of vehicles served per cycle had an effect on left-turn headway. This effect was referred to as "traffic pressure" in a preceding section. Bonneson recommended the following equation for predicting the minimum discharge headway of a left-turn movement as a function of radius and traffic pressure:

$$H_{lt} = 1.58 + \frac{0.830}{R^{0.245}} - 0.0121 v_l' \quad (8)$$

where:

v_l' = demand flow rate per lane (i.e., traffic pressure), vpcpl.

As discussed in a previous section, the HCM (3) describes many additional factors that can affect discharge headway (e.g., lane width, vehicle classification, etc.). To avoid confounding the effect of these factors with those specifically being considered in this study (e.g., turn radius), the study sites were selected to have as near ideal conditions as possible for all non-relevant factors.

Model Calibration. The field data were used to calibrate the left-turn movement minimum discharge headway model, as shown in Equation 9. Statistics describing the model's predictive performance and the range of each model variable are provided in Table 10.

$$H_{lt} = 1.55 \left(1 + \frac{1.71}{R} \right) (1 - 0.00630 v_l') (1 + 0.868 t_g) \quad (9)$$

with

$$t_g = \left(\frac{g}{C} \right) I_g + 0.27 (1 - I_g) \quad (10)$$

where:

H_{lt} = left-turn movement minimum discharge headway, sec/veh;

R = radius of curvature of the left-turn travel path (at center of path), m;

v_l' = demand flow rate per lane (i.e., traffic pressure), vpcpl;

g = effective green time where platoon motion (flow) can occur, sec;

C = cycle length, sec;

t_g = signalization variable ($0.0 < t_g < g_x$);

I_g = indicator variable (1.0 if $g/C < g_x$, 0.0 otherwise); and

g_x = maximum g/C ratio (larger g/C ratios have no additional effect on headway).

As the statistics in Table 10 indicate, the calibrated model explains only about five percent of the variability in the headway data. The remaining 95 percent of the variability is primarily due to the inherent randomness in headway data (as described in a preceding section dealing with the saturation flow rate model for through movements). Nevertheless, the statistics in Table 10 indicate

recommended as the ideal saturation flow rate for a left-turn movements (although it is recognized that the radius of the travel path is assumed to be infinitely long). The definition of ideal conditions was also used to derive the following adjustment factors:

$$f_R = \frac{1}{1 + \frac{1.71}{R}} \quad (12)$$

$$f_v = \frac{1}{1.07 - 0.00672 v_l'} \quad (13)$$

$$f_{g/C} = \frac{1}{0.810 + 0.703 t_g} \quad (14)$$

$$t_g = \left(\frac{g}{C}\right) I_g + 0.27 (1 - I_g) \quad (15)$$

where:

R = radius of curvature of the left-turn travel path (at center of path), m;

v_l' = demand flow rate per lane (i.e., traffic pressure), vpcpl;

g = effective green time where platoon motion (flow) can occur, sec;

C = cycle length, sec;

t_g = signalization variable ($0.0 < t_g < 0.27$); and

I_g = indicator variable (1.0 if $g/C < 0.27$, 0.0 otherwise).

Sensitivity Analysis. The calibrated model was used to examine the effect of radius, traffic pressure, and g/C ratio on the saturation flow rate of left-turn movements. This effect is shown in Figure 7 for the respective characteristics. The trend lines shown in this figure reflect left-turn volumes of 3 and 10 vpcpl and g/C ratios of 0.16 and 0.27. These ranges were selected to be inclusive of about 90 percent of the observations in the database. In general, the model has a trend of increasing saturation flow rate with radius. The range of traffic volume (i.e., pressure) considered makes a difference of about 100 pcphgpl in saturation flow rate. In contrast, the g/C ratio has almost twice the effect as traffic pressure (i.e., a change of about 160 pcphgpl). Of course, g/C ratio has no effect when the ratio increases beyond 0.27

a statistically significant relationship between minimum discharge headway, radius, traffic pressure, and g/C ratio. The minimum precision of the average headway estimate is about ± 0.007 sec/veh.

Table 10. Calibrated left-turn movement minimum discharge headway model

Model Statistics		Value		
R^2 :		0.05		
Root Mean Square Error:		0.44 sec/veh		
Observations:		4,153		
Range of Model Variables				
Variable	Variable Name	Units	Minimum	Maximum
H_h	Left-turn movement min. discharge headway	sec/veh	0.83	3.5
R	Radius of curvature of travel path	meters	15	98
v_i'	Demand flow rate per lane (traffic pressure)	vpcpl	5	26
g/C	Effective green to cycle length ratio	na	0.06	0.55

Saturation Flow Rate Model. The calibrated left-turn movement minimum discharge headway model was converted into an equivalent saturation flow rate model. The form of this model was patterned after that used in Chapter 9 of the HCM (3). Specifically, the saturation flow rate for a particular location is estimated as the product of the ideal saturation flow rate and the various site-specific adjustment factors. In this context, the adjustment factors found in this research relate to the effect of traffic pressure, signal timing, and turn radius. The basic form of the model is:

$$s_l = s_0 \times f_R \times f_v \times f_{g/C} \quad (11)$$

where:

- s_l = saturation flow rate per lane under prevailing conditions, vphgpl;
- s_0 = saturation flow rate per lane under ideal conditions, pcphgpl;
- f_R = adjustment factor for the radius of the travel path;
- f_v = adjustment factor for volume level (i.e., traffic pressure); and
- $f_{g/C}$ = adjustment factor for signal timing.

The ideal saturation flow rate represents the saturation flow rate when not affected by any external environmental factors, atypical vehicles, and constrained geometrics. In this regard, the saturation flow rate would be equal to the ideal rate when all factor effects are optimum for efficient traffic flow and the corresponding adjustment factors are equal to 1.0. Based on this definition, it was determined that an infinite radius, a traffic pressure of 10.0 vpcpl, and a g/C ratio greater than 0.27 were representative of ideal conditions for left-turn movements.

Using the aforementioned definition of ideal conditions and associated parametric values, the resulting ideal saturation flow rate s_0 was derived from Equation 9 as 2,010 pcphgpl. As the precision of this estimate (i.e., about ± 12 pcphgpl) includes 2,000 pcphgpl, this latter value is

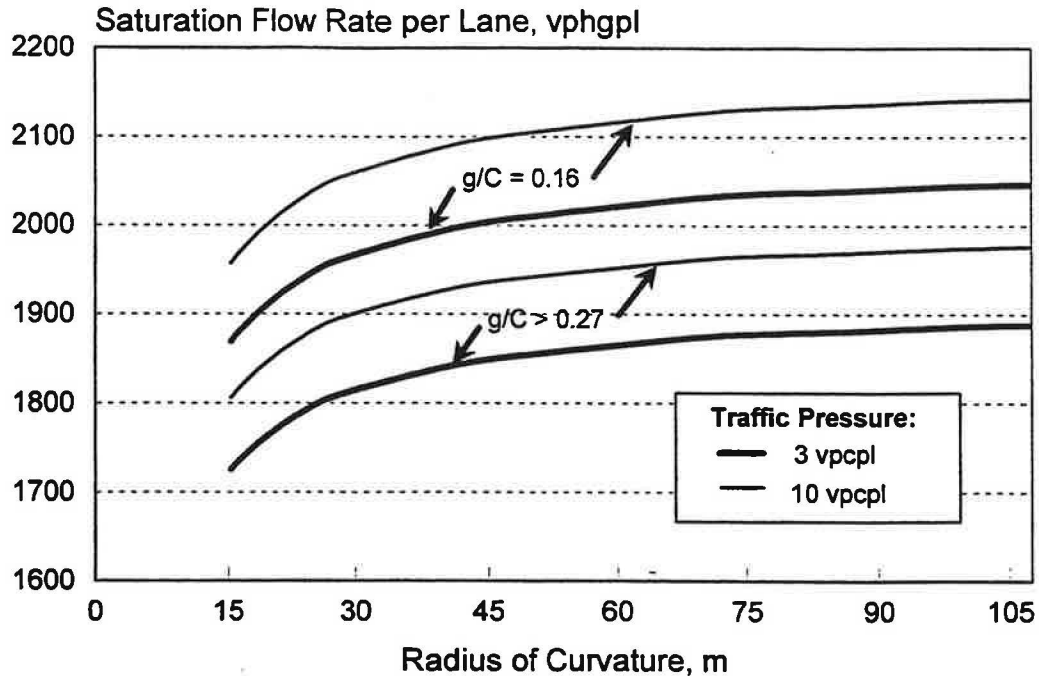


Figure 7. Effect of traffic pressure, signal timing, and radius on left-turn movement saturation flow rate.

2.4.3 Start-Up Lost Time for Through Movements

The time “lost” at the start of a phase stems from the fact that the headways of the vehicles in the first few queue positions are larger than those of vehicles in the higher queue positions. The headways of these first few queued vehicles are large because of the acceleration they are undergoing as they cross the stop line. Once the vehicles in these positions near the “desired” discharge speed, their headways converge to the minimum discharge headway. Thus, factors that influence this speed (e.g., distance to queue, radius, lane width, etc.) also affect minimum discharge headway and saturation flow rate. As a result, there is an inherent relationship between saturation flow rate and start-up lost time. This section describes the calibration of the start-up lost time model for through movements. As with the saturation flow rate model, a start-up lost time model is developed for the left-turn movement in a subsequent section.

Model Calibration. An examination of the start-up lost time and saturation flow rate data indicated that there was a strong linear relationship between the two variables. Based on this examination, a linear model form was calibrated using the field data. The calibrated model for through traffic is:

$$l_s = -4.64 + 0.00373 s_l \quad (16)$$

where:

l_s = start-up lost time for through traffic, sec; and

s_l = saturation flow rate per lane for through traffic under prevailing conditions; vphgpl.

Statistics describing the model's predictive performance and the range of each model variable are presented in Table 11. As the statistics in this table indicate, the calibrated model explains about 41 percent of the variability in the data which is indicative of a strong correlation between start-up lost time and saturation flow rate. Based on the root mean square error and number of observations, the minimum precision of the average start-up lost time estimate is about ± 0.02 sec.

Table 11. Calibrated through movement start-up lost time model

Model Statistics		Value		
R^2 :		0.41		
Root Mean Square Error:		1.07 sec		
Observations:		1,927		
Range of Model Variables				
Variable	Variable Name	Units	Minimum	Maximum
l_s	Start-up lost time	sec	-3.1	8.0
s_f	Saturation flow rate per lane	vphgpl	1,257	3,326

Sensitivity Analysis. The calibrated model of start-up lost time was used to examine the sensitivity of through movement start-up lost time to saturation flow rate. This relationship is shown in Figure 8. As this figure indicates, the start-up lost times for saturation flow rates of 1,800 and 1,900 vphgpl are about 2.0 and 2.5 seconds, respectively. These values are slightly larger than the 1.0 to 2.0 seconds recommended in Chapter 2 of the HCM (3).

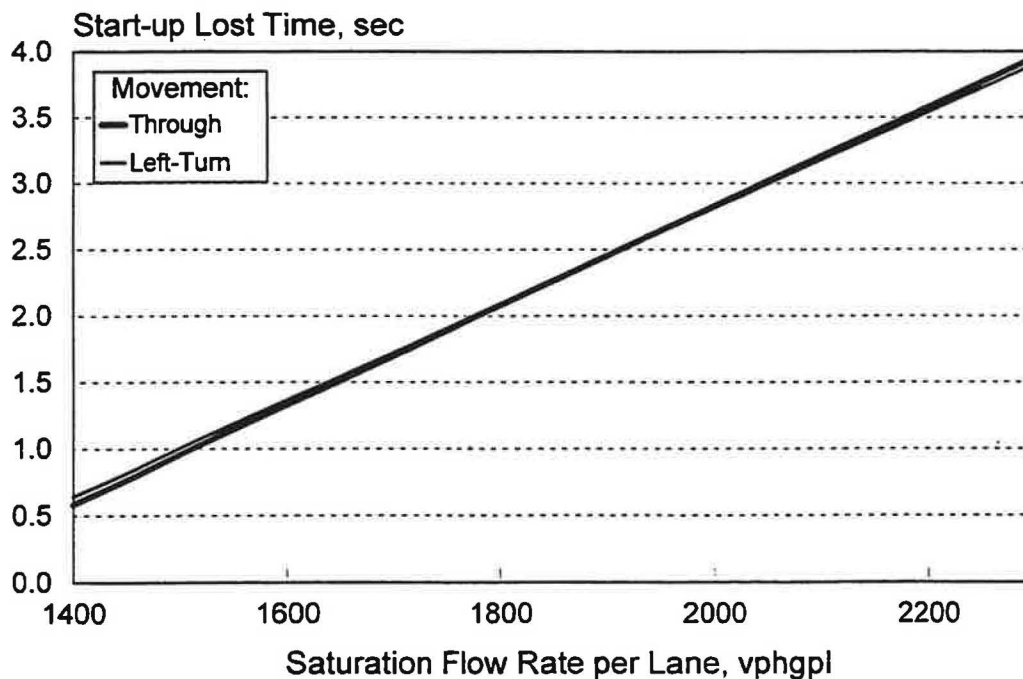


Figure 8. Expected through movement start-up lost time as a function of saturation flow rate.

2.4.5 Clearance Lost Time

When the yellow indication is presented at the end of a phase, drivers close to the intersection generally continue on through the intersection because stopping would be impossible or, at least, very uncomfortable. Thus, these "clearing" drivers tend to use the first few seconds of the yellow interval. The remaining portion of the yellow interval that is not used by the average clearing driver plus the red clearance interval is defined as the clearance lost time. Based on this definition, the following equation can be used to compute clearance lost time:

$$l_e = Y + R_c - g_Y \quad (18)$$

where:

- l_e = clearance lost time, sec;
- Y = yellow interval, sec;
- R_c = red clearance interval, sec; and
- g_Y = effective green extension into the yellow interval, sec.

It was hypothesized that extent of driver encroachment into the yellow interval could be affected by the clearing vehicle's speed, width of the intersection, and delay if not clearing. Thus, model calibration focused on defining a relationship between green extension, speed, intersection width, and signal timing (as a surrogate for delay). Green extension was quantified as the average duration of the yellow interval used during phases where it was observed to be used to some degree. Phases where the yellow interval was not used were excluded from the analysis because the reason for this lack of use was not determinable from the data. In general, the yellow interval was used in about 27 percent of the phases, although this frequency varied widely among the twelve study sites.

Model Calibration. After an exploratory analysis of variance, several influential factors were identified that had a significant effect on end use. These factors were related to the duration of green extension using the following linear model form:

$$g_Y = 1.48 + 0.014 SL + 6.40 (X - 0.88) I_X \quad (19)$$

where:

- g_Y = effective green extension into the yellow interval, sec;
- SL = approach speed limit, km/h;
- X = volume-to-capacity ratio for the lane group; and
- I_X = indicator variable (1.0 if $X > b_3$, 0.0 otherwise).

The analysis considered 1,044 signal phases with observed driver use of the yellow (and in some cases, red clearance) interval. These phases were observed at twelve interchange ramp terminals and at twelve intersection approaches. The green extension data used in this analysis represent observations made for both left-turn and through movements. The left-turns at the interchanges were made from either the off-ramp or the arterial. Statistics describing the model's predictive performance and the range of each model variable are presented in Table 13.

2.4.4 Start-up Lost Time for Left-Turn Movements

Model Calibration. An examination of the saturation flow rate and start-up lost time data for the left-turn movements studied indicated that a linear relationship existed between the two characteristics, similar to that found for the through movements. Thus, a linear model form was also calibrated for the left-turn movement data. The calibrated model for left turns is:

$$l_s = -4.43 + 0.00362 s_l \quad (17)$$

where:

l_s = start-up lost time for left turn traffic, sec; and

s_l = saturation flow rate per lane for left turns under prevailing conditions, vphgpl.

Statistics describing the model's predictive performance and the range of each model variable are presented in Table 12. As the statistics in this table indicate, the calibrated model explains about 34 percent of the variability in the headway data which is indicative of a good correlation between start-up lost time and saturation flow rate. The minimum precision of the average start-up lost time estimate is about ± 0.04 sec.

Table 12. Calibrated left-turn movement start-up lost time model

Model Statistics		Value		
R^2 :		0.34		
Root Mean Square Error:		1.14 sec		
Observations:		714		
Range of Model Variables				
Variable	Variable Name	Units	Minimum	Maximum
l_s	Start-up lost time	sec	-1.5	7.8
s_l	Saturation flow rate per lane	pcphgpl	1,339	2,770

Sensitivity Analysis. The calibrated model of start-up lost time was used to examine the sensitivity of left-turn movement start-up lost time to saturation flow rate. This relationship is also shown in Figure 8. As this figure indicates, the start-up lost times for saturation flow rates of 1,800 and 1,900 pcphgpl are about 2.0 and 2.5 seconds, respectively. These values are slightly larger than the 1.0 to 2.0 seconds recommended in Chapter 2 of the HCM (3).

Figure 8 facilitates comparison of the relationships between start-up lost time and saturation flow rate for left and through movements. The comparison indicates that there is very little difference between the two movements in this regard. Therefore, it appears reasonable to conclude that the effect of saturation flow rate on start-up lost time is independent of movement type.

2.4.5 Clearance Lost Time

When the yellow indication is presented at the end of a phase, drivers close to the intersection generally continue on through the intersection because stopping would be impossible or, at least, very uncomfortable. Thus, these "clearing" drivers tend to use the first few seconds of the yellow interval. The remaining portion of the yellow interval that is not used by the average clearing driver plus the red clearance interval is defined as the clearance lost time. Based on this definition, the following equation can be used to compute clearance lost time:

$$l_c = Y + R_c - g_Y \quad (18)$$

where:

l_c = clearance lost time, sec;

Y = yellow interval, sec;

R_c = red clearance interval, sec; and

g_Y = effective green extension into the yellow interval, sec.

It was hypothesized that extent of driver encroachment into the yellow interval could be affected by the clearing vehicle's speed, width of the intersection, and delay if not clearing. Thus, model calibration focused on defining a relationship between green extension, speed, intersection width, and signal timing (as a surrogate for delay). Green extension was quantified as the average duration of the yellow interval used during phases where it was observed to be used to some degree. Phases where the yellow interval was not used were excluded from the analysis because the reason for this lack of use was not determinable from the data. In general, the yellow interval was used in about 27 percent of the phases, although this frequency varied widely among the twelve study sites.

Model Calibration. After an exploratory analysis of variance, several influential factors were identified that had a significant effect on end use. These factors were related to the duration of green extension using the following linear model form:

$$g_Y = 1.48 + 0.014 SL + 6.40 (X - 0.88) I_X \quad (19)$$

where:

g_Y = effective green extension into the yellow interval, sec;

SL = approach speed limit, km/h;

X = volume-to-capacity ratio for the lane group; and

I_X = indicator variable (1.0 if $X > b$, 0.0 otherwise).

The analysis considered 1,044 signal phases with observed driver use of the yellow (and in some cases, red clearance) interval. These phases were observed at twelve interchange ramp terminals and at twelve intersection approaches. The green extension data used in this analysis represent observations made for both left-turn and through movements. The left-turns at the interchanges were made from either the off-ramp or the arterial. Statistics describing the model's predictive performance and the range of each model variable are presented in Table 13.

Table 13. Calibrated green extension model

Model Statistics		Value		
R^2 :		0.11		
Root Mean Square Error:		1.33 seconds		
Observations:		1,044		
Range of Model Variables				
Variable	Variable Name	Units	Minimum	Maximum
g_r	Effective green extension into the yellow interval	sec	0.02	7.3
SL	Approach speed limit	km/h	56	72
X_i	Volume-to-capacity ratio in lane i	na	0.08	1.3

The R^2 statistic in Table 13 indicates that the calibrated model accounts for eleven percent of the variability in the green extension data. The remaining variability is likely due to random sources; although, some of it may be due to differences among the study sites (that was not explained by speed limit and volume-to-capacity ratio). Nevertheless, it is believed that the calibrated model provides a relatively good fit to the data and that it can be used to predict the average green extension with reasonable precision (i.e., a minimum of ± 0.04 sec.).

Sensitivity Analysis. The calibrated model was used to examine the effect of speed limit and volume-to-capacity ratio on clearance lost time. Prior to conducting this examination, it was necessary to define the duration of the yellow and red clearance intervals. Recognizing that the yellow interval duration is often dependent on the approach speed and that there are a wide range of methods being used to determine yellow interval duration, it was decided to set the yellow interval equal to 0.062 times the approach speed (i.e., $Y = 0.062 * \text{speed limit (km/h)}$). This approach yields values generally consistent with other methods or policies. The red clearance interval was established as 1.0 second. Using these values, the clearance lost time was computed for a range of speed limits and volume-to-capacity ratios. The results of this examination are shown in Figure 9.

As Figure 9 illustrates, clearance lost time increases with approach speed limit and decreases with increasing volume-to-capacity ratio. In general, it ranges from 1.0 to 3.0 seconds for typical speed limits and uncongested conditions. This range compares with the 1.2 to 2.8-second range for clearance lost time suggested in Chapter 2 of the HCM (3). Clearance lost time increases with speed because of a corresponding increase in the yellow interval; however, it should be noted that this effect is offset to some degree by the increase in green extension associated with higher speeds.

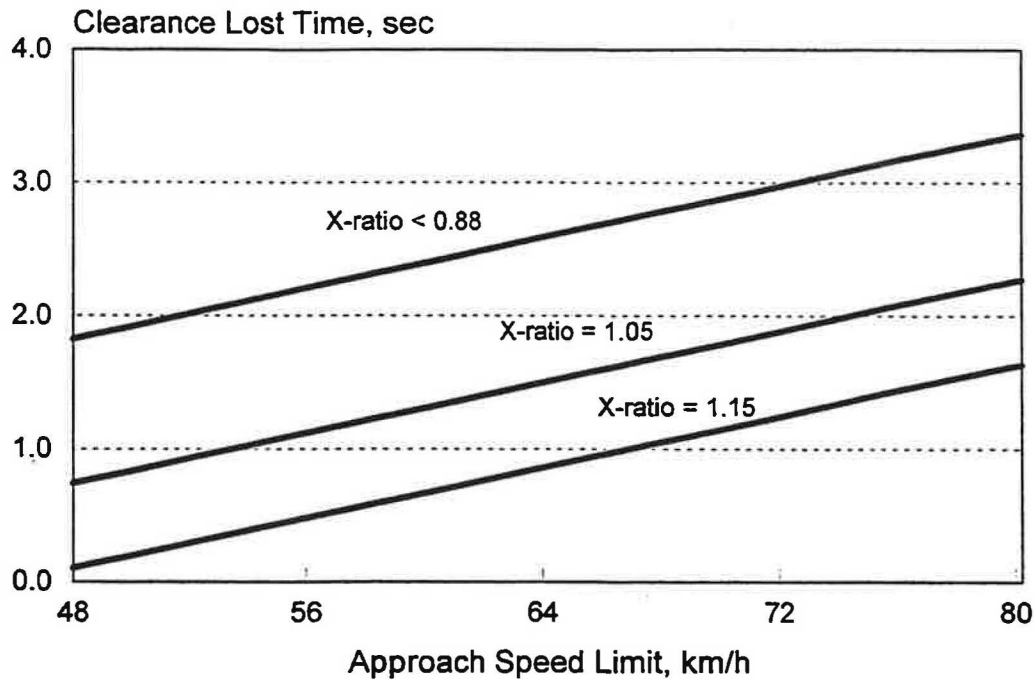


Figure 9. Effect of approach speed limit on clearance lost time.

2.4.6 Lane Utilization

The quality of service provided by a signalized intersection is highly dependent on the volume-to-capacity ratio of the intersection and its associated signal phases. One consideration in determining demand volume for the phase is the distribution of traffic among the lanes it serves. More specifically, these would be the lanes available to a "lane group," as defined in Chapter 9 of the HCM (3). Obviously, if the traffic for a given lane group is concentrated in only one of the several available lanes, then the phase duration would need to be long enough to serve traffic in this one lane. Alternatively, if the phase duration is not increased, then the capacity of the lane group is effectively reduced by the degree of underutilization of its lower volume lanes.

This section describes the calibration and examination of a model for predicting the lane utilization factor. This factor is traditionally used in a capacity analysis to adjust the lane group volume such that the resulting, adjusted volume reflects the traffic demand in the lane with the highest demand. Equation 1 was used to compute the lane utilization factor for each signal cycle at each site and lane group studied.

Model Development. The lane utilization model developed in this research is based on a quantitative description of the two problems previously described: (1) drivers not distributing themselves as evenly as possible, and (2) drivers prepositioning for a downstream turn. The first problem is more fundamental in nature and deals specifically with the random nature of vehicle

arrivals per cycle and the extent that drivers collectively can and will distribute themselves among available traffic lanes. Unbalanced lane use stemming from this problem would be found in any multi-lane lane group on an intersection approach. The second problem is of a site-specific nature as it relates to the effects of downstream turn movements on a driver's lane choice at the upstream intersection. This problem would not necessarily be found in all multi-lane lane groups.

A theoretic model that describes both of the aforementioned lane use problems was developed for this research and calibrated with field data (see Appendix C). This model is applicable to interchanges, adjacent intersections, and other intersections where prepositioning may occur. The form of this model is:

$$U = [1 + 0.423 (\frac{N-1}{2v'}) + 0.433 N \sqrt{\frac{N-1}{2v'}}] (1 - I_p) + [1.05 \frac{\text{Max}(v'_{dl}, v'_{dr}) N}{v'}] I_p \quad (20)$$

where:

- U = lane utilization factor for the lane group;
- v'_i = demand flow rate in lane i , $i = 1, 2, \dots, N$, vpcpl;
- v'_{dl} = flow rate in the lane group that will be turning left at the downstream intersection, vpc;
- v'_{dr} = flow rate in the lane group that will be turning right at the downstream intersection, vpc;
- N = number of lanes in the lane group, lanes;
- I_p = indicator variable (1.0 if $\text{Max}(v'_{dl}, v'_{dr})/v' > 1/N$, 0.0 otherwise); and
- $\text{Max}(v'_{dl}, v'_{dr})$ = larger of v'_{dl} and v'_{dr} .

Statistics describing the model's predictive performance and the range of each model variable are provided in Table 14. As the statistics provided in this table indicate, the calibrated model provides a reasonably good fit to the data. The minimum precision of the average lane utilization factor estimate is about ± 0.01 , based on the root mean square error and the number of observations. The R^2 of 0.18 is lower than values traditionally expected; however, it must be remembered that there is considerable random variability in the lane utilization factor. This variability stems from the fact that two random variables (i.e., v'_{max} and v') are being used in the computation of the lane utilization factor. Thus, the variability in this factor represents the combined variability of the two underlying random variables.

Sensitivity Analysis. Figure 10 illustrates the relationship between lane utilization, lane group flow rate, and number-of-lanes, as predicted by the calibrated lane utilization model. In general, the lane utilization factor increases with number of lanes and decreases with increasing volume. It should be noted that the predicted lane utilization factors exceed the values recommended by the HCM (3) (i.e., 1.05 at two through lanes, 1.10 at three through lanes).

Also shown in Figure 10 are the lane utilization factors predicted by a model developed by Fambro *et al* (9, 10). This model is coined the "TTI Model" in reference to the authors' affiliation. It was calibrated to ten traffic movements at nine signalized intersections. In general, the TTI Model

Table 14. Calibrated lane utilization model

Model Statistics		Value		
R^2 :		0.18		
Root Mean Square Error:		0.11		
Observations:		97		
Range of Model Variables				
Variable	Variable Definition	Units	Minimum	Maximum
U	Lane utilization factor	na	1.0	1.56
v'_{max}	Maximum lane flow rate in any lane	vpcpl	5.8	40.3
N	Number of lanes in the lane group	na	2	4
v'_{dl}	No. of vehicles turning left downstream	vpc	0	23.3
v'_{dr}	No. of vehicles turning right downstream	vpc	0	16.1
v'	Demand flow rate for the lane group	vpc	12.4	69.7

predictions compare favorably with those of the calibrated lane utilization model; however, the agreement is best at the higher flow rates. This agreement is partly due to the fact that both data bases had the majority of their observations in this higher range of flow rates. This agreement suggests that the calibrated lane utilization model may be applicable to all signalized intersections.

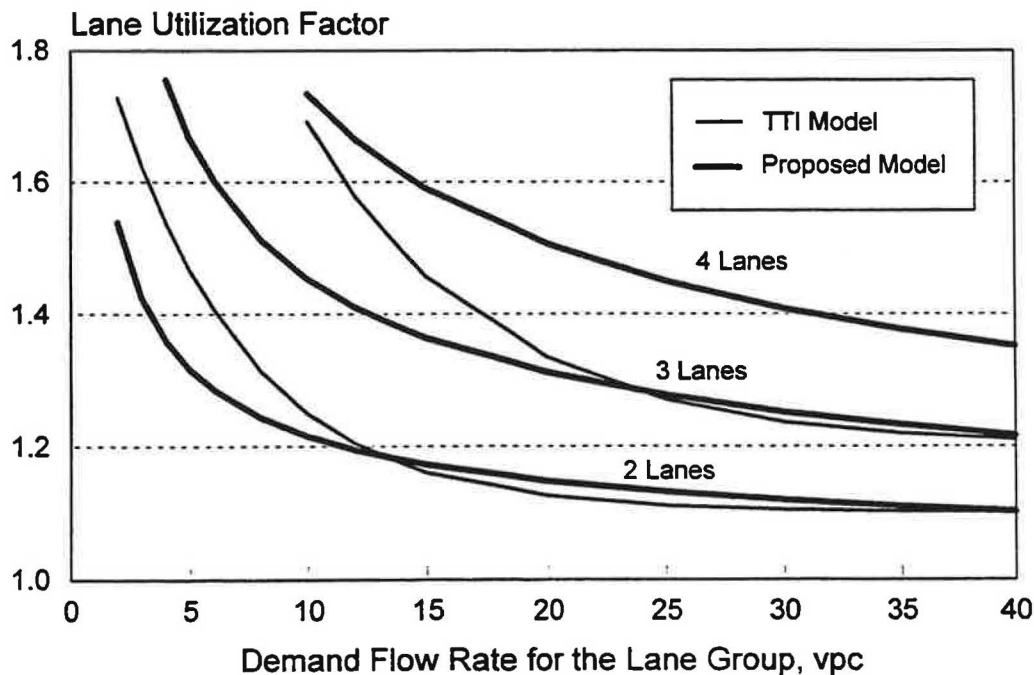


Figure 10. Effect of flow rate and number of lanes on the lane utilization factor.

2.5 TRAFFIC CONTROL, SPILLBACK AND PERFORMANCE

The research problem statement of NCHRP 3-47 identified queue spillback as being the primary traffic operational "problem" that confronts traffic engineers trying to improve traffic flow at high-volume signalized interchanges. Moreover, the national survey of traffic engineers conducted in the initial stages of this research also confirmed that queue spillback and related capacity issues were believed to be major operational problems observed at signalized interchanges.

This research program has produced some new and important operational findings regarding spillback that might help to better explain the operational dynamics occurring during oversaturation conditions. An important finding is that traffic response to control inputs during undersaturation conditions is basically the reverse sensitivity to what occurs during oversaturation conditions. Traffic control plans that are designed to provide priority arterial flow during undersaturation may not provide the same relative priority and expected performance during oversaturation. A series of microscopic traffic simulation studies are presented which illustrate the response sensitivities of traffic signal systems observed for throughput (arterial volume) and for traffic delay experienced on internal links along the arterial and on external approach movements feeding the arterial.

2.5.1 Experimental Testbed

The TRAF-NETSIM simulation model (5) was employed to study closely spaced intersections often found at signalized interchanges. NETSIM was chosen because of its capability to simulate congested traffic conditions, including spillback, and its supporting graphics for visualizing the experimental process. The four-intersection study testbed is depicted in Figure 11. The signalized ramp terminals of a representative diamond interchange are noted as being intersections j and k in Figure 11. The outer two signalized intersections along the crossing arterial are noted as being intersections i and l .

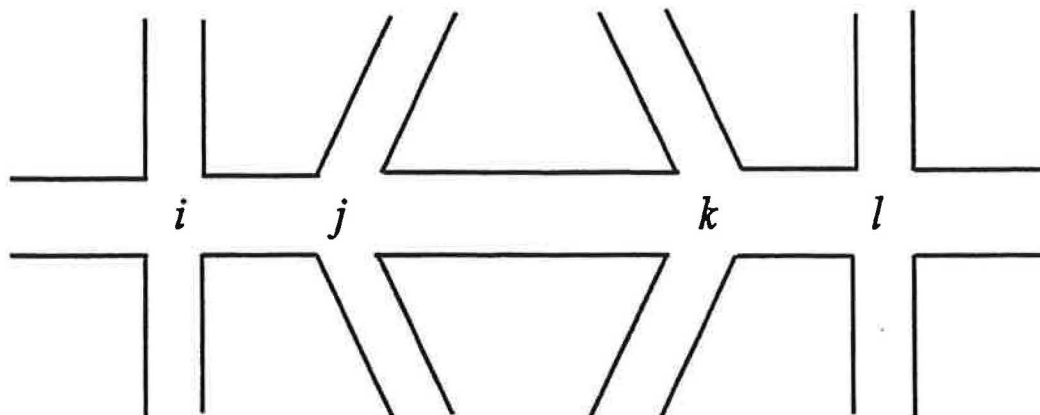


Figure 11. Experimental testbed for signalized interchange with crossing arterial.

Experimentation with the traffic performance on individual links $j-k$, such as between the two signals of the interchange, demonstrate how the traffic performance varies with control inputs. Two traffic measures examined were (1) arterial throughput (the smaller of the traffic demand or service capacity), and (2) traffic delay experienced using the link (due to the traffic signal). Only research issues were examined and demonstrated. Such topics as showing how capacity and delay vary with cycle time were accepted as known technology (3) and were not studied herein. See Appendix A for further details on existing procedures and recent research on the subject (11).

An arterial street is a connected chain of links such that link $i-j$ is connected to link $j-k$, etc. For traffic flow in the $i-j-k$ direction, traffic signal j is defined as the downstream node of link $i-j$ and the upstream node of link $j-k$. Thus, precedence and dependency relationships exist between links and may be operative at any time conditions warrant. Determining when conditions are critical is a necessary part of any traffic analysis methodology.

2.5.2 Equation of Continuity

The well-known equation of continuity serves as the fundamental theory of traffic flow on all types of traffic links, including impeded and congested operations. The equation of continuity, also known as the input-output model, is

$$n(L, t) = n_0 + \sum v_m t - \sum c_m t \quad n(L, t) \leq n_{\max} \quad (21)$$

where:

$n(L, t)$	=	number of vehicles operating on the link of length L at time t , vehicles;
n_0	=	number of vehicles operating on the link at the start of period, vehicles;
v_m	=	total arrival flow into head of link destined to movement m , vph;
c_m	=	output flow \leq capacity of link serving movement m , vph;
k_q	=	stopline queue storage density of 143 vpkmpl, or a storage spacing of 7.0 m/veh (23 ft/veh); and
n_{\max}	=	maximum number of vehicles that can store on link, vehicles,

The equation of continuity can be used to examine the boundaries of flow and dependencies for a wide range of operations, including undersaturated and oversaturated conditions.

2.5.3 Undersaturated Conditions

Undersaturated traffic conditions are those wherein the traffic demand on an approach to a signal phase is less than the operational capacity of the signal phase that serves it. All other lane groups that may interact with this traffic movement must also be undersaturated. Principal factors which affect whether a phase is undersaturated include:

1. the arrival traffic demand;
2. the nominal phase capacity serving the demand; and
3. any impediments to saturation flow of the subject phase due to spillback.

Undersaturation is normally thought of as being a deterministic condition where the existing traffic demand is less than the nominal phase capacity (3). Moreover, undersaturated conditions may also be thought of as not having any flow dependency problems. However, this research has shown that closely spaced signalized intersections, whose traffic signals are poorly timed, can have both demand starvation on the link and still cause flow blockages on the next upstream link due to queue spillback, even during nominally undersaturated conditions. Demand starvation results in wasted green at the downstream signal. The flow blockages due to queue spillback can then cause oversaturation to occur. Results of simulation experiments will be presented that demonstrate these findings. Three examples will be used to illustrate the range of traffic situations that might be encountered in the field and the resulting operational responses that might be observed as existing signal timings are evaluated and possible timing changes envisioned.

Arterial Dominance. This traffic pattern is 100% arterial through traffic. Queue spillback, blockage, and other impediments to flow can occur on short links that have poorly timed traffic signals, even during undersaturated conditions along the arterial. A water transportation example illustrates the worst-case signal timing situation. Consider traffic operations along the Panama Canal. There, one objective is to minimize the water flow along the links of the canal (arterial) per passage of a ship (signal cycle). This minimum flow is accomplished by using very short links and never having both upstream and downstream gates (greens) open simultaneously. Thus, using the equation of continuity, the maximum flow, $f_c(L, C)$ (in vpspl), that can occur on a link, of length (meters), where no simultaneous input-output flows occur during maximum storage conditions is:

$$f_c(L, C) = \frac{N_{\max}}{C} = \frac{k_q L}{C} \approx \frac{L}{7C} \quad (22)$$

The minimum (critical) upstream or downstream green, g_c , required to fill or dissipate, respectively, the critical flow without simultaneous signal coordination is:

$$g_c = \frac{k_q L}{s} \approx \frac{L}{3.5} \quad (23)$$

Phases longer than g_c run the risk of not being "effectively" green unless simultaneous input-output flows occur during critical storage conditions. Equation 22 can also be solved for C_c , given that $f_c(L, C) = v$, the arrival volume, to determine the critical cycle, C_c , that might produce upstream blockage and/or downstream green starvation of the output phase. The resulting equation is

$$C_c = \frac{k_q L}{v} \quad (24)$$

These concepts are illustrated in the following NETSIM traffic simulations. Two connected 100-meter links $i-j$ and $j-k$ are assumed to have only arterial through traffic $v = 1400$ vph on two lanes which are served by signal phases having a cycle of 120 seconds, effective green times $g = 49$ seconds, and $s = 1900$ vphpl, providing green ratios of $g/C = 0.41$ and a nominal phase capacity $c = 1558$ vph, which produce a nominal v/c ratio of 0.9 at each signal. Equations 22-24 predict a critical flow of 428 vphpl, a critical green of about 27 seconds, and a critical cycle of 73 seconds for each 100-meter link. The short links are susceptible both to having upstream input blockage due to queue spillback and also to simultaneously (within the same cycle, but later in the cycle) experiencing demand starvation. These operational problems are possible even though all signals are nominally undersaturated when no spillback/blockage occurs. Continuing with the NETSIM study, the signal offset for link $i-j$ was fixed at 10 seconds to provide good arterial progression and minimal delay, and it was not varied. The offset for link $j-k$ was varied over the entire 120-second cycle. The results from several studies of arterial throughput and delay follow.

Figure 12 shows that throughput flow problems are occurring on the short downstream link $j-k$ primarily due to "demand starvation." For offsets of 40 seconds or more, the throughput volume drops 39%, from a nominal two-lane flow of about 1400 vph to a flow rate of about 850 vph. This reduction in flow has the outward appearance of being caused by a drop in phase capacity due to either having a reduced effective green or saturation flow. Of course, no downstream impediment actually exists and any control strategy designed on this false premise would be misguided.

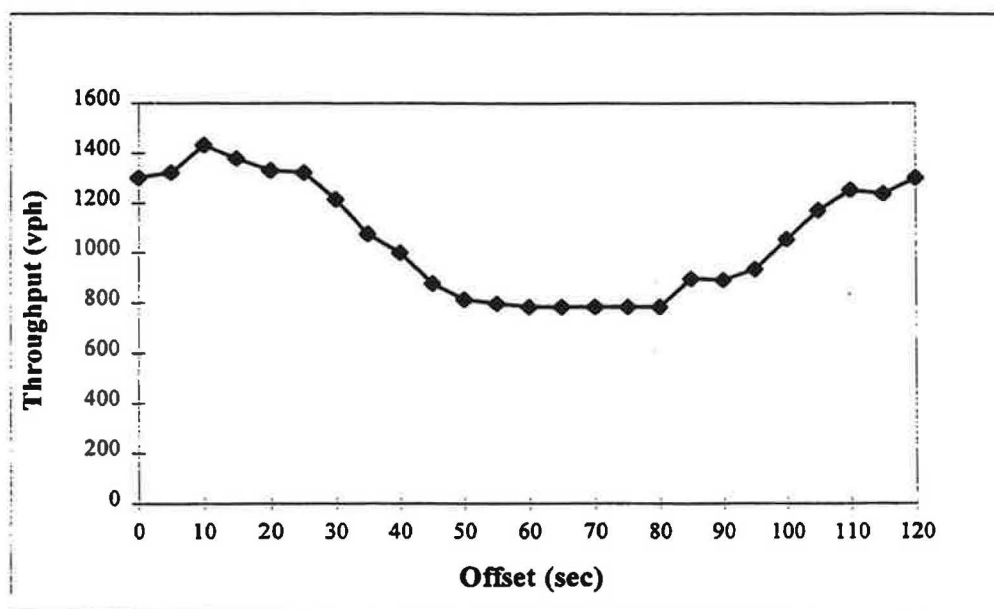


Figure 12. Throughput variation on downstream link with change in offset for undersaturated conditions and arterial dominance.

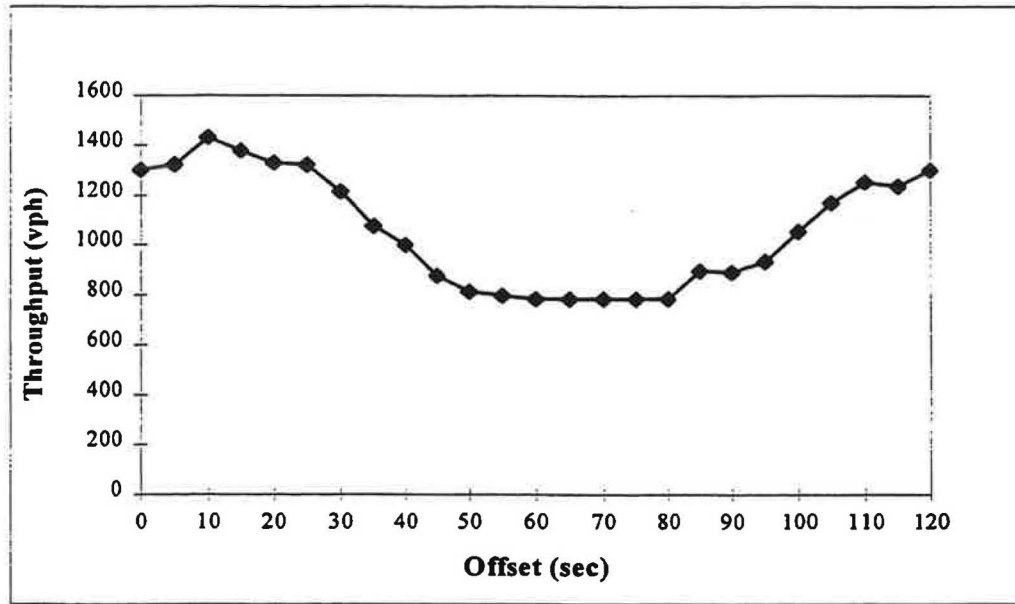


Figure 14. Throughput variation on upstream link with offset variation on downstream link for undersaturated conditions and arterial dominance.

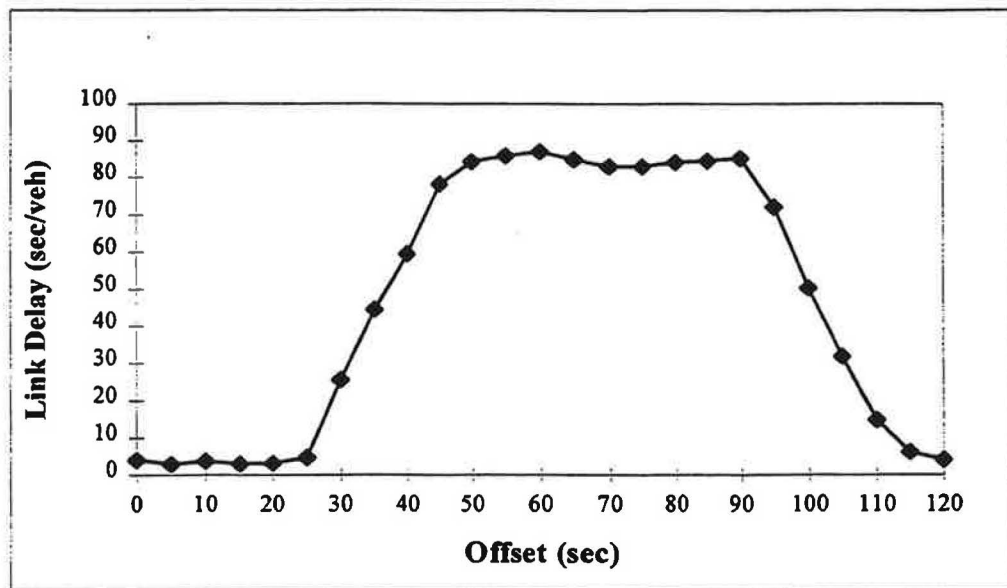


Figure 15. Delay variation on upstream link with offset variation on downstream link for undersaturated conditions and arterial dominance.

These results demonstrate that good/bad signal timing of the next downstream closely-spaced signal can seriously impact the connecting upstream link, even during undersaturated conditions. Moreover, traffic operations on the causal link would likewise suffer when signal timings are poor and blockages occur, but strangely not as much as would have occurred if the link had been longer because there is no place to store longer queues. None of these studies show the difficulties that cross street traffic may have in gaining access to and using the major arterial facility because all link traffic is composed of arterial through traffic. Later studies will illustrate this problem area.

Arterial Predominance. This traffic pattern has 80 percent arterial traffic, which is thought to be typical of nominal arterial streets outside of interchanges and away from shopping malls, other heavy traffic generators, and circulation systems like in downtown areas. Simulation results of arterial throughput are not illustrated, but they show a slight and expected moderation in the effects of demand starvation because of the ability to feed some traffic into the empty arterial during these conditions. As shown in Figure 13, traffic delay experienced upon arrival at a traffic signal is known to vary with arrival pattern (3). The more traffic that arrives on a fixed green per cycle, the less the average delay per vehicle. On the other hand, the higher the proportion of traffic that arrives on red, the higher the delay. During undersaturated conditions, these proportions can vary and so can the resulting delay, as Figure 16 demonstrates based on the NETSIM simulations.

Figure 16 shows that link delay during undersaturation is very sensitive to signal offset for closely spaced intersections where almost no platoon dispersion has occurred from the upstream signal. A fully dispersed platoon would become random flow so that delay would show no sensitivity to changes in signal offset. A link would have to be very long to provide the time to fully disperse a platoon that is composed of mostly arterial through traffic.

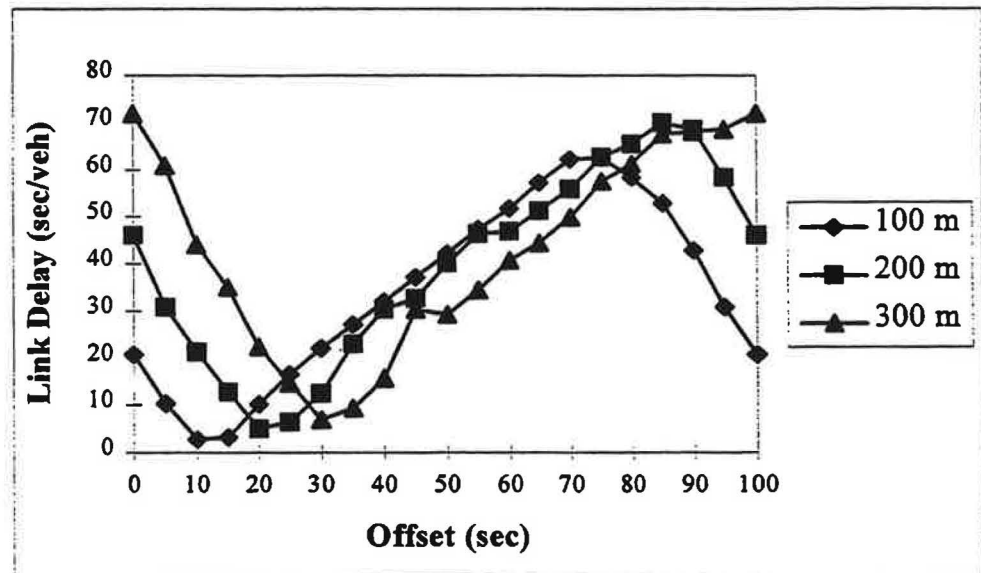


Figure 16. Delay variation on link with offset for nominal arterial conditions.

Balanced Pattern. Assume that the input flows to the links $i-j$ and $j-k$ are now all nearly balanced, rather than coming from just one input (like the arterial dominance case above). In these experiments using NETSIM, the through, left-on, and right-on movements to the head of the link were 50, 25 and 25% of the total downstream link volume, which were served by green ratios of 36, 26 and 26 %, of a 100-second cycle, respectively. Figure 17 presents the NETSIM simulation results for the observed throughput on link $i-j$ for links of 100, 200 and 300 meters long when the connected links $i-j$ and $j-k$ are both undersaturated at v/c ratios of 0.8. The throughput (of 1200 vph) changed only slightly as the signal offset θ_{ij} (the time between the start of arterial through greens $i-j$) varies over the cycle during these undersaturated conditions for any of the link distances studied. Some queue spillback effects are noted for the 100 meter link for a small range of offsets.

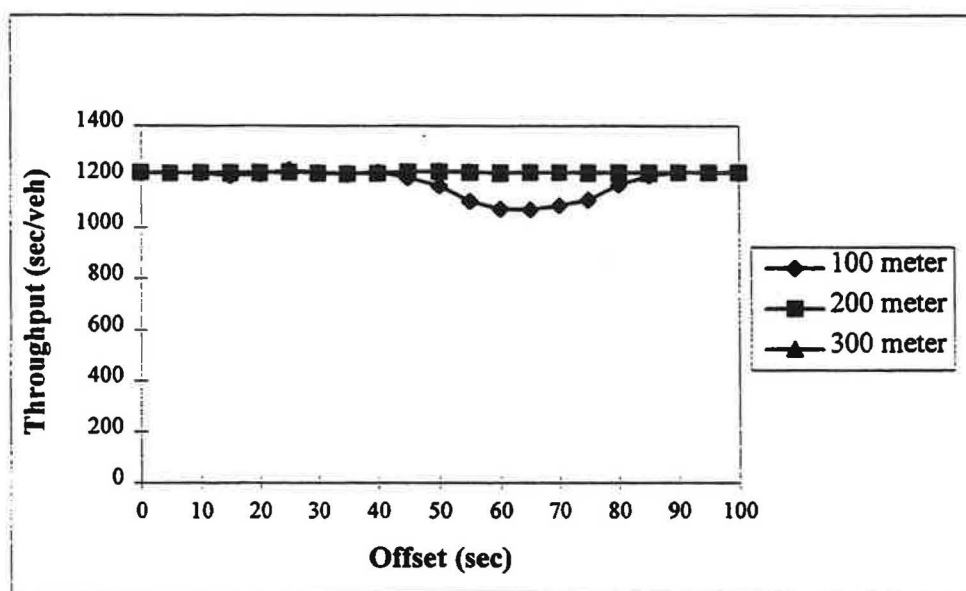


Figure 17. Throughput variation on link with offset change for balanced flow patterns.

Traffic delays incurred when arrival flows are nearly balanced and undersaturated show almost no sensitivity (or predictable change) to signal offset, as demonstrated in Figure 18. This lack of response sensitivity is totally different from that depicted in Figures 13 and 16 for arterial dominated traffic patterns. While the arrival flows are not random, they are nearly uniformly distributed. Selection of progression adjustment factors for delay estimation in HCM-level analyses should reflect this finding. No benefit of progression should be assumed or expected for any signal timing plan developed when upstream flows are nearly constant throughout the cycle. Signal control strategies can only improve link operations by providing a larger green ratio (g/C).

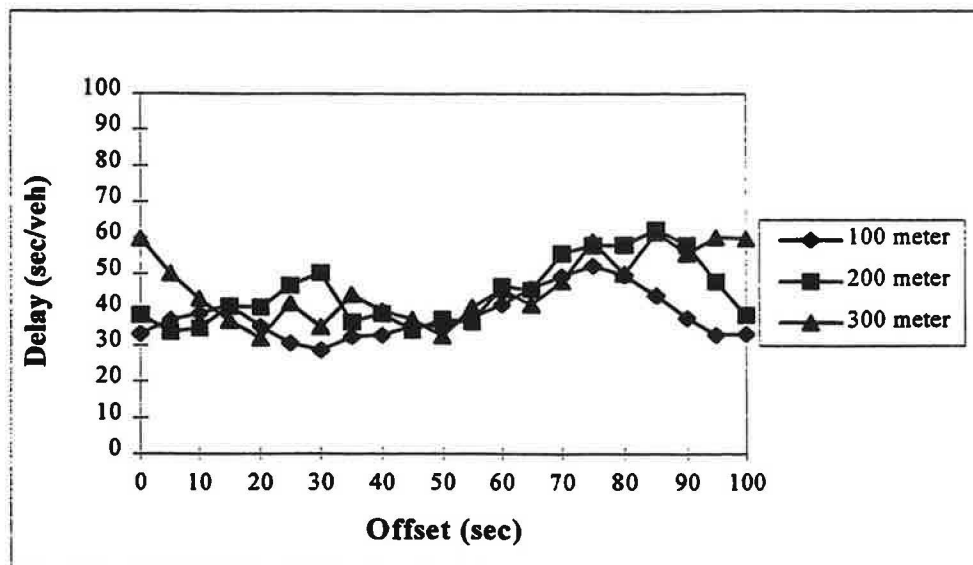


Figure 18. Variation in link delay with offset for undersaturated conditions and balanced flow patterns.

One benefit of balanced input flows is that simultaneous displays of input-output greens and resulting flows occur, and occur more frequently when storage is critical. This feature increases the minimum flow that can occur on short links from that given by Equation 22 by the amount of flow that simultaneously occurs per cycle. This increase in flow may quickly reach the output limit set by the existing capacity of the downstream signal when multiple upstream turning lanes are present.

2.5.2 Oversaturated Conditions

During oversaturated conditions, when upstream traffic demand exceeds downstream signal capacity, queue spillback along the affected links will routinely fill during the signal cycle, and link flow becomes highly output dependent, rather than upstream demand dependent. As the following NETSIM simulation experiments show, variations in flow and delay do occur during oversaturation conditions depending on the length of the link, upstream traffic patterns, and signal offset. However, it may be surprising which upstream movements are impacted the most. As with undersaturated conditions, three traffic patterns will be examined by simulation.

Arterial Dominance. The initial study of oversaturation assumes that all of the input traffic to the two-link arterial testbed ($i-j-k$) has all arterial through traffic with an input traffic demand of 1.5 times the downstream signal capacity. Signal timings provide a cycle of 100 seconds and the arterial green splits are 54% and 36% of the cycle.

Figure 19 presents variations in flow generated by NETSIM for a 100-m length when all of the link traffic is through traffic. As shown earlier, this traffic case is most susceptible to “demand starvation” as the noticeable drop in throughput indicates over a range of link signal offsets. The queue spillback during oversaturation would be blocking the upstream signal, and arrival flows to the link can input a value greater than the downstream signal capacity in this case.

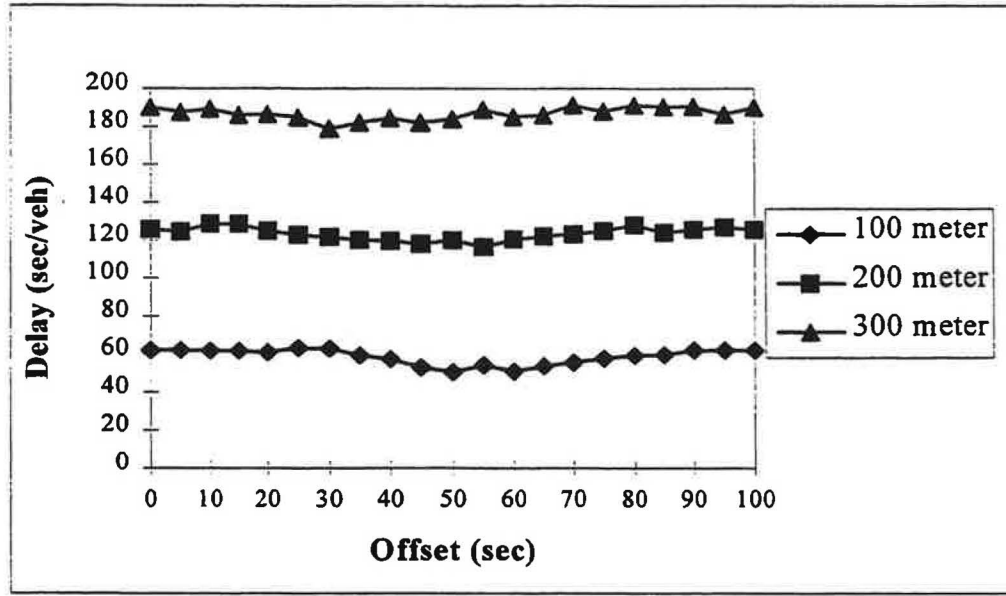


Figure 20. Traffic delay on link with variation in offset for oversaturated conditions and predominantly arterial traffic pattern.

As derived in Appendix D, the minimum average travel speed on the fully loaded link during oversaturation is given by

$$u_l = \frac{u_s}{(1 + \beta)} \quad (25)$$

where:

- u_l = minimum link travel speed during saturation, km/hr;
- u_s = speed at saturation flow, km/hr;
- r, g = effective red (r) and green (g) of the downstream phase, sec, and
- β = $(k_q r)/(k_s g) \approx 2.79 r/g$ of the downstream phase.

The average maximum link delay can be calculated as the difference between the overall link travel time and the baseline running time at the approach running speed u_a . The maximum link delay for a link of length L would be

$$d_l = \frac{L}{u_l} - \frac{L}{u_a} \quad (26)$$

Should the link also experience demand starvation, the “effective” link length experiencing delay should be reduced to reflect the percentage of the cycle demand starvation occurs.

Balanced Pattern. Assume that the input flows to the links $i-j$ and $j-k$ are now all nearly balanced, rather than coming from just one input (like arterial dominance above). In these experiments using NETSIM, the through, left-on, and right-on movements to the head of the link were 50, 25 and 25% of the total downstream link volume, which were served by green ratios of 36, 26 and 26 %, of a 100-second cycle, respectively, to yield the targeted oversaturation v/c ratio of 1.5.

Figure 21 presents NETSIM simulations of observed throughput on link $j-k$ for links of 100, 200 and 300 meters long when the link $j-k$ is oversaturated at a v/c ratio of 1.5. The total throughput on link $j-k$ is seen to change very little with signal offset because the individual upstream movements can keep the downstream link filled sufficiently during its green to maintain saturation output flows. That is, during the time the downstream signal is green, the equation of continuity provides that:

$$n(L, g) = n_0 + \sum v_m g - \sum s_m g \quad 0 \leq n(L, g) \leq n_{\pi} \quad (27)$$

where:

- $n(L, g)$ = number of vehicles operating on the link of length L at end of green, vehicles;
- n_0 = number of vehicles operating on the link at start of green vehicles;
- v_m = total arrival flow into link during green destined to movement m , vph;
- s_m = output saturation flow of movement m , subject to $s_m g \leq$ capacity of link serving output movement m , vphg; and
- n_{max} = maximum number of vehicles that can store on link, vehicles, $k_q L$ with a typical storage density of 143 vpkmpl, or a storage spacing of 7.0 m/veh (23 ft/veh).

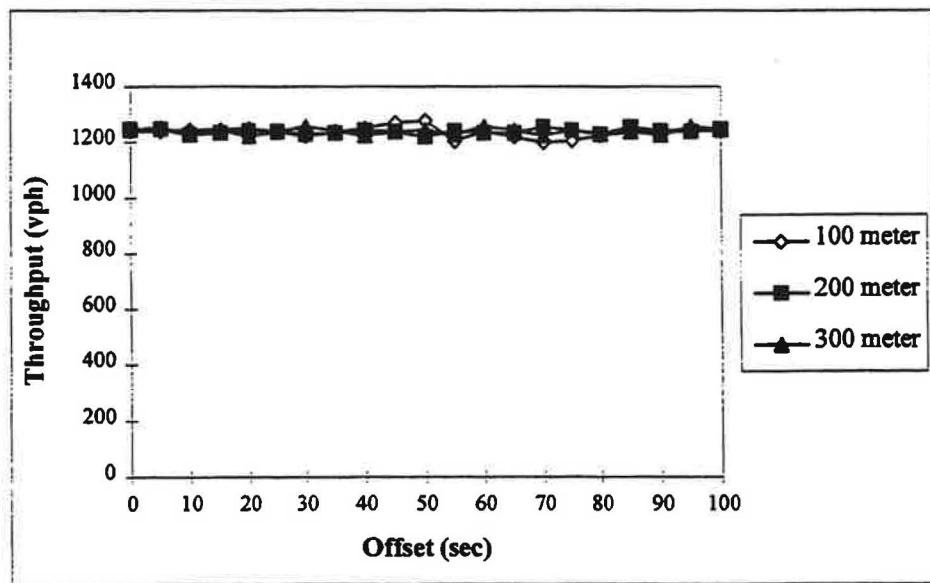


Figure 21. Throughput variation on link with change in offset for oversaturation and balanced traffic pattern.

As long as the arrival flow to the link plus the queue storage at start of downstream green exceeds the downstream phase capacity, the throughput on the link will not change with offset, θ_{ij} . However, the link's signal offset θ_{ij} does control which upstream feeding movements benefit from the available, albeit insufficient, link capacity and which movements get little or no service. Figures 22-24 demonstrate this finding for link spacings of 100, 200 and 300 meters. Note that the total input volumes for the three spacings equal the throughputs shown in Figure 21.

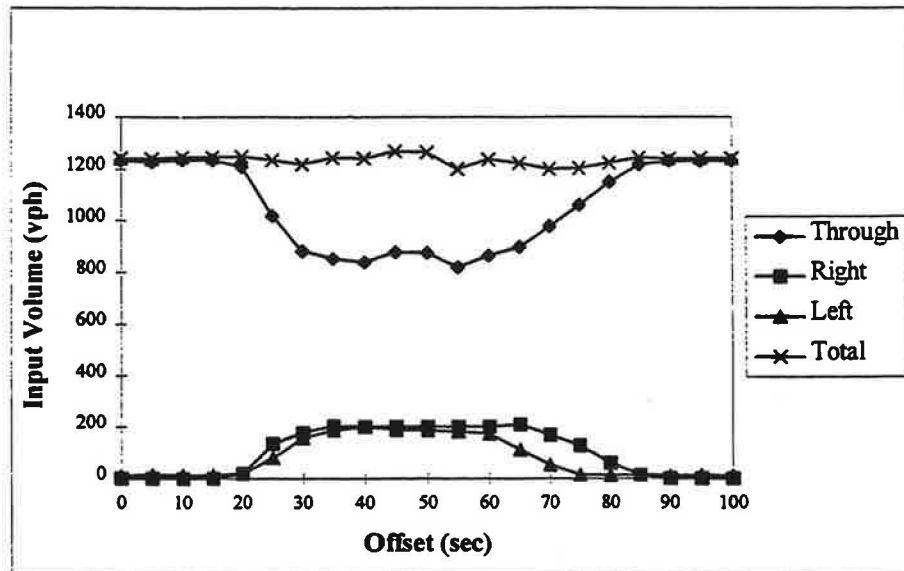


Figure 22. Input volume variation with change in offset for 100 meter link, oversaturated condition and balanced traffic pattern.

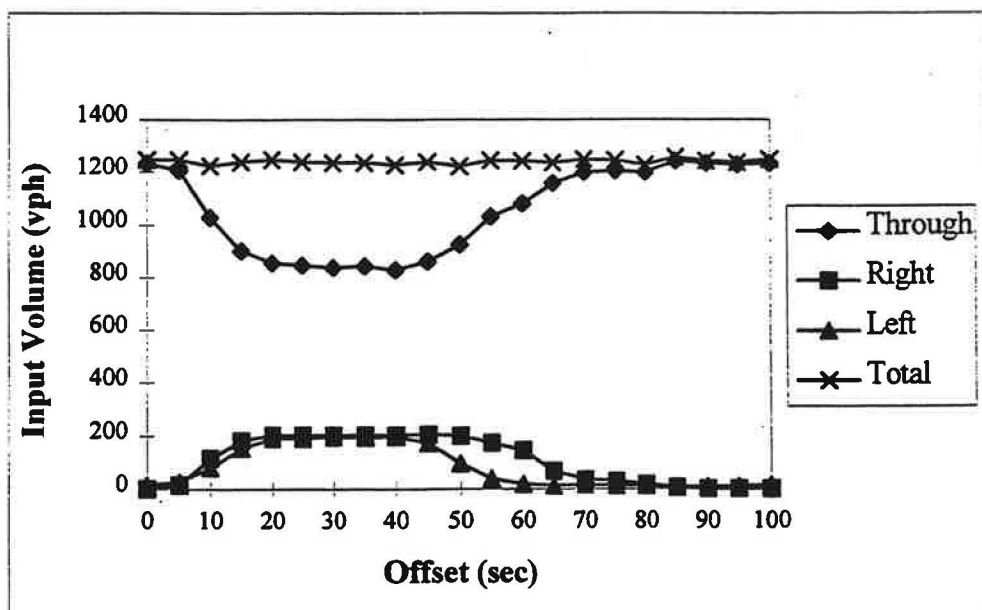


Figure 23. Input volume variation with change in offset for 200 meter link, oversaturated condition and balanced traffic pattern.

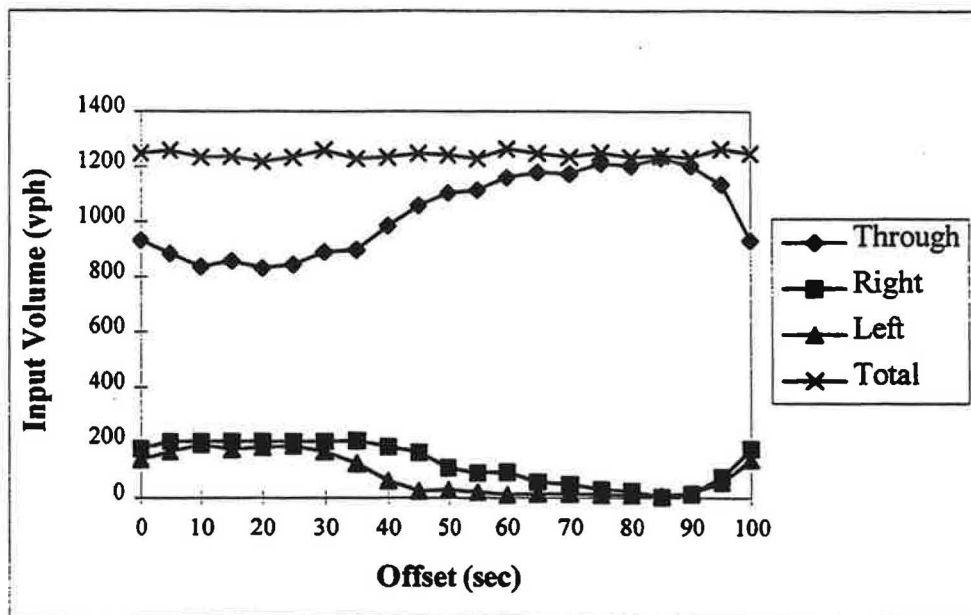


Figure 24. Input volume variation with change in offset for 300 meter link, oversaturated condition and balanced traffic pattern.

Balanced traffic patterns produced arrival delay results on the link shown in Figure 25 that are similar to Figure 20. Figure 20 confirmed Equation 26 in that the maximum delay that can be experienced on a link is primarily a function of the length of the link, and it is not sensitive to excessive arrival volumes. Figure 25 shows that the link delay is basically insensitive to offset during oversaturation for most traffic patterns. Maximum delay is, however, a function of the link length and the effective capacity (green ratio) of the downstream signal. Thus, average travel speed is a good measure of level of service along an arterial, as used in Ch. 11—Signalized Arterials of the HCM (3), but total travel time (or delay) is a better measure of disutility, or cost.

The resulting traffic delays experienced on the three input movements to the link reflect the limited capacity available and allocated to each one by changing the downstream signal offset. Traffic delays observed on the exterior input movements to the upstream intersection follow a consistent but reciprocal pattern to observed flows in that when capacity goes down delay goes up, as Figure 26 depicts for a 100 meter link. Upstream input movement delays are seen to be highly affected by the selection of downstream signal offset. Similar highly sensitive delay patterns, adjusted for travel time differences, were observed in studies for 200 m and 300 m lengths having the same relatively balanced upstream input flows. Should it be desired to favor one movement over the others, then new models are needed to predict what the desired control offset should be.

A powerful new computer algorithm, called the PDX Model, has been developed which provides promise of being able to assess the probable outcome of queue spillback and demand starvation on throughput and delay over all traffic patterns and volume conditions, including oversaturation. This model is described in Chapter Three and Appendix D.

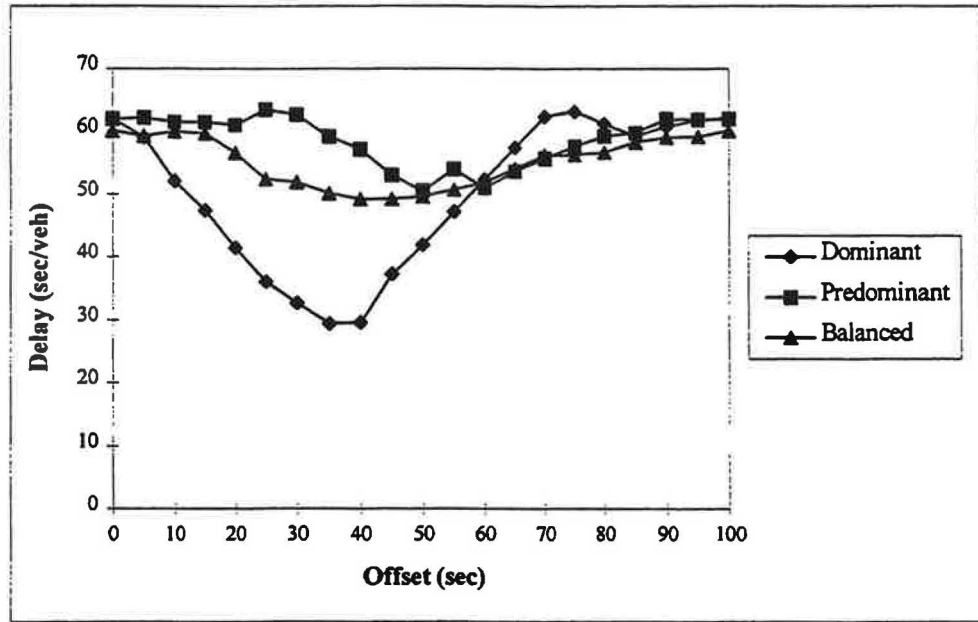


Figure 25. Link delay experienced for three traffic patterns for 100-meter link and oversaturated conditions.

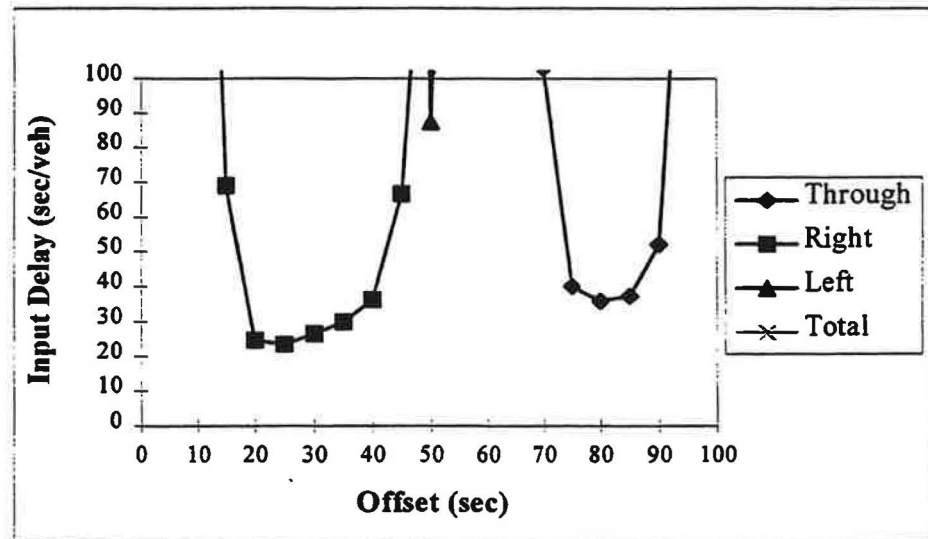


Figure 26. Traffic delay experienced on upstream input movements as downstream offset changes for oversaturated conditions and balanced traffic pattern.

2.6 ARTERIAL WEAVING SPEED

This section describes two models that can be used to evaluate the performance of selected traffic movements in weaving sections on arterial cross streets in interchange areas. This performance is evaluated in terms of the speeds of both the weaving and non-weaving movements. The weaving maneuver that is considered is the off-ramp right-turn movement that weaves across the arterial to make a left-turn at the downstream signalized intersection. Although several other weaving maneuvers exist in interchange areas, the off-ramp weave maneuver is generally found to have the largest volume and to be the most disruptive to arterial traffic flow.

Henceforth, the weaving problem described in this section is referred to as "arterial" weaving. This terminology is adopted to clearly indicate that the weaving studied in this research occurs on streets whose traffic flow is periodically interrupted by traffic signals, as compared to the more extensively studied weaving that occurs on uninterrupted flow facilities (such as freeways).

2.6.1 Data Collection

The data for this study were collected at six study sites in four states. Each study site consisted of a section of urban arterial located between a freeway off-ramp and a closely-spaced signalized intersection. During each field study, flow rates, travel times, travel distance, and stopped delays were measured for seven different travel paths through the weaving section (e.g., upstream entry as a through movement and downstream exit as a right-turn, etc.). Three different types of traffic control (i.e., signal, yield, and uncontrolled) are represented in the database for the off-ramp right-turn movement. Additional details of the data collection effort are provided in Appendix B.

2.6.2 Calibrated Models

Model calibration consisted of using linear regression techniques to calibrate several candidate maneuver speed model formulations. Two models were ultimately identified as having the best fit to the data. One model predicts the weaving maneuver speed and the other model predicts the arterial maneuver speed.

The dependent variable considered in both of these models is the maneuver speed. Maneuver speed is defined as an average running speed and represents the ratio of travel distance to travel time within the study section. The distance (and time) are measured from the point of entry to point where the vehicle first stops in a queue, stops at the stop line, or crosses the stop line and exits the study section, whichever is shortest (or occurs first). Therefore, the corresponding maneuver distance (and time) varies from vehicle to vehicle. It also varies among the two maneuver types (i.e., arterial and weaving) as the weaving maneuver often has to accelerate from a stopped (or slowed) condition whereas the arterial maneuver generally enters the weaving section at speed. The calibrated weaving maneuver speed model is:

$$U_{m,w} = 3.741 U_a^{0.408} e^{(-11.045 (1 - P_U) V_w / 3600)} \quad (28)$$

where:

- $U_{m,w}$ = average maneuver speed for weaving vehicles, m/s;
- U_a = average arterial speed entering the weaving section, m/s;
- P_U = probability of a weaving vehicle being unblocked (i.e., able to change lanes freely); and
- V_w = average weaving flow rate, vph.

This model relates the weaving maneuver speed to the average speed of arterial vehicles entering the weaving section. This latter speed was measured as a spot speed at the point of entry to the arterial weaving section. Hence, it represents the "desired" speed of arterial drivers for the given arterial volume conditions when there is no weaving activity.

The calibrated arterial maneuver speed model is:

$$U_{m,a} = 1.986 U_a^{0.717} e^{(-0.634 V_a/3600)} \quad (29)$$

where:

- $U_{m,a}$ = average maneuver speed for arterial through vehicles, m/s.
- U_a = average arterial speed entering the weaving section, m/s; and
- V_a = average arterial flow rate entering the weaving section, vph.

This model relates the arterial maneuver speed to the average speed of arterial vehicles as they enter the weaving section. The arterial maneuver speed decreases with increasing arterial flow rate. This latter flow rate is strongly correlated with the weaving flow rate and, hence, indirectly accounts for the level of weaving activity.

One of the independent variables used in the weaving maneuver speed model is the "probability of a weaving vehicle being unblocked" P_U . This variable relates to the portion of time that the end of the off-ramp (i.e., the beginning of the weaving section) is not blocked by the passing of the arterial traffic stream. The blocked condition is represented by the formation of platoons in the arterial traffic stream (induced by signalization or random bunching). The quantity P_U can be computed as:

$$P_U = \left(1 - \frac{V_a}{s_l N_l} \right)^{N_l} (1 - I_L) + \left(1 - \frac{V_a}{s_l N_l} \right) I_L \geq 0 \quad (30)$$

where:

- P_U = probability of a weaving vehicle being unblocked;
- s_l = saturation flow rate per lane under prevailing conditions ($\approx 1,800$), vphpl;
- N_l = number of arterial through lanes in the subject direction, lanes;
- I_L = indicator variable (1.0 if $D_{m,w} > 90 (N_l - 1)$, 0.0 otherwise);
- $D_{m,w}$ = average maneuver distance for weaving vehicles ($= L_w - L_{q,w}$), m; and
- $L_{q,w}$ = average length of queue joined by weaving vehicles, m.

Table 15 lists several statistics that indicate the quality-of-fit for each maneuver speed model. As these statistics suggest, the weaving maneuver speed model accounts for 16 percent of the variability in the maneuver speed data. Likewise, the arterial maneuver speed model accounts for 22 percent of the variability in the data. In all cases, the independent variables included in the model were found to be strongly correlated with maneuver speed. All tests were conducted with a 95 percent level of confidence. The root mean square error (or standard error) of each model, combined with the number of observations, yields a minimum precision of ± 0.10 and ± 0.16 m/s for estimates of the average weaving and arterial maneuver speeds, respectively.

Table 15. Maneuver speed model statistics

Maneuver Speed Model	Observations	R ²	Root Mean Square Error	Precision
Weaving	421	0.16	2.02 m/s	± 0.10 m/s
Arterial	324	0.22	2.94 m/s	± 0.16 m/s

Table 16 shows the range of values in the weaving database for the independent and dependent variables included in the maneuver speed models. The variables are listed according to the applicable model. The values in this table indicate the range over which each model is considered valid. In general, the models were calibrated with sites having two or three arterial through lanes (in the subject direction), closely spaced intersections, and a wide range of arterial flow rates.

Table 16. Range of independent and dependent variables

Model	Variable	Variable Name	Units	Minimum ²	Maximum ²
Both ¹	N_t	Arterial through lanes in the subject direction	--	2	3
	U_a	Average arterial entry speed	m/s	8.9	21.8
	V_a	Avg. arterial flow rate entering the weaving section	vph	640	1,924
Weaving	P_U	Probability of a weaving vehicle being unblocked	—	0.27	0.82
	V_w	Average weaving flow rate	vph	88	270
	$u_{m,w}$	Weaving maneuver speed	m/s	2.8	18.7
Arterial	$u_{m,a}$	Arterial maneuver speed	m/s	1.9	23.4

Notes:

- 1 - Variables used by both the weaving and arterial maneuver speed models.
- 2 - All average values are based on a 15-minute intervals.

2.6.3 Sensitivity Analysis

Figure 27 illustrates the effect of arterial flow rate on weaving and arterial maneuver speed. This figure shows the behavior of both models when all other factors are held constant. The values selected for these factors represent their respective average values as found in the database. The range of flow rates over which the two models are compared is larger than the corresponding range in the database. This extension was undertaken to show the overall behavior of each model when extrapolated to extreme (but realistic) values.

The trends in Figure 27 show that both models predict an exponentially decreasing maneuver speed with increasing arterial flow rate. This trend is somewhat consistent with the traditional speed-flow relationship for uninterrupted traffic streams in uncongested conditions. This figure also shows that the arterial maneuver speed is always higher than the weaving maneuver speed for the same flow rate. This trend is reasonable since the arterial vehicles enter the weaving section at speed while the weaving vehicles often must accelerate from a stopped (or slowed) condition when departing the off-ramp. The trend toward convergence of the two models at higher flow rates is also reasonable as the weaving maneuver speed should approach the arterial maneuver speed as the capacity of the weaving section is neared.

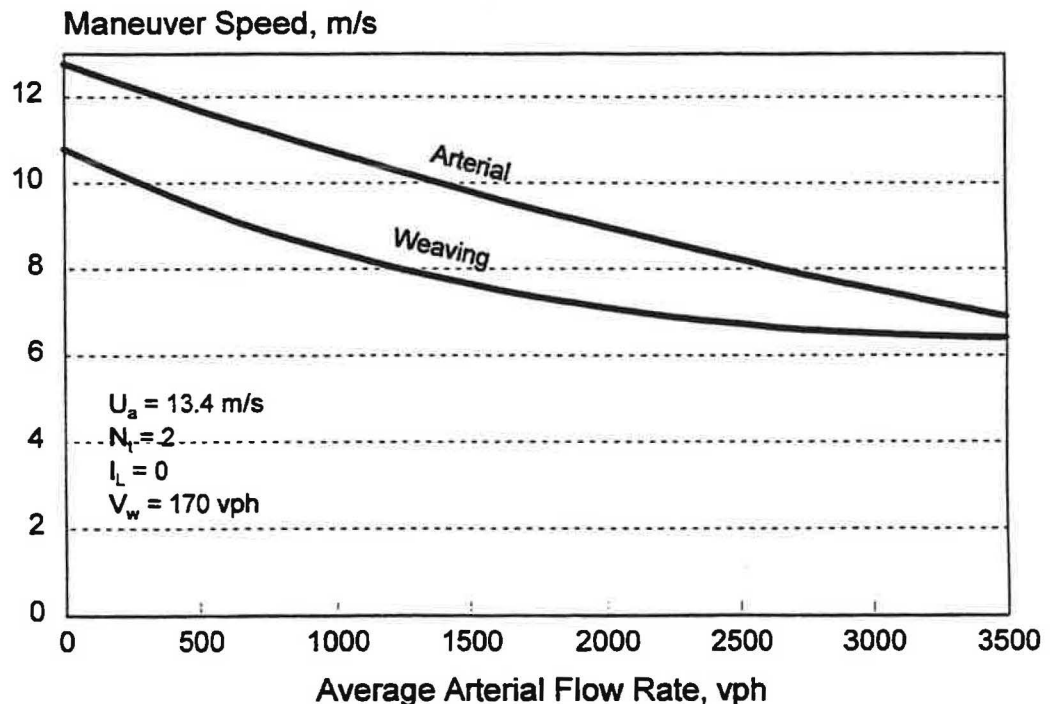


Figure 27. Effect of arterial flow rate on weaving and arterial maneuver speeds.

2.7 RAMP WEAVING CAPACITY MODEL

The section of the cross arterial roadway between an interchange ramp terminal and a closely-spaced downstream intersection generally experiences operational problems, reduced capacity, and deteriorated Levels of Service (LOS) when the ramp-to-intersection weaving is heavy and difficult to perform. The more difficult traffic maneuver to perform usually is the off-ramp right turn trying to cross and then turn left at the next downstream intersection. When the downstream intersection is signalized, additional queuing in the left turn lane shortens the effective weaving length, resulting in increased operational problems.

An additional operational constraint is the physical capacity of the ramp-arterial crossing maneuver. This maneuver usually operates like a freeway merge operation during rush-hour conditions because even free right-turn maneuvers are usually performed from a stopped position in queue. The Highway Capacity Manual (3) does not address arterial weaving. This section will present a method for estimating arterial crossing capacity based on NETSIM traffic simulation studies. Both random and progressed flow conditions along the arterial can be evaluated. Models to predict operating speeds in arterial weaving sections are presented in Appendix E.

2.7.1 Study Methodology

The experimental testbed shown in Figure 28 was coded in TRAF-NETSIM to simulate the study conditions. An arterial free speed of 60 km/h was assumed. The distance between the ramp terminal and the downstream intersection was 200 meters. The ramp traffic, on yield control, made a right turn onto the arterial and then made a left turn at the downstream intersection. The arterial traffic went through the downstream intersection without making any turns. The strategy was to heavily load the cross weave with abundant demand, i.e., maintain a standing off-ramp queue so that the maximum ramp crossing volume could be observed for different operating conditions.

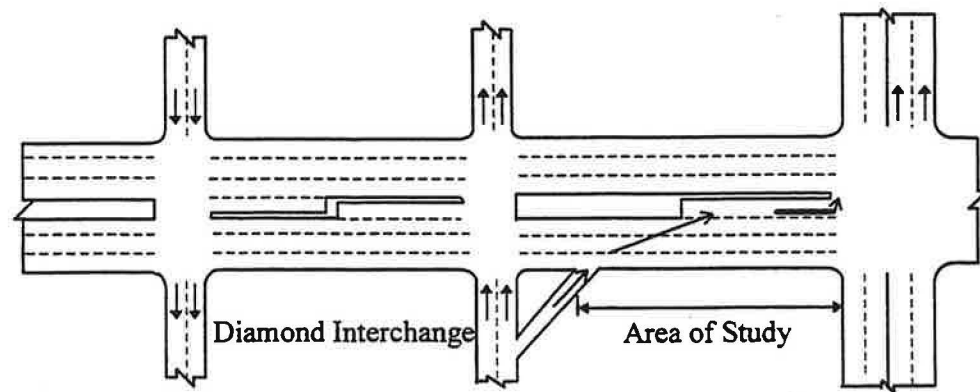


Figure 28. Arterial testbed for ramp-to-arterial weaving analysis.

Preliminary testing revealed that the weaving capacity from the ramp terminal follows the pattern of a negative exponential function with increasing arterial volume. Thus, negative exponential regression analysis was performed to model the weaving capacity. The basic form of the exponential regression equation for the predicting ramp capacity is shown below.

$$Q_R = \frac{Q_t e^{-\alpha Q_t}}{1 - e^{-\beta Q_t}} \quad (31)$$

where :

Q_R	=	ramp crossing/weaving volume (vph);
Q_t	=	arterial through volume (vph);
α	=	coefficient of the model = $T_c / 3600$;
β	=	coefficient of the model = $H_s / 3600$;
T_c	=	critical gap of ramp weave, sec, and
H_s	=	minimum follow-up headway, sec.

The coefficients of the exponential equation, α and β , for random flow were determined on the basis of the simulations for various arterial through volume conditions. The coefficients α and β were computed by inputting the simulated arterial and the ramp crossing volumes into SAS, a statistical software analysis package (4), and performing the desired regression analysis.

For the random flow conditions, the arterial traffic was varied from 100 vph to 2000 vph. Weaving across one, two, and three arterial lanes was studied for the volume conditions noted. Also, the effect of the change in decile gap acceptance distribution in NETSIM was studied.

For progressed flow conditions, the arterial traffic was varied from a v/c of 0.2 (500 vph) to a v/c of 0.8 (2000 vph) for a three lane arterial. A cycle length of 100 seconds and a clearance interval of four seconds per phase were also assumed. Various PF ranging from 0.1 to 1.8 were also simulated by varying the percent vehicles arriving on green (PVG) at the upstream intersection.

2.7.2 Study Results

The next section consists of the results obtained in the various cases involving random flow conditions along the arterial. Also, the computed coefficients for determining the ramp crossing volumes for different arterial flow conditions are presented. Changes in the gap acceptance distribution were observed to affect the ramp crossing volume. The second section presents the calibration coefficients for the proposed negative exponential equation for computing the ramp crossing volume for different arterial through volumes. The third section covers the results of simulations involving several volume conditions and different progression factors. The effect of progression on the ramp crossing volume is discussed in detail in this section. The development of the final model form and the methodology used to predict the ramp capacity across the arterial weaving section for various progression factors are presented in Chapter 3 and Appendix E.

Random Flow. Initial ramp capacity studies were conducted with NETSIM assuming that the cross arterial had no progression and random flow. Moreover, preliminary testing of the model assessed the sensitivity of capacity to the default gap acceptance function provided in the model. The 1994 HCM states that the critical gap for a right turn from a Yield sign onto a major street could be taken as 5.5 seconds (3). TRAF-NETSIM assumes a decile distribution wherein the default median value is taken as 6.4 seconds. In order to simulate the HCM recommended distribution, Card Type 145 in TRAF-NETSIM was coded to produce a decile distribution having a median value of 5.5 seconds. Hence the data file with the new decile distribution and an upstream link length of 365 meters was simulated for random flow.

The effect of changing the decile gap distribution for three lanes can be seen in Figure 29. Due to the lower (better) gap acceptance criteria, more ramp vehicles can make a right turn onto the arterial. Though the trend is similar, the ramp crossing volume for the HCM decile distribution is slightly higher than the TRAF-NETSIM default decile distribution. Following a review of the gap acceptance study results shown in Figure 29, it was arbitrarily decided to continue using the NETSIM default distribution in subsequent model building.

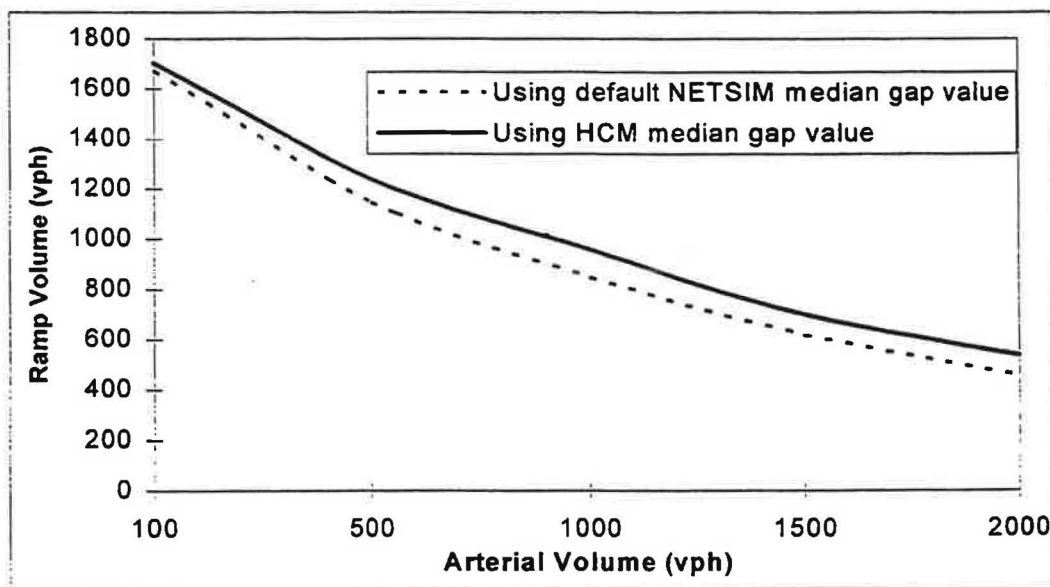


Figure 29. Effect of NETSIM decile gap distribution for three-lane arterial.

The effect of the number of lanes on ramp crossing volumes is illustrated in Figure 30. The drop in the ramp crossing volume is sharper with an increase in the number of vehicles on the one-lane arterial because all the vehicles have to use the single lane so the number of acceptable gaps available is reduced. For the two and the three lane cases, the same number of vehicles are distributed over two or three lanes, as the case may be, and there is a lesser effect on the ramp crossing vehicles. The net increase in the vehicles per hour per lane for the one lane arterial case is largest and hence its ramp capacity is affected the most.

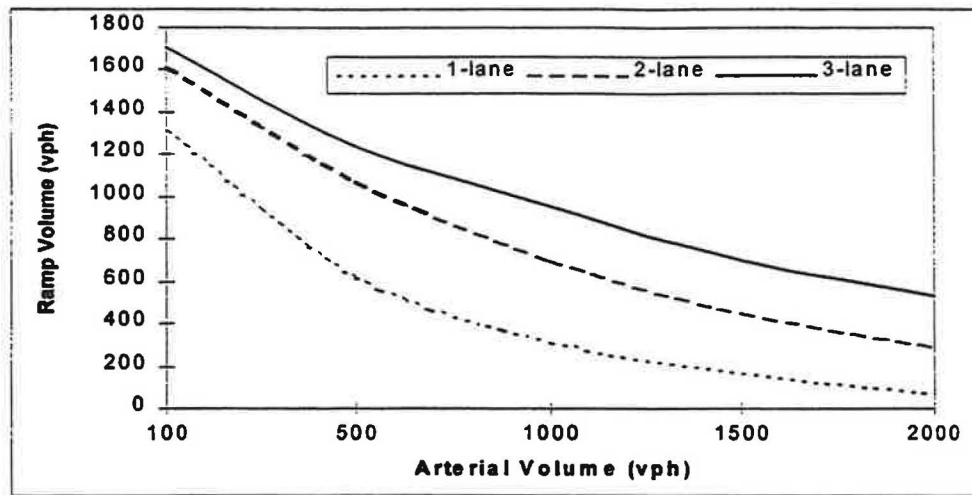


Figure 30. Effect of number of lanes on maximum ramp volume.

Observations of the simulation results of Figure 30 suggest that an exponential model would reasonably fit the interchange ramp capacity results generated by NETSIM. The values of ramp capacity were obtained by simulation of the desired conditions and the coefficients of the model were determined using SAS, a statistical analysis software package (4). Figure 31 shows how well the model fits the traffic simulation program values. The points indicate the average of ten simulation runs while the lines indicate the trend using the calibrated exponential model.

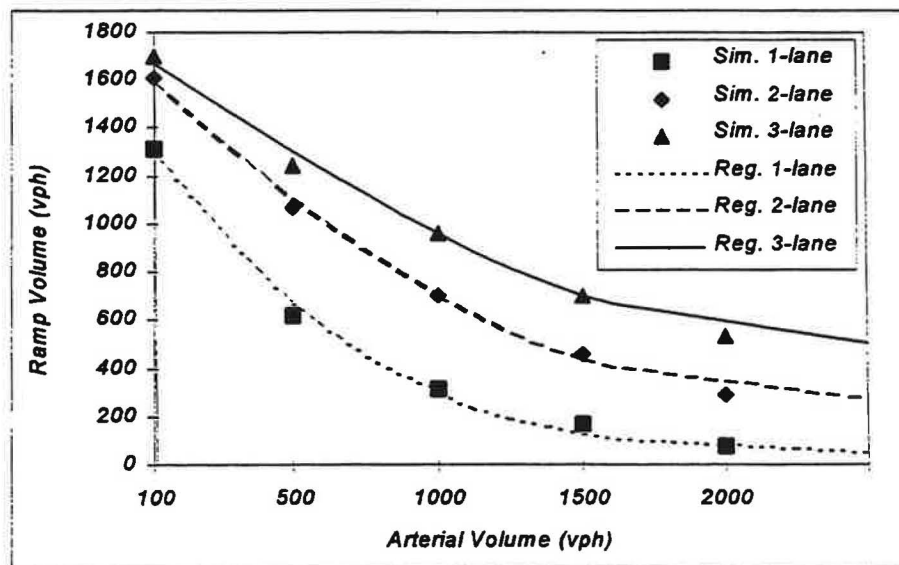


Figure 31. Comparisons of ramp capacity for simulation and exponential regression model results.

Applying the exponential regression analysis in SAS, the average of the observed volume data for each case was used to estimate the unknown coefficients α and β in the exponential equation. α multiplied by 3600, denoted as T_c , is the average critical gap time of the corresponding lane configuration; while β multiplied by 3600, denoted as H_s , is the minimum headway of the arterial weaving section. Table 17 shows the coefficients α and β of the exponential model computed for one, two and three lane arterials. The coefficients in the proposed exponential equation are accurate estimations of the TRAF-NETSIM simulated operations in terms of standard errors and their variances. Table 18 illustrates the coefficients of the model on a per lane basis. For the per lane analysis, the results of the one, two and three lane cases were pooled and regressed. It can be observed that the values of T_c and H_s are close to that of the one lane case.

Table 17. Coefficients of the exponential regression model

Lanes	1-lane		2-lane		3-lane	
Coefficients	α	β	α	β	α	β
Exponential	0.00195	0.000657	0.00118	0.000574	0.00088	0.000565
R ² Value	0.9977		0.9995		0.9989	
Conversion of Coefficients	T_c	H_s	T_c	H_s	T_c	H_s
Values (sec.)	7.02	2.36	4.26	2.06	3.17	2.03

Table 18. Coefficients of the exponential regression model on a per lane basis

Coefficients	Exponential	R ² Value	Conversion of Coefficients	Values (sec.)
α	0.002091	0.9970	T_c	7.52
β	0.000583		H_s	2.10

Progressed Flow. The NETSIM simulations were used to determine ramp crossing volumes for progressed arterial flow. Different progression factors were analyzed, ranging from PF s of 0.1 to 1.8. A PF value of 1.0 is essentially uncoordinated, uniformly distributed flow. Progression factors from unity reflect the degree of platooning of the dominant flow. Volume-to-capacity ratios of 0.2, 0.4, 0.6, 0.7 and 0.8 on the upstream feeding movements were studied for a three-lane arterial. In order to simulate various PF , vehicles were emitted from the upstream intersection operating in two phases to create two platoons flowing downstream such that one platoon arrives on red and the other platoon arrives on green. At the merge point, the notion of red and green only characterizes the degree of platooning in the arterial flow, as there is no signal at the merge point.

Figure 32 summarizes the experimental results. Polynomial regression equations for these plots were determined using SAS. During low volume conditions on the arterial, little change occurred in ramp crossing capacity for different PF s. As the volume on the arterial increased, the ramp crossing volume decreased significantly due to fewer acceptable gaps for the weaving maneuver. For higher volumes, the change in the ramp crossing volumes for various PF s becomes more significant. A PF of 1.0 is considered random flow and the ramp crossing volume is the least for a PF of 1.0 when compared to other PF s between 0.1 to 1.8. The flow graph takes the shape of a parabola which has its minimum at a PF of 1.0. Figure 32 clearly indicates the trend of ramp crossing volumes for v/c ratios of 0.2, 0.4, 0.6, 0.7 and 0.8, respectively. The difference in ramp crossing volume for a PF of 0.1 to that of a PF of 1.0 increases with an increase in arterial volume. In other words, the difference between the ramp crossing volumes for a PF of 0.1 to that of a PF of 1.0 increases with an increase in the v/c ratios on the arterial. Since the green ratio (green time/cycle length) is the same along both the approaches of the arterial, the PVG for a PF less than 1.0 corresponds to the PVR for a PF greater than 1.0 and vice-versa. Plots of ramp crossing volumes for PF less than 1.0 are a mirror image of plots for PF greater than 1.0 about the axis of PF of 1.0.

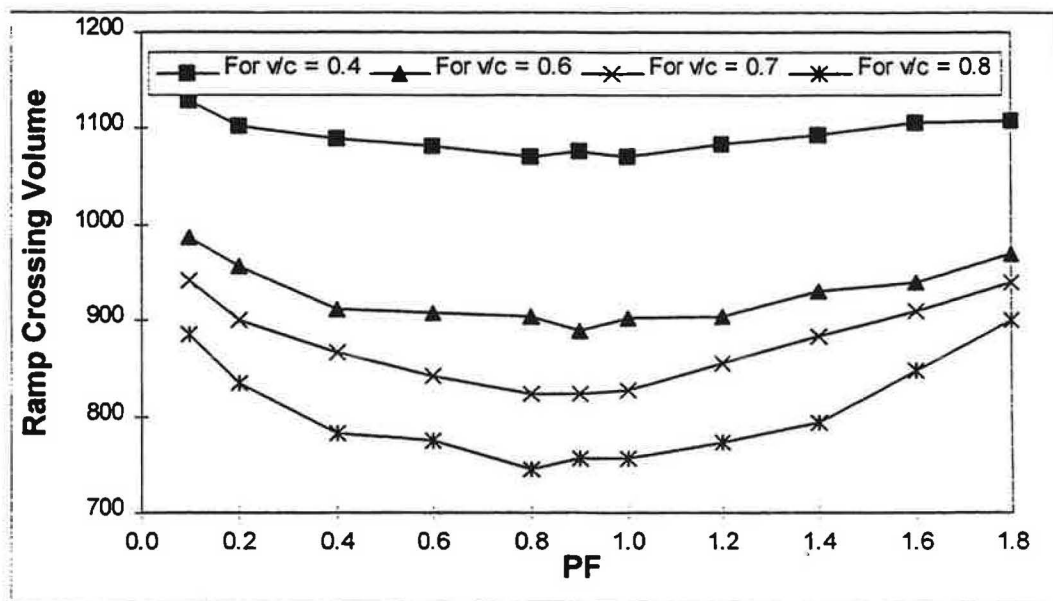


Figure 32. Effect of PF on ramp crossing volume for various v/c .

Adjustment Factors for Progression. In order to further simplify the simulation results, the regression equations from the graphs for various v/c ratios were used to determine individual values of ramp crossing volume. A PF of 1.0, also considered as random flow, was used as the basis for development of the adjustment factors. The factors for other PF were computed by determining the ratio of the value at PF of 1.0 to that of another PF . Because the curves were parabolic and the values on one side of the curve were mirror images of the other, adjustment factors for PF from 0.1

to 1.0 were computed. Figure 33 presents adjustment factors for various PF ranging from 0.1 to 1.0 for volume-to-capacity (v/c) ratios of 0.4, 0.6, 0.7 and 0.8.

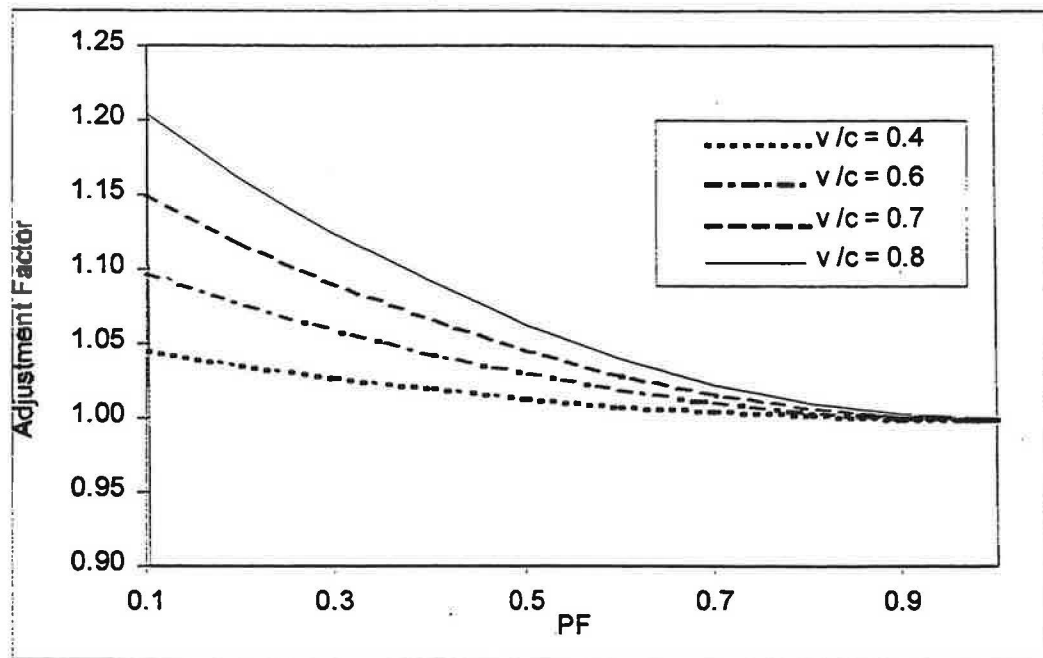


Figure 33. Capacity adjustment factors for various progression factors.

Table 19 shows the actual (average) capacity adjustment factors simulated for various progression factors ranging from 0.1 to 1.0. Table 20 provides related capacity adjustment factors obtained using the exponential equation (Equation 31) shown below as developed from Table 19 average results using SAS. Note that average arterial lane volumes V are used in Equation 31 to provide a more convenient data input format.

$$f_{PF} = 1 + 0.015 * e^{[0.0044 * V - 3.05 * PF]} \quad (32)$$

where:

- f_{PF} = ramp weaving capacity adjustment factor;
- PF = progression factor; and
- V = arterial volume per hour per lane (vphpl).

From a comparison of Tables 19 and 20, it can be seen that the above exponential equation follows a close fit of the actual average adjustment factors. The sum of squares error (SSE) was determined to be 0.00217.

2.7.3 Determination of Ramp Capacity during Random Flow

The ramp capacity during random flow, denoted by Q_R can be determined by inputting the total arterial volume in Equation 31. The computed coefficients α , β shown in Table 17 have been computed on the basis of number of lanes in the arterial section and can be used in the exponential equation. The results of one, two and three lane arterials were pooled and regressed to compute the common coefficients for single and multilane arterials. Depending on the degree of accuracy required by the user, the different coefficients could be used for predicting ramp capacity.

Table 19. Actual adjustment factors for PF of 0.1 to 1.0

v/c	Progression Factors, PF									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.4	1.046	1.036	1.028	1.020	1.014	1.009	1.005	1.002	1.001	1.000
0.6	1.098	1.077	1.059	1.043	1.030	1.009	1.011	1.005	1.001	1.000
0.7	1.150	1.118	1.090	1.066	1.046	1.019	1.017	1.007	1.002	1.000
0.8	1.205	1.162	1.124	1.091	1.063	1.030	1.023	1.010	1.003	1.000

Table 20. Computed adjustment factors for PF of 0.1 to 1.0

v/c	Progression Factors, PF									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.4	1.048	1.035	1.026	1.019	1.014	1.010	1.008	1.006	1.004	1.003
0.6	1.100	1.074	1.054	1.040	1.029	1.021	1.016	1.012	1.009	1.006
0.7	1.142	1.105	1.077	1.057	1.042	1.031	1.023	1.017	1.012	1.009
0.8	1.208	1.153	1.113	1.083	1.061	1.045	1.033	1.025	1.018	1.013

2.7.4 Adjustment for Sneakers

Comparison of the simulation results between random flow and progressed flow at a PF of 1.0 revealed that the ramp crossing volume for a PF of 1.0 was higher. The difference between the ramp crossing volumes between random and progressed flow increased with an increase in arterial volumes. This difference in ramp crossing volumes was attributed to the sneakers crossing during the two phase change intervals i.e., sneakers (S_m). In other words, the ramp vehicles completed the weaving maneuver by making use of the large gap available to the ramp vehicles during the two phase change intervals of four seconds each at the upstream intersection. The random flow conditions had a situation wherein the upstream intersection had 100 percent green on the arterial movement and hence the effect of sneakers was not observed. On the average, approximately three vehicles were completing the weaving maneuver during each phase change interval. The effect of

sneakers was confirmed visually by observing the animation of the simulation for the required conditions in GTRAF. Thus, the ramp crossing capacity, adjusted for sneakers, would be

$$Q'_R = Q_R + S_m \quad (33)$$

where:

$$\begin{aligned} Q'_R &= \text{ramp capacity adjusted for sneakers (vph);} \\ Q_R &= \text{ramp capacity for random flow (vph); and} \\ S_m &= \text{sneaker volume (vph).} \end{aligned}$$

2.7.5 Application of Adjustment Factors

In order to obtain the ramp capacity for different progression factors, the adjustment factors for progression, f_{PF} , needs to be multiplied to the ramp capacity which has been adjusted for sneakers as follows:

$$Q_{PF} = Q'_R * f_{PF} \quad (34)$$

where:

$$\begin{aligned} Q_{PF} &= \text{ramp capacity adjusted for progression (vph);} \\ Q'_R &= \text{ramp capacity for random flow (vph); and} \\ f_{PF} &= \text{adjustment factor for progression.} \end{aligned}$$

The application methodology of this formulation of arterial weaving will be presented in Chapter 3. Moreover, field studies of arterial weaving operations were conducted and are described in detail in Appendix E (12). Several empirical models of maneuver speeds and delays as related to local conditions are provided. These studies were extremely tedious and time consuming. Other initial NETSIM simulation studies of arterial/ramp weaving operations were conducted and reported (13). All of these studies showed the benefit of increased signal separation between the interchange and the next downstream signal together with the benefit of arterial signal coordination during undersaturated conditions.

CHAPTER 3

INTERPRETATION, APPRAISAL, APPLICATIONS

3.1 CAPACITY CHARACTERISTICS

This section summarizes the models developed for capacity prediction. These models can be used to predict capacity-related traffic characteristics for a range of conditions common to interchanges and closely-spaced signalized intersections. The specific characteristics considered include saturation flow rate, start-up lost time, end lost time, and lane utilization.

The capacity of lane group represents the maximum number of vehicles that can be served by the group's traffic lanes during the time allocated to it within the signal cycle. A lane group flowing at capacity is typically characterized by the continuous discharge of queued traffic for the duration of the signal phase that controls it. To this extent, lane group capacity is dependent on the discharge characteristics of the departing traffic queue. These characteristics include the time lost at the start of the phase due to driver reaction time and acceleration, the saturation flow rate, and the time lost at the end of the phase due to a necessary change interval. The following equation is commonly used to compute the capacity of a lane group:

$$c = \frac{gs}{3,600} \quad (35)$$

with

$$g = [G + Y + R_c - (l_s + l_e)] \cap CP \quad (36)$$

where:

- c' = capacity of the lane group, vpc;
- g = effective green time where platoon motion (flow) can occur, sec;
- s = saturation flow rate for the lane group under prevailing conditions, vphg;
- G = green signal interval, sec;
- Y = yellow interval, sec;
- R_c = red clearance interval, sec;
- l_s = start-up lost time, sec;
- l_e = clearance lost time, sec; and
- CP = clear period during cycle/phase when subject flow is unblocked, sec.

Many of the variables defined in Equations 35 and 36 represent the basic set of capacity characteristics. Models for estimating these characteristics (i.e., saturation flow rate, start-up lost time, and clearance lost time) for all traffic movements at interchange ramp terminals and closely-spaced intersections are described in the following sections.

3.1.1 Saturation Flow Rate Model

The saturation flow rate for a lane group can be predicted using the following equation:

$$s = N \times s_0 \times f_w \times f_{HV} \times f_g \times f_p \times f_{bb} \times f_{RT} \times f_{LT} \times f_D \times f_v \quad (37)$$

where:

- s = saturation flow rate for the lane group under prevailing conditions, vphg;
- s_0 = saturation flow rate per lane under ideal conditions, pcphgpl;
- f_w = adjustment factor for lane width;
- f_{HV} = adjustment factor for heavy vehicles;
- f_g = adjustment factor for approach grade;
- f_p = adjustment factor for parking;
- f_{bb} = adjustment factor for bus blockage;
- f_{RT} = adjustment factor for right-turns in the lane group;
- f_{LT} = adjustment factor for left-turns in the lane group;
- f_D = adjustment factor for distance to downstream queue at green onset; and
- f_v = adjustment factor for volume level (i.e., traffic pressure).

The first seven adjustment factors listed in this equation represent those described in the HCM (3, Chapter 9). The last two factors were developed for this research and are specifically applicable to interchanges and closely-spaced signalized intersections.

A third adjustment factor f_R was developed for this research that quantifies the effect of turn radius on the saturation flow rate of a left or right-turn movement. This factor is introduced by revising the adjustment factors for protected turn movements provided in the HCM. These revised right and left-turn adjustment factors are:

$$f_{RT} = \frac{1}{1 + P_{RT} \left(\frac{1}{f_R} - 1 \right)} \quad : \text{Protected, Shared Right-Turn Lane} \quad (38)$$

$$f_{RT} = f_R \quad : \text{Protected, Exclusive Right-Turn Lane}$$

$$f_{LT} = \frac{1}{1 + P_{LT} \left(\frac{1}{f_R} - 1 \right)} \quad : \text{Protected, Shared Left-Turn Lane} \quad (39)$$

$$f_{LT} = f_R \quad : \text{Protected, Exclusive Left-Turn Lane}$$

where:

f_R = adjustment factor for the radius of the travel path (based on the radius of the applicable left or right-turn movement);
 P_{RT} = portion of right-turns in the lane group; and
 P_{LT} = portion of left-turns in the lane group.

Adjustment factors for permissive-only and protected-permissive phasing can also be developed using Equations 38 and 39 as a basis.

The ideal saturation flow rate s_0 represents the saturation flow rate of a lane that is not affected by any external environmental factors (e.g., grade), non-passenger-car vehicles (e.g., trucks), and constrained geometrics (e.g., less than 3.6-meter lane widths, curved travel path). In this regard, the saturation flow rate would be equal to the ideal rate when all factor effects are optimum for efficient traffic flow and the corresponding adjustment factors are equal to 1.0. Based on this definition, it was determined that ideal conditions were represented by an infinite distance-to-queue, non-spillback conditions, a tangent alignment (i.e., an infinite radius), and a traffic pressure representing that typically experienced during peak traffic periods. Under these conditions, 2,000 pcphgpl is recommended as the ideal saturation flow rate for all traffic movements at interchanges.

One potential factor that was evaluated but not included in Equation 37 is that of phase duration. Some evidence was found that drivers adopted larger saturation flow rates for phases of short duration than for those of long duration. However, the relationship was not felt to be conclusive for both left-turn and through movements, thus, more research is believed to be needed before an adjustment for phase duration can be recommended for general applications.

Distance-to-Queue Adjustment Factor. The distance-to-queue adjustment factor f_D accounts for the adverse effect of downstream queues on the discharge rate of an upstream traffic movement. In general, the saturation flow rate is low for movements that have a downstream queue relatively near at the start of the phase; it is high for movements that are not faced with a downstream queue at the start of the phase. Thus, the distance-to-queue adjustment factor is based on the distance to the back of the downstream queue at the start of the subject phase.

The variable "distance to queue" is measured from the subject (or upstream) movement stop line to the "effective" back of queue. The effective back of queue represents the location of the back of queue if all vehicles on the downstream street segment (moving or stopped) at the start of the phase were joined into a stopped queue. If there are no moving vehicles at the start of the phase, then the effective and actual distance to queue are the same. If there are no vehicles on the downstream segment at the start of the phase, then the effective distance to queue would equal the distance to the through movement stop line at the downstream intersection. The distance-to-queue is computed as:

$$D = L - \frac{n_s}{N_d} L_v \quad (40)$$

with

$$L_v = (1 - P_{HV}) L_{pc} + P_{HV} L_{HV} \quad (41)$$

where:

- D = effective distance to the back of downstream queue (or stop line if no queue) at the start of the subject (or upstream) phase, m;
- L = distance between the subject and downstream intersection stop lines (i.e., link length), m;
- n_s = number of vehicles on the downstream street segment (moving or queued) at the start of the subject phase, veh;
- N_d = number of through lanes on the downstream segment, lanes;
- L_v = average lane length occupied by a queued vehicle, m/veh;
- L_{pc} = lane length occupied by a queued passenger car (≈ 7.0 m/pc), m/pc;
- L_{HV} = lane length occupied by a queued heavy vehicle (≈ 13 m/veh), m/veh; and
- P_{HV} = portion of heavy vehicles in the traffic stream.

The effect of distance-to-queue is also dependent on whether spillback occurs during the subject phase. Spillback is characterized by the backward propagation of a downstream queue into the upstream intersection such that one or more of the upstream intersection movements are effectively blocked from discharging during some or all of their respective signal phase. If spillback occurs during the phase, the saturation flow rate prior to the occurrence of the spillback is much lower than it would be if there were no spillback. Thus, the magnitude of the adjustment to the saturation flow rate is dependent on whether spillback occurs during the subject phase. A procedure for determining if queue spillback occurs is provided Appendix C. The distance-to-queue adjustment factor is tabulated in Table 21. It can also be computed using the following equation:

$$f_D = \begin{cases} \frac{1}{1 + \frac{8.13}{D}} & : \text{ no spillback} \\ \frac{1}{1 + \frac{21.8}{D}} & : \text{ with spillback} \end{cases} \quad (42)$$

where:

- f_D = adjustment factor for distance to downstream queue at green onset; and
- D = effective distance to the back of downstream queue (or stop line if no queue) at the start of the subject (or upstream) phase, m.

Table 21. Adjustment factor for distance-to-queue (f_D)

Distance to Back of Queue at Start of Subject Phase, m	Spillback Condition	
	No Spillback	With Spillback
15	0.649	0.408
30	0.787	0.579
60	0.881	0.734
120	0.937	0.846
180	0.957	0.892
240	0.967	0.917
300	0.974	0.932
360	0.978	0.943

Turn Radius Adjustment Factor. Traffic movements that discharge along a curved travel path do so at rates lower than those of through movements. The effect of travel path radius is tabulated in Table 22. It can also be computed using the following equation:

$$f_R = \frac{1}{1 + \frac{1.71}{R}} \quad (43)$$

where:

f_R = adjustment factor for the radius of the travel path; and

R = radius of curvature of the left-turn travel path (at center of path), m.

Table 22. Adjustment factor for turn radius (f_R)

Radius of the Travel Path, m	Movement Type	
	Left & Right-Turn	Through
8	0.824	1.000
15	0.898	1.000
30	0.946	1.000
45	0.963	1.000
60	0.972	1.000
75	0.978	1.000
90	0.981	1.000
105	0.984	1.000

This factor was calibrated for left-turn traffic movements; however, a comparison of this adjustment factor to that developed by others for right-turn movements indicates close agreement. Therefore, this factor is recommended for use with both left and right-turn movements.

Traffic Pressure Adjustment Factor. Saturation flow rates are generally found to be higher during peak traffic demand periods than during off-peak periods. This trend is explained by the effect of "traffic pressure." In general, it is believed that traffic pressure reflects the presence of a large number of aggressive drivers (e.g., commuters) during high-volume traffic conditions. They demonstrate this aggressive behavior by accepting shorter headways during queue discharge than would less-aggressive drivers. As these aggressive drivers are typically traveling during the morning and evening peak traffic periods, they are found to represent a significant portion of the traffic demand associated with these periods.

The effect of traffic pressure was found to vary by traffic movement. Specifically, the left-turn movements tended to be more affected by pressure as their saturation flow rates varied more widely than those of the through movements for similar conditions. It is possible that this difference between movements stems from the longer delays typically associated with left-turn movements.

Based on the preceding definition, it is logical that the effect of traffic pressure is strongly correlated with the demand flow rate in the subject lane group. Thus, the effect of traffic pressure, as represented by traffic volume, is tabulated in Table 23 for each movement type. It can also be computed using the following equation:

$$f_v = \begin{cases} \frac{1}{1.07 - 0.00672 v_l'} & : \text{left-turn} \\ \frac{1}{1.07 - 0.00486 v_l'} & : \text{through or right-turn} \end{cases} \quad (44)$$

where:

f_v = adjustment factor for volume level (i.e., traffic pressure); and
 v_l' = demand flow rate per lane (i.e., traffic pressure), vpcpl.

As noted previously, ideal conditions were defined to include a traffic pressure effect representative of peak traffic periods. In this regard, traffic pressures under ideal conditions were defined as 10 and 15 vpcpl for the left-turn and through movements, respectively. These flow rates are conservative in their representation of higher volume conditions as they exceed 80 to 90 percent of all traffic demands that were observed at the field study sites. One consequence of this approach to defining ideal conditions is that it is possible for the traffic pressure adjustment factor to have values above 1.0 when the flow rate is extremely high.

Table 23. Adjustment factor for volume level (i.e., traffic pressure) (f_v)

Traffic Volume, vpcpl	Movement Type	
	Left-Turn	Through & Right-Turn
3.0	0.953	0.947
6.0	0.971	0.961
9.0	0.991	0.974
12.0	1.011	0.988
15.0	1.032	1.003
18.0	1.054	1.018
21.0	1.077	1.033
24.0	1.100	1.049

3.1.2 Start-up Lost Time

The start-up lost time associated with a discharging traffic queue varies with its saturation flow rate. More specifically, start-up lost time increases with saturation flow rate because it takes more time for the discharging queue to attain the higher speed associated with the higher saturation flow rate. The recommended start-up lost times corresponding to a range of saturation flow rates are provided in Table 24. These values can also be computed using the following equation:

$$l_s = -4.54 + 0.00368 s_f \geq 0.0 \quad (45)$$

where:

l_s = start-up lost time, sec; and

s_f = saturation flow rate per lane under prevailing conditions ($= s/N$), vphgpl.

The recommended start-up lost times are applicable to left, through, and right-turn movements.

Table 24. Start-up lost time (l_s)

Saturation Flow Rate, vphgpl	Start-up Lost Time, sec
1,400	0.61
1,500	0.98
1,600	1.35
1,700	1.71
1,800	2.08
1,900	2.45
2,000	2.82
2,100	3.18

3.1.3 Clearance Lost Time

The time lost at the end of the phase represents the portion of the change interval that is unavailable for traffic service. This time is intended to provide a small time separation between the traffic movements associated with successive signal phases and, thereby, promote a safe change in right-of-way allocation. Clearance lost time at the end of the phase can be computed as:

$$l_e = Y + R_c - g_Y \quad (46)$$

where:

l_e = clearance lost time at end of phase, sec;

Y = yellow interval, sec;

R_c = red clearance interval, sec; and

g_Y = effective green extension into the yellow interval, sec.

For speeds in the range of 64 to 76 km/h and volume-to-capacity ratios of 0.88 or less, the average green extension is 2.5 seconds. This average value is recommended for use in most capacity analysis. More appropriate values can be computed from Equation C-51 in Appendix C when the speeds are outside the range of 64 to 76 km/h or when X for the analysis period exceeds 0.88.

3.1.4 Lane Utilization

Drivers do not distribute themselves evenly among the traffic lanes available to a lane group. As a consequence, the lane of highest demand has a higher volume-to-capacity ratio and the possibility of more delay than the other lanes in the group. The HCM (3) recognizes this phenomena and offers the use of a lane utilization factor U to adjust the lane group volume such that it represents the flow rate in the lane of highest demand. The adjusted lane group volume can be computed as:

$$v' = v_g' U \quad (47)$$

where:

v' = demand flow rate for the lane group, vpc;

v_g' = unadjusted demand flow rate for the lane group, vpc; and

U = lane utilization factor for the lane group.

There are two reasons for an uneven distribution of traffic volume among the available lanes. One reason is the inherent randomness in the number of vehicles in each lane. While drivers may prefer the lesser-used lanes because of the potential for reduced travel time, they are not always successful in getting to them for a wide variety of reasons (e.g., unable to ascertain which lane is truly lowest in volume, lane change prevented by vehicle in adjacent lane, driver is not motivated to change lanes, etc.). Thus, it is almost a certainty that one lane of a multi-lane lane group will have more vehicles in it than the other lanes, during any given cycle.

The second reason for an uneven distribution of traffic volume among the available lanes is driver desire to "preposition" for a turn maneuver at a downstream intersection. This activity commonly occurs on the arterial cross street at interchange ramp terminals and at any associated closely-spaced intersections. This prepositioning causes drivers to concentrate in one lane at the upstream ramp terminal or intersection in anticipation of a turn onto the freeway at the downstream ramp terminal.

Random Behavior Model. The lane utilization factor for an intersection lane group that does not experience prepositioning (e.g., at an isolated intersection) is dependent on the degree of randomness in the group's collective lane-choice decisions. In addition, it is also based on the group's traffic volume on a "per cycle" basis. In general, lane utilization is more uneven for groups with low demands or a high degree of randomness or both. The recommended lane utilization factors for random lane-choice decisions are provided in Table 25. These values can also be computed using the following equation:

$$U_r = 1 + 0.423 \left(\frac{N-1}{2 v'} \right) + 0.433 N \sqrt{\frac{N-1}{2 v'}} \quad (48)$$

where:

U_r = lane utilization factor for random lane-choice decisions;

v' = demand flow rate for the lane group, vpc; and

N = number of lanes in the lane group.

The recommended lane utilization factors are applicable to left, through, and right-turn lane groups.

Table 25. Lane utilization factors for lane groups with random lane choice (U_r)

Lane Group Flow Rate, vpc	Number of Lanes in the Lane Group			
	1	2	3	4
5.0	1.00	1.32	1.67	2.08
10.0	1.00	1.22	1.45	1.74
15.0	1.00	1.17	1.36	1.59
20.0	1.00	1.15	1.31	1.51
25.0	1.00	1.13	1.28	1.45
30.0	1.00	1.12	1.25	1.41
35.0	1.00	1.11	1.23	1.38
40.0	1.00	1.10	1.22	1.35

Prepositioning Model. The lane utilization factor for interchanges and associated closely-spaced intersections can be strongly influenced by drivers prepositioning for downstream turns. The magnitude of the effect is largely a site-specific characteristic, depending on the number of vehicles turning left or right at the downstream intersection (or ramp terminal). If this information is available (such as from a turn movement count where downstream destination is also recorded), the lane utilization factor can be computed as:

$$U_p = 1.05 \frac{\text{Max}(v'_{dl}, v'_{dr}) N}{v'} \quad (49)$$

where:

U_p = lane utilization factor for prepositioning;

v'_{dl} = number of vehicles in the subject lane group that will be turning left at the downstream intersection, vpc;

v'_{dr} = number of vehicles in the subject lane group that will be turning right at the downstream intersection, vpc; and

$\text{Max}(v'_{dl}, v'_{dr})$ = larger of v'_{dl} and v'_{dr} .

Lane Utilization Factor. The possibility of prepositioning must be evaluated to determine whether to use Equation 48 or 49 to estimate the lane utilization factor U . This possibility can be determined from the following test:

$$\begin{aligned} \frac{\text{Max}(v'_{dl}, v'_{dr})}{v'} &> \frac{1}{N} : \text{prepositioning} \\ \frac{\text{Max}(v'_{dl}, v'_{dr})}{v'} &\leq \frac{1}{N} : \text{no prepositioning} \end{aligned} \quad (50)$$

Based on the outcome of this test, the lane utilization factor is computed as:

$$U = \begin{cases} U_p & : \text{prepositioning} \\ U_r & : \text{no prepositioning} \end{cases} \quad (51)$$

The distance to the downstream intersection could effect the propensity of drivers to preposition into their desired lane. In recognition of this effect, the test equation is recommend only for intersections located in interchange areas or near other closely-spaced intersections where the inter-signal distances are less than 300 meters.

3.2 INTERCHANGE CAPACITY ESTIMATION METHODOLOGY

3.2.1 Signal Timing

A two-level signalized interchange is essentially composed of two closely spaced intersections (except for a SPUI) connected by an internal arterial link. As shown in Figure 34, each intersection will have an arterial input phase (*a*), and most will have a ramp input phase (*b*) together with an arterial left turn phase (*c*). Elapsing cycle time is upward in the example phase sequences depicted in Figure 34. All diamond interchanges and two-quadrant parclo will have all three phases per intersection in some configuration. Some four-quadrant parclos (Parclo BB) do not need the ramp phase (*b*) and others (Parclo AA) not the arterial left turn phase (*c*). Thus, for most interchange cases, three separate protected phases serve each ramp terminal. In 1973, Munjal (12) graphically examined the three critical conflicting phases at each intersection of a diamond interchange. This phase notation was adopted by the widely used diamond interchange computer program, PASSER III, developed by Texas Transportation Institute in 1977 (13).

Almost all signal timing plans used at two-level interchanges today can be described by using the *a:b:c* phase sequence in four combinations together with a related signal offset between the two intersections. Table 24 illustrates these phase sequence combinations. Various phase overlap combinations are also possible, depending on the interchange type and signal offset.

Table 24. Basic Signal Phase Sequences at Interchanges

Phase Combination	Left-Side Intersection	Right-Side Intersection	Signal Sequence
1	<i>abc</i>	<i>abc</i>	Lead-Lead
2	<i>abc</i>	<i>acb</i>	Lead-Lag
3	<i>acb</i>	<i>abc</i>	Lag-Lead
4	<i>acb</i>	<i>acb</i>	Lag-Lag

Figure 35 presents typical signal timing plans for some common interchange types for illustrative purposes. The four-phase strategy is depicted for diamond interchanges. As can be seen, partial cloverleaves (parclos), in contrast to traditional diamond interchanges, provide some application variation but do not change the basic concepts. Two-quad parclos have three phases per intersection; whereas, four-quad parclos have only two phases per side, deleting the ramp Phase *b* (Parclo BB) or arterial left turn Phase *c* (Parclo AA). The two-quad parclos may have Phase *c* in the inbound direction to the interchange (Parclo AA) or in the outbound direction (Parclo BB). Moreover, the ramp phases (Phase *b*) may be in advance of the cross street (Parclo AA), as in a conventional diamond, or beyond (Parclo BB). Parclos provide one distinguishing phasing difference to diamonds in that right-turn (sometimes free) signal overlaps are common. Single-point urban interchanges (SPUI) basically employ a conventional intersection phasing sequence, using

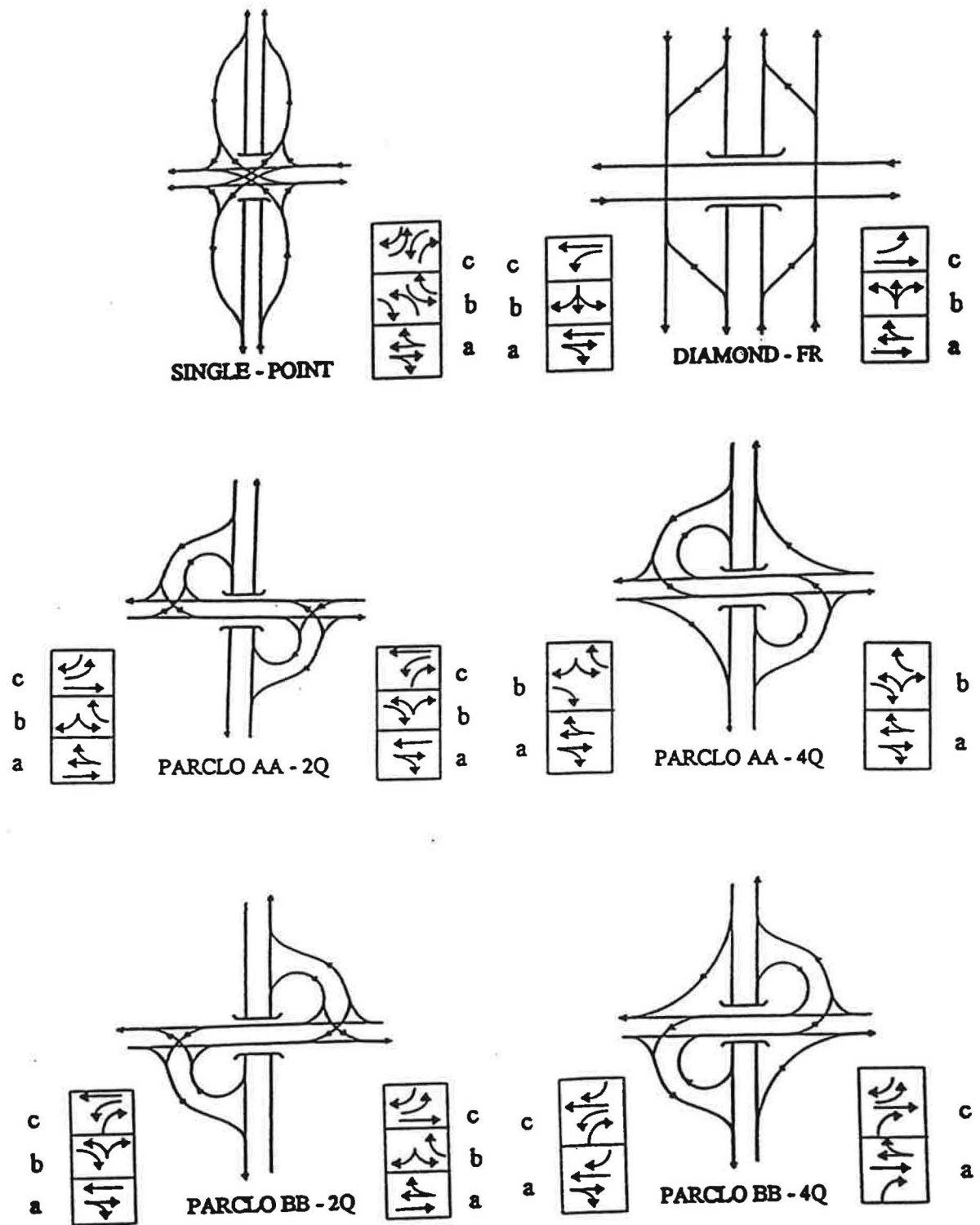
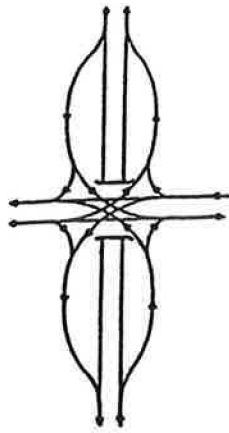
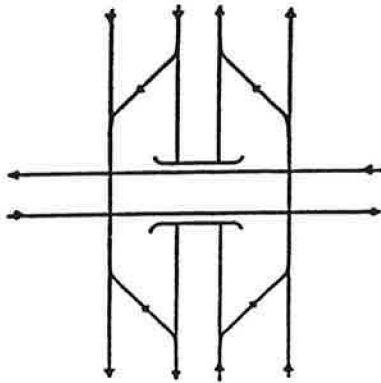
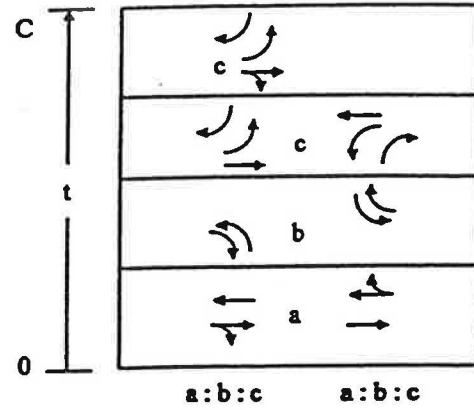


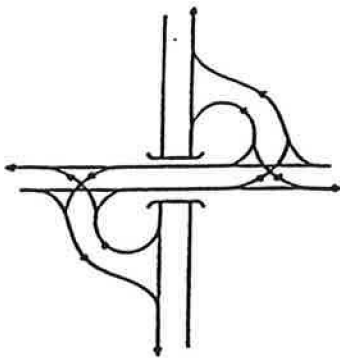
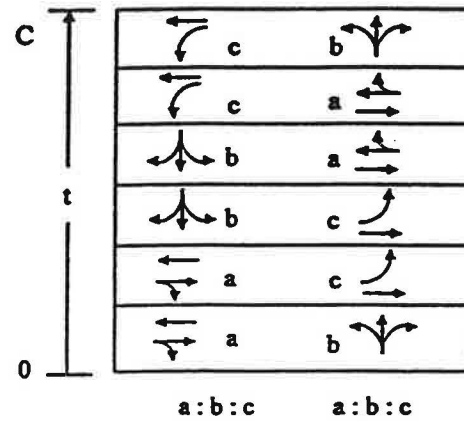
Figure 34. Interchange Signal Phasing.



SINGLE - POINT



DIAMOND - FR



PARCLO BB - 2Q

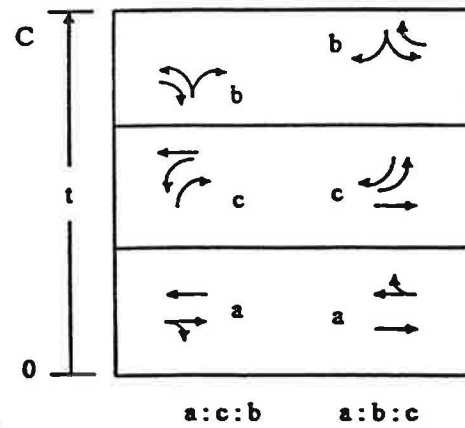


Figure 35. Interchange Signal Phase Sequences.

dual-ring operations on the crossing arterial and a single phase for the ramps. Longer phase lost times arise due to the longer clearance intervals generally employed to clear the large intersection area. Cross-street right turns have two protected phases, whereas, the ramps have only one protected phase. Thus, the ramp right-turn capacity at a SPUI may be less than at a diamond or parclo, and it should be carefully analyzed for queue spillback into the adjacent left turn lane on the ramp.

Traffic flow through an interchange depends on the signal timing, traffic mix and geometric features. Throughput depends on the capability to first enter the interchange and then exit the facility. The more restrictive case is usually entry into the interchange, although cross street left turn lanes can become overloaded under some traffic patterns. Seldom is the arterial outbound phase a restraint unless it is blocked or impeded by queue spillback from a closely-spaced downstream intersection. Operations at some two-quadrant parclo may be an exception.

3.2.2 Independent Signal Operations

At low traffic volumes and/or wide intersection spacings, the two intersections (u, d) could be operated independently of each other. Let ϕ be the total phase duration (green + yellow + red clearance intervals) such that the effective green, g , is $g = \phi - l$ where $l = l_s + l_e + l_b$ is the total phase lost time. At one intersection, u , the three conflicting phases (a, b, c) must add to one cycle

$$\phi_{ua} + \phi_{ub} + \phi_{uc} = C_u \quad (52)$$

and at the second intersection, d , the sum of three conflicting phase times may add to another cycle

$$\phi_{da} + \phi_{db} + \phi_{dc} = C_d \quad (53)$$

subject to each phase, m , satisfying its minimum green requirements $\phi_m \geq \text{Min. } \phi_m$.

As long as the two signals operate independently, cycle times can be different to accommodate variable traffic demands, green splits can be provided without constraints to better satisfy those demands, and capacities are at a maximum. Capacity principles described in Chapter 9 - Signalized Intersections of the HCM would apply, as follows. For a representative intersection and cycle, C_i , the sum of critical phases is

$$\phi_{ia} + \phi_{ib} + \phi_{ic} = C_i \quad (54)$$

which is equivalent in time to

$$(g + l)_{ia} + (g + l)_{ib} + (g + l)_{ic} = C_i \quad (55)$$

Letting L_3 equal the sum of the three lost times per phase, the total effective green time per cycle is

$$g_{ia} + g_{ib} + g_{ic} = C_i - L_3 \quad (56)$$

And since the v/c ratio for a phase or related lane group " m " is given by

$$X_{im} = \left(\frac{vC}{sg} \right)_m \quad (57)$$

Solving for the effective green, g_m and substituting into Equation 56 for the available effective green time, yields a more general expression of variables

$$\left(\frac{vC}{sX} \right)_{ia} + \left(\frac{vC}{sX} \right)_{ib} + \left(\frac{vC}{sX} \right)_{ic} = C_i - L_3 \quad (58)$$

Dividing Equation 58 by the cycle results in the fundamental capacity equation for the individual intersections at the interchange of

$$\left(\frac{v}{sX} \right)_{ia} + \left(\frac{v}{sX} \right)_{ib} + \left(\frac{v}{sX} \right)_{ic} = \frac{C_i - L_3}{C_i} \quad (59)$$

Letting $Y = v/s$ be the flow ratio of demand flow to saturation flow, then

$$\left(\frac{Y}{X} \right)_{ia} + \left(\frac{Y}{X} \right)_{ib} + \left(\frac{Y}{X} \right)_{ic} = \frac{C_i - L_3}{C_i} \quad (60)$$

Defining a new term, the pie ratio for a critical phase m , to be

$$\pi_{im} = \left(\frac{Y}{X} \right)_{im} = \left(\frac{v}{sX} \right)_{im} \quad (61)$$

which is the proportion of the available green time needed (used) to serve the demand volume for a given geometry and degree of saturation, then the total intersection pie ratio for each intersection, i , π_i , must equal the proportion of the cycle time available for moving traffic

$$\pi_i = \sum_{m=1}^{m=3} \pi_{im} = \frac{C_i - L_3}{C_i} = 1 - \frac{L_3}{C_i} \quad (62)$$

For planning and design purposes, the fundamental capacity equation (Equation 59) may be solved for critical service volumes for given average/design assumptions. First, a signal timing strategy is usually assumed that provides equal v/c ratios for all critical phases such that

$$X_{ia} = X_{ib} = X_{ic} = X_{im} = X_i \quad (63)$$

following the well-known Webster strategy, which also tends to minimize intersection delay. Thus, for equal degrees of saturation X_{im} for all critical movements m , $X_{im} = X_i$, Equation 59 becomes

$$\frac{1}{X_i} \left(\frac{v}{s} \right)_{ia} + \frac{1}{X_i} \left(\frac{v}{s} \right)_{ib} + \frac{1}{X_i} \left(\frac{v}{s} \right)_{ic} = 1 - \frac{L_3}{C_i} \quad (64)$$

so that for the three critical phases at intersection i , the total intersection flow ratio is

$$Y_i = \sum_{m=1}^{m=3} Y_{im} = \sum_{m=1}^{m=3} \left(\frac{v}{s} \right)_{im} = X_i \left(1 - \frac{L_3}{C_i} \right) \quad (65)$$

For planning and design purposes, it is also convenient to work with locally adjusted average saturation flow values, and per lane volumes, to estimate the resulting flow ratios as

$$Y_{im} = \left(\frac{v}{s} \right)_{im} = \frac{1}{s_o \prod_j (f_j)} \left(\frac{v}{N} \right)_{im} \quad (66)$$

The equivalent flow ratio on a per lane basis (l) is

$$Y_{ilm} = \frac{1}{s_{ilm}} \left(\frac{v}{N} \right)_{ilm} = \frac{v_{ilm}}{s_{ilm}} \quad (67)$$

such that the sum of the equivalent critical lane flow ratios are

$$\frac{v_{ila}}{s_{ila}} + \frac{v_{ilb}}{s_{ilb}} + \frac{v_{ilc}}{s_{ilc}} = X_i \left(1 - \frac{L_3}{C_i} \right) \quad (68)$$

Assuming that $s_{ila} = s_{ilb} = s_{ilc}$ and that v_{ilb} and v_{ilc} have been adjusted slightly to equivalent through volumes (CV_{ilb} , CV_{ilc}) to keep the initial flow ratios the same, then the sum of the "equivalent through vehicle" critical lane service volumes for a given X_i would be

$$CV_{il} = \sum_{m=1}^{m=3} CV_{ilm} = s_{ila} X_i \left(1 - \frac{L_3}{C_i} \right) \quad (69)$$

The average allowable service flow per lane per critical phase, P_{il} , would be

$$P_{il} = \frac{CV_{il}}{3} = s_{ila} X_i \left(\frac{1}{3} - \frac{l}{C_i} \right) \quad (70)$$

where l is the average lost time per critical phase. The "equivalent through vehicle" basis can represent passenger cars only, or the average fleet mix for the locale by adjusting s_{ila} accordingly. Assuming that $s_{ila} = 1800$ vphgpl, $X_i = 1.0$ for capacity flow, $l = 4.0$ sec/phase, and $C_i = 100$ sec; then from Equation 69 the total critical lane capacity flow for intersection i would be

$$CV_{il} = 1800 \times 1.0 \left(1.0 - \frac{3(4)}{100} \right) = 1584 \text{ vphpl} \quad (71)$$

and the resulting average critical lane capacity per phase is

$$P_{il} = 1800 \times 1.0 \left(0.333 - \frac{4}{100} \right) = 528 \text{ vphpl} \quad (72)$$

Other critical lane service volumes can be calculated for the degree of saturation selected, or vice versa to conduct an operational analysis of the existing situation using Equation 69.

3.2.3 Coordinated Intersections

Traffic signals at most interchanges are coordinated to improve overall traffic operations because the intersections are usually closely spaced and traffic volumes are often high. Under these conditions, coordination generally improves operational reliability and reduces internal queuing, queue spillback into upstream intersections, and the threat of operational gridlock. At lower traffic volumes, cross arterial progression can also be provided in most cases to reduce queuing. The overall quality of operations depends on the features of the signal system deployed, the progression provided, and other factors.

Coordinated signal operation implies that the two intersections no longer can operate independently. The first coordination constraint applied usually is that the two cycles, C_u and C_d , must be the same length at a particular point in time. This cycle constraint is true whether the interchange is controlled by one controller unit or by two. Thus

$$C = C_i = C_u = C_d \quad (73)$$

Delays and queuing incurred on the two external approaches to the lower-volume intersection will increase due to operations at a suboptimal (higher) cycle, assuming that the interchange would operate at the cycle of the higher-volume intersection. Signal coordination within the interchange can almost eliminate outbound delays, but arterial coordination is required to mitigate arterial approach delays incurred on the inbound Phase α .

Cycle timing is critical to interchange operations for many complex and interrelated reasons. Capacity analysis must recognize this central control parameter. The first consideration in establishing cycle time is to determine whether interchange control is coordinated with the cross street arterial or not. If yes, then cycle times will be fixed for various coordination time periods. The interchange's local control may be coordinated pretimed or coordinated (semi) actuated, but the system cycle length is the same in either case.

The second consideration in establishing cycle time, given that cross arterial coordination is not effected, is whether the interchange's controller(s) is pretimed or coordinated (semi) actuated. Local interchange coordination is presumed. Pretimed systems are fixed to prescribed durations regardless of current traffic demands and local capacity provided. Presumably forecasted traffic demands and estimated roadway capacity were considered in the initial selection of cycle times, but traffic conditions may have changed and operations deteriorated. For locally coordinated actuated control, field measurements are highly recommended over unproven analytical models or engineering judgement. Actuated control immediately responds to existing queues and random volumes, but tends to be unstable in congested conditions. Unpredictable green splits and extremely long cycles may result which makes modeling extremely difficult.

Another critical operational feature affecting interchange traffic operations is whether the green splits at the two intersections depend on one another. If two separate controllers are used (or

independent rings are provided using one controller), then capacity estimates are best given by Equation 69. If phasing is dependent (e.g. only one controller is used with dependent phases), then more complex demand analysis is required consistent with the type of interchange and phasing used. Operations at "four-phase" diamonds is a classic example, but other overlap timings exist.

3.2.4 Coordinated Diamond Interchanges

A popular diamond interchange signal timing strategy is "four-phase with two overlaps." In addition to having many of the above timing features (e.g., same cycle, a fixed sequence, and a fixed offset), this strategy also provides quality platoon progression for the arterial traffic in both directions of flow through the interchange by either special controller design, or by judicious signal timing. By either method, the following signal timing relationship must occur (14) for four-phase with two overlaps signalization to result

$$G_{ua} + G_{ub} + G_{da} + G_{db} = C_i + \Phi \quad (74)$$

where it is presumed that G_{ua} and G_{ub} are the thru and ramp phases, respectively, and Φ is the total interchange "overlap" for both directions of flow through the interchange (14). This operational requirement provides great progression for the arterial traffic passing through the interchange. However, the sum of the four external phases serving traffic input to the interchange is fixed for a constant cycle and does not have full flexibility to optimally adjust to all possible traffic patterns that might arise at the interchange. In addition, this constraint (Equation 74) further implies that the sum of the two internal left turns within the interchange is also fixed (14) at

$$G_{uc} + G_{dc} = C_i - \Phi \quad (75)$$

because the sum of the conflicting phases must equal two cycles (Equations 73 and 74). This fact may be a significant constraint on the optimal solution depending on the cycle time, minimum green times required, and traffic pattern being serviced (13).

The total interchange overlap is equal to twice the operational travel time between the two signals, i.e., the link travel time down and back. Either of two possible travel times may arise, depending on whether arriving arterial vehicles are assumed to be stopped at the initial stop line or arriving in an open lane in a coordinated platoon. Traditionally, the former case has been assumed which presumes stopline queuing and results in longer travel times due to the extra time needed to accelerate from a stop to the running speed. The former case generates longer overlaps. Conditions which result in minimal queuing at green onset, such as low volume levels per lane and the presence of good signal coordination along the crossing arterial, may reduce the effective overlap. A simple kinematic equation is given in the PASSER III user's manual for estimating total overlap, assuming platoon acceleration from the stop line after a 0.5 second perception-reaction time, is:

$$\Phi = 2 [0.50 + \sqrt{0.137 L}] \quad (76)$$

for an interchange ramp spacing of L meters, subject to a maximum speed of 50 km/h. Different maximum speeds and overlaps may be appropriate where conditions warrant. Effects of grades, speed limits, signal coordination, and general operations should be considered.

The computer signal timing program PASSER III, developed at Texas Transportation Institute, contains these strategies for developing optimal signal timing for all forms of two-level signalized diamond interchanges. PASSER III also contains delay/difference-in-relative offset algorithms, somewhat like TRANSYT 7F, to evaluate traffic performance for a given set of signal timings (13).

Clearly, capacity analysis for signalized two-level interchanges requires knowledge of the same traffic, geometric and signal operations variables as signalized intersections. Estimation of effective green time and saturation flow are more complex, but the same concepts apply. If the more robust definition of effective green as proposed herein is utilized for intersections and interchanges, then the major difference is in the complexity of determining the appropriate values for interchanges when queue interactions are expected due to high-volumes and closely-spaced intersections.

The efficiency of analysis of interchanges is likely to be more software sensitive for several reasons. Whereas the basic input volumes do not change for optional at-grade intersections, the equivalent turning movement volumes for interchanges vary widely as the type of interchange form changes. Diamonds and parclo have major differences in intersection-equivalent turning movement volumes which should be assessed efficiently. Analysis of the effects of queue spillback is critical to an accurate assessment of closely-spaced, high-volume signalized interchange operations. These assessments become more feasible with the use of interactive computer software.

3.3 RECOMMENDED CHANGES TO THE HIGHWAY CAPACITY MANUAL

Operational analysis results in the determination of capacity and level of service for each lane group at the interchange and for the total interchange. Data must provide a description of the traffic volumes, geometrics, signal timings, and environmental conditions. An existing or proposed interchange may be analyzed, with data forecasting assumptions made as necessary. The quality of the analysis being highly dependent on the quality of the forecast. Because operational analysis of signalized interchanges is quite complex, its description has been divided into six distinct modules, as depicted in Figure 35 and described as follows:

3.3.1 Input Module

All required information and data upon which subsequent computations are based is defined in the input module. This module includes all necessary data on interchange geometry, traffic volumes, signalization, lane assignments, and environmental conditions. It is used to provide a convenient summary for the remainder of the operational analysis.

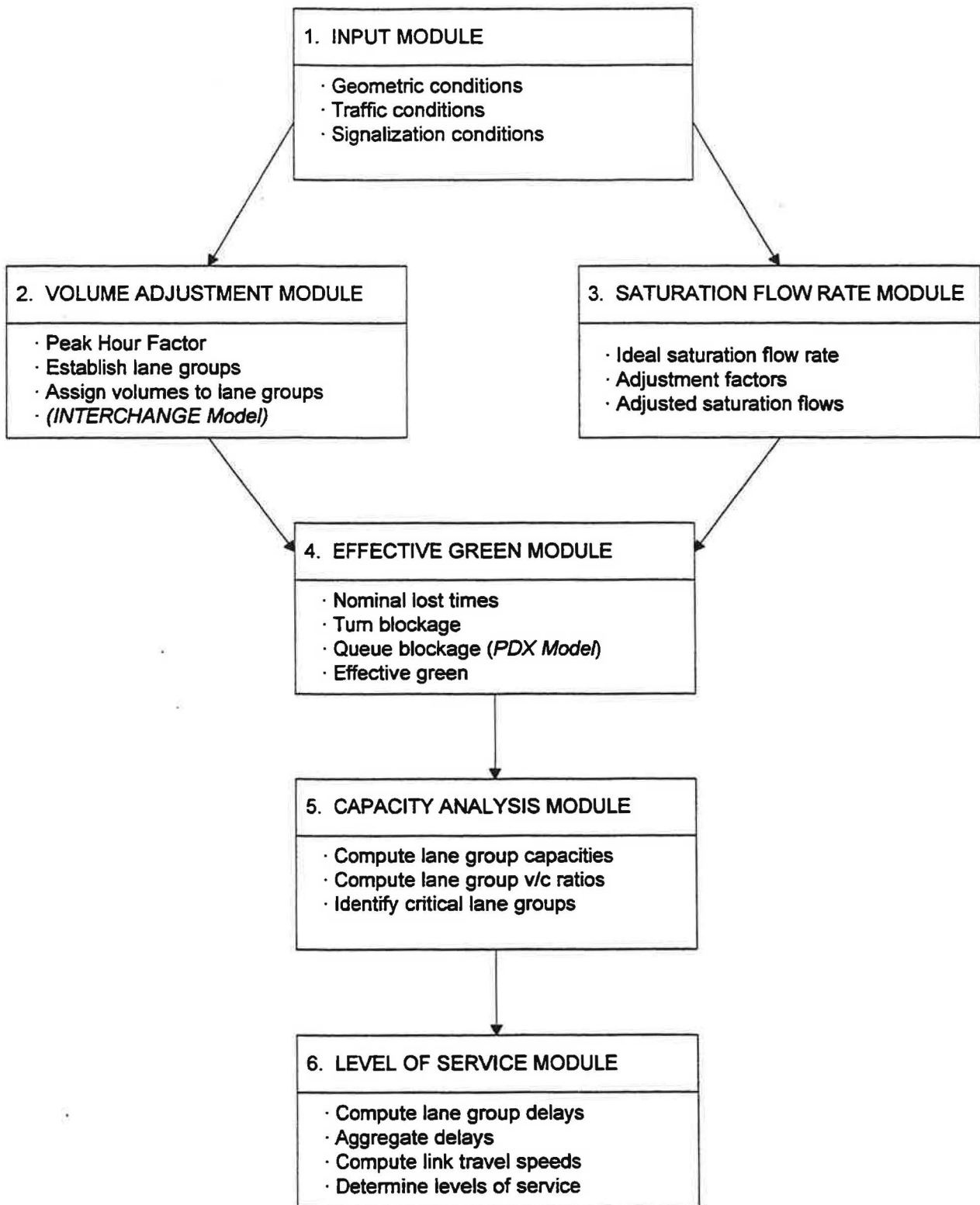


Figure 35. Recommended interchange operational analysis procedure.

3.3.2 Volume Adjustment Module

Vehicular demand volumes are generally stated in terms of vehicles per hour for a peak hour. The volume adjustment module converts these hourly traffic volumes to flow rates (in vph) for the peak 15-min analysis period. Any special traffic control assignment of turning movements to lanes should be included in the definition of lane groups in this module. Should a critical lane analysis be desired, lane utilization factors could be included at this step in the analysis. The proposed new computer software "INTERCHANGE" can aid in the preparation of input volumes for alternative interchange forms (15). This software can currently process ten interchange forms plus an at-grade intersection. The architecture of the data base is depicted in Figure 36. Once any one form is input, all other data bases are automatically available by selection. Figure 37 presents an example screen of this model. Section 3.4 below expands our vision of INTERCHANGE and Appendix F further describes the features, applications, and proposed enhancements of the software.

3.3.3 Saturation Flow Rate Module

The saturation flow rate is computed for each of the lane groups established for analysis. The flow rate is based upon the adjustment of an "ideal" saturation flow rate (recommended to be 2,000 pcphgpl) to reflect prevailing conditions. The saturation flow rate should reflect the nature of the flow only when flow is actually occurring. Blockages to flow during the related signal phase should be treated as extra lost time, as determined below, and not as impediment factors as is currently performed for permitted left turns in the 1994 HCM (3).

The saturation flow rate is the flow in vehicles per hour that could be accommodated by the lane group assuming that the green phase was always serving the lane group's traffic during the hour. Computations begin by selecting an "ideal" saturation flow rate, recommended to be 2,000 passenger cars per hour of effective green time per lane (pcphgpl), and then adjusting this value to prevailing conditions which may not be ideal. The recommended saturation flow model is:

$$s = s_o N f_w f_{HV} f_g f_p f_{bb} f_v f_{RT} f_{LT} f_D \quad (77)$$

where:

s	=	saturation flow rate for lane group under prevailing conditions, vphg;
s_o	=	saturation flow rate per lane under ideal conditions (2,000), pcphgpl;
N	=	number of lanes for the lane group;
f_w	=	adjustment factor for average lane width;
f_{HV}	=	adjustment factor for heavy vehicles;
f_g	=	adjustment factor for approach grade;
f_p	=	adjustment factor for parking;
f_{bb}	=	adjustment factor for bus blockage;
f_v	=	adjustment factor for volume (traffic pressure) level;
f_{RT}	=	adjustment factor for right turns (and turn radius); and
f_{LT}	=	adjustment factor for left turns (and turn radius).
f_D	=	adjustment factor for distance to downstream queue at green onset.

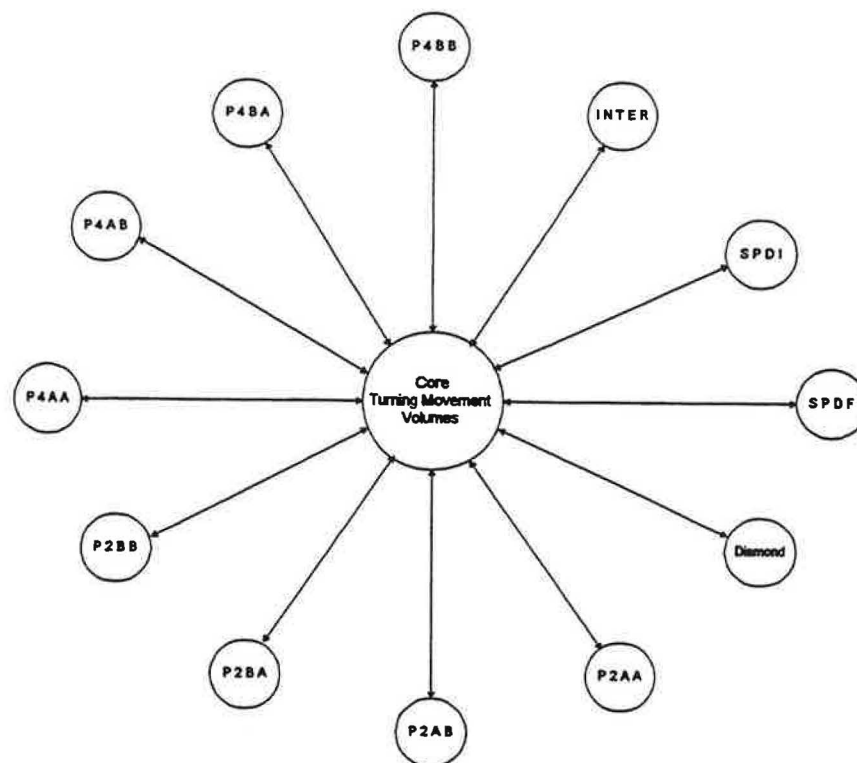


Figure 36. Database architecture of INTERCHANGE software.

INTERCHANGE

File

Input Screen

Turning Movement	Number of Vol.	Sat. Lanes	Green Time	Red Time
			sec	sec
SR	500			
SL	100			
ER	80			
ET	180			
EL	200			
EU	10			
ETR	200			
WT	600			
WR	200			
WL	600			
WU	10			
NB	300			
NL	200			

Choose One

- ☐ Diamond Interchange
- ☒ Parclo 2-Quad AA
- ☐ Parclo 2-Quad AB
- ☐ Parclo 2-Quad BA
- ☐ Parclo 2-Quad BB
- ☐ Parclo 4-Quad AA
- ☐ Parclo 4-Quad AB
- ☐ Parclo 4-Quad BA
- ☐ Parclo 4-Quad BB
- ☐ Single-Point Diam.
- ☐ Core Intersection

Parclo 2AA

Turning Movement Volumes

Directions

Fill in turning movement volumes on form

Max

Figure 37. INTERCHANGE sample input screen for parclo 2-quad AA.

3.3.4 Effective Green Module

In order to more accurately calculate capacity, queue spillback, and traffic delay during signal operations, the effective green time when platoon motion occurs at the stopline is determined. All components of signal phase lost time are identified and analyzed. These include those lost times related to nominal platoon start-up and phase clearance, intersection geometry, opposed turn queue blockages, and queue spillback during high-volume conditions. Moreover, the traditional platoon start-up lost time has been found to relate to the saturation flow found in the prior module. The resulting output from this module is the "effective green" of the individual lane groups when flow occurs at the adjusted saturation flow rate determined above. The effective green is then determined from the following relationship:

$$g = [G + Y + R_c - (l_s + l_e)] \cap CP \quad (78)$$

or emphasizing the equivalent green signal interval, G , then

$$g = [-l_s + G + g_Y] \cap CP \quad (79)$$

where:

g	=	effective green time serving lane group, sec;
G	=	signal green interval of phase serving lane group, sec;
Y	=	signal yellow warning interval of phase serving lane group, sec;
R_c	=	all-red clearance of phase serving lane group, sec;
l_s	=	platoon start-up lost time as related to saturation flow, sec;
l_e	=	platoon clearance lost time, $Y + R_c - g_Y$, sec;
g_Y	=	effective green extension into yellow, sec; and
CP	=	clear period during cycle/phase when subject flow is unblocked, sec.

The platoon start-up lost time includes lost times due to perception-reaction time to signal onset and to incremental acceleration times needed to reach the speed at saturation flow. Thus, the platoon start-up lost time increases as saturation flow increases because it takes motorists longer to attain the higher speed associated with the higher quality of flow. For left, through and right-turn movements, the relationship is:

$$l_s = -4.54 + 0.00368 s_f \geq 0.0 \quad (80)$$

where:

l_s	=	platoon start-up lost time, sec; and
s_f	=	saturation flow per lane for the subject movement, s/N , vphgpl.

Drivers continued use of the initial yellow period as the phase change occurs produces an extension of the effective green into the yellow warning interval. This effective extension of the green is denoted as g_Y and as EU , the end use. The remainder of the yellow and red clearance ($Y_r + R_c$), commonly called the end lost time (I_e), is assumed to provide no flow as a safe phase transition must occur. Interchanges having long red clearances will likewise have large clearance lost times as all of the red clearance for the subject movement is lost to flow. The green extension into the yellow warning interval is determined from:

$$g_Y = 1.48 + 0.0144 SL + 6.40 (X - 0.88) I_X \quad (81)$$

where:

g_Y	=	extension of saturation flow into yellow interval, sec;
SL	=	approach speed limit, km/h;
X	=	phase volume-to-capacity ratio; and
I_X	=	indicator variable (1.0 if $X > 0.88$, 0.0 otherwise).

The above equation has a complicating dependency relationship. To proceed, assume that the initial X is less than 0.88 so that either $I_X = 0.0$ in Equation 81 or that $g_Y = 2.5$ seconds, the default value.

An important modification to the present HCM methodology for signalized intersections is recommended for the definition of effective green. This definitional change is in respect to the requirement that flow should be occurring across the stop line of the subject lane group during effective green. This method will result in improved capacity and delay estimation, particularly the latter. This definition is very robust in that it allows the timely determination of several potential green blockage phenomena as conditions warrant. One flow blockage case arises when permitted (opposed) left turns must wait for the opposing queue to clear before filtering across opposing through volumes in a gap-acceptance mode of operations. Appendix A describes this process and the relevant queue blockage models.

A much more complex blockage problem may arise due to queue spillback during oversaturated or near-saturated conditions. Here, both blockages and saturation flow reductions can arise, and both should be separately evaluated where possible. The Clear Period during the cycle/phase is defined in this module as the time when output flow from the stopline during green can occur. This green window or Clear Period (from blockage) must be estimated either by traffic flow theory, or by the PDX Model developed within this research effort. Criteria have been provided in Chapter 2 and Appendix D to assess the potential for queue spillback blocking and upstream intersection, but the actual process is highly complex and iterative. Some solutions may only approximate the real situation, requiring field calibration and evaluation to produce reliable results. Moreover, links which are oversaturated cannot flow at a rate faster than the output capacity of the link. The signal offset of the link allocates this limited capacity to the upstream feeder lane groups. See Equation 93 for one method for estimating the signal offset to the reference phase of interest at the upstream signal. This offset defines the start of the Clear Period at the upstream signal which has a duration equal to the duration of the downstream signal.

3.3.5 Capacity Analysis Module

Volumes, saturation flow rates, and effective green ratios are processed to compute the capacity and v/c ratios for each lane group and the critical v/c ratio for the intersections of the interchange. The equation used to evaluate the resulting v/c ratio, X , is:

$$X = \frac{v C}{s g} \quad (82)$$

where:

X	=	v/c ratio (degree of saturation) for the lane group;
v	=	arrival traffic demand (volume), vph;
s	=	saturation flow rate for the lane group, vphg;
g	=	effective green time for the serving phase, sec; and
C	=	average cycle length for the subject intersection, sec.

3.3.6 Level of Service Module

Level-of-service for interchanges is based on a combination of degree of saturation, delay and overall travel speed on the link. Delays can be aggregated for approaches and for the interchange as a whole to estimate overall impacts. A lane group must be unblocked and undersaturated for any level of service above F to be identified. Link operating speeds can be used to evaluate levels of service as per arterial systems and to better quantify level-of-service F, if desired.

Vehicular delay is recognized by the 1994 HCM as being a significant traffic performance measure and, consequently, is used as the sole criterion for the level of service provided for isolated intersections. Federal Highway Administration has sponsored a research project in coordination with the Highway Capacity Committee of TRB to provide a recommended update for HCM chapters on isolated signalized intersections (Ch.9) and on coordinated signalized arterials (Ch.11). This research has just recently been completed and is being reviewed by HCQS committee of TRB. The following is a summary of the current arterial recommendations provided by the cited researchers (11). This delay methodology is also recommended for lane groups at interchanges that are either on isolated approaches or are undersaturated links along the crossing arterial. Evaluation of oversaturated links or links experiencing spillback should be based on the overall travel speed on the link using the level of service criteria for signalized arterials (Ch.11).

Generalized Delay Model. The proposed generalized delay model for signalized intersections and arterial streets (interrupted traffic flow facilities) for a subject lane group (phase) is (11).

$$d = d_1 + d_2 + d_3 \quad (83)$$

where:

- d = average total delay per vehicle for vehicles arriving during the analysis period, sec/veh;
 d_1 = uniform delay, sec/veh;
 d_2 = incremental delay due to random arrivals and overflow queues, sec/veh; and
 d_3 = incremental delay due to non-zero queues at the start of the analysis period, sec/veh.

where the uniform, or so-called first-term, delay is

$$d_1 = 0.5 C \frac{(1 - g/C)^2}{1 - (g/C) \min(X, 1.0)} PF f_{pp} \quad (84)$$

and the second-term of delay is

$$d_2 = 900 T \left[(X-1) + \sqrt{(X-1)^2 + \frac{8 k I X}{T c}} \right] \quad (85)$$

If $X < 1$ and zero queue exists at the end of the analysis period, then

$$d_3 = \left(3600 \frac{n_i}{c} \right) \frac{0.5 n_i}{T c (1-X)} \quad (86)$$

If $X < 1$ and non-zero queue exists at the end of the analysis period, then

$$d_3 = \left(3600 \frac{n_i}{c} \right) - 1800 T (1-X) \quad (87)$$

If $X \geq 1$, or oversaturation is present, then

$$d_3 = \left(3600 \frac{n_i}{c} \right) \quad (88)$$

If $X < 1$ and zero queue exists at the start of the analysis period, then

$$d_3 = 0 \quad (89)$$

where:

C	=	average cycle length, sec;
g	=	average effective green time, sec;
X	=	degree of saturation for subject lane group;
PF	=	progression adjustment factor;
f_{pp}	=	early/late arrival platoon progression adjustment factor;
T	=	analysis period in hours, in which the model parameters are fixed;
k	=	delay parameter for given arrival and service distributions;
I	=	variance-to-mean ratio of arrivals/cycle at a point;
c	=	capacity of the lane group, veh/hr; and
n_i	=	queue at the start of the analysis period.

Isolated Intersections. Considering isolated signalized intersections for a 15-minute analysis period under pretimed control and no initial queue, which essentially describes the nominal analysis conditions of Chapter 9 of the 1994 HCM, the following default values would be used: $PF = 1.0$, $f_{pp} = 1.0$, $T = 0.25$ hours, $k = 0.50$, $I = 1.0$ and $d_j > 0$. The resulting total delay equation for a lane group for an isolated approach (or one having random arrivals) would be

$$d = d_1 + d_2 \quad (90)$$

$$d = 0.5C \frac{(1 - g/C)^2}{1 - (g/C) \min(x, 1.0)} + 225 \left[(X - 1) + \sqrt{(X - 1)^2 + \frac{16X}{c}} \right] \quad (91)$$

which is almost the same model for pretimed control as has been used in the HCM since 1985.

Signalized Arterials. The generalized delay model for interrupted flow facilities can be used for coordinated approaches by selecting appropriate values for coordinated conditions and type of traffic control. Table 25 provides the progression adjustment factor (PF) for the first term of the delay equation based on arrival type (AT) together with the early/late arrival factor (f_p) which also depends on the degree of saturation of the lane group.

For coordinated intersections, the following equation is proposed for 15-minute analysis periods

$$d = 0.5C \frac{(1 - g/C)^2}{1 - (g/C) \min(X, 1.0)} PF f_{pp} + 225 \left[(X - 1) + \sqrt{(X - 1)^2 + \frac{32 k I X}{c}} \right] \quad (92)$$

Table 25. Uniform Delay Adjustment Factors

Progression Adjustment Factor (PF)						
Green Ratio (g/C)	Arrival Type (AT)					
	AT-1	AT-2	AT-3	AT-4	AT-5	AT-6
0.20	1.167	1.083	1.000	0.917	0.833	0.750
0.30	1.268	1.143	1.000	0.857	0.714	0.571
0.40	1.445	1.222	1.000	0.778	0.555	0.333
0.50	1.667	1.333	1.000	0.667	0.333	0.000
0.60	2.001	1.500	1.000	0.500	0.000	0.000
0.70	2.556	1.778	1.000	0.222	0.000	0.000
Default R_p	0.333	0.667	1.000	1.333	1.667	2.000
$PF = (1 - P)/(1 - g/C); R_p = R/(g/C)$						
Early/Late Arrival Factor (f_{pp})						
Degree of Saturation (X)	Arrival Type (AT)					
	AT-1	AT-2	AT-3	AT-4	AT-5	AT-6
0.2	1.000	0.880	1.000	1.240	1.000	1.000
0.4	1.000	0.910	1.000	1.180	1.000	1.000
0.6	1.000	0.940	1.000	1.120	1.000	1.000
0.8	1.000	0.970	1.000	1.060	1.000	1.000
1.0	1.000	1.000	1.000	1.000	1.000	1.000
$f_{pp(x=0)}$	1.000	0.850	1.000	1.300	1.000	1.000
$f_{pp} = f_{pp(x=0)} + (1 - f_{pp(x=0)}) x$						

Source: Reference 11.

Delay Results. For isolated approaches having random arrival flow the model (Eq. 91) predicts that vehicular delay increases as volumes and resulting v/c ratio (degree of saturation) increase, as depicted in Figure 38 for an isolated intersection approach. At low volumes (v/c ratio), delays are caused primarily by vehicles arriving on red. At higher volumes nearing saturation ($X = 1.0$), some cycles fail to completely clear the number of vehicles arriving (randomly) per cycle and these vehicles, which would previously not be delayed hardly at all (because they arrived on late green), are now being delayed a full red duration plus adding to the delay of subsequent arrivals. At even higher arrival volumes that routinely exceed capacity, queuing and delay continue to increase for each cycle that oversaturation exists with theoretical delays being limited only by the ending of the oversaturation period, T . Extremely large delays are theoretically possible if adequate queue storage is available. A practical maximum X value of 1.2 is employed by the HCS software to limit the delay calculated, but no maximum delay value has been selected. NCHRP 3-47 research has shown that the maximum delay can be calculated and is related to link length. Upper bounds on delay should be calculated for the above HCM-based delay models using Equations 25 and 26, especially for links less than 300 meters long.

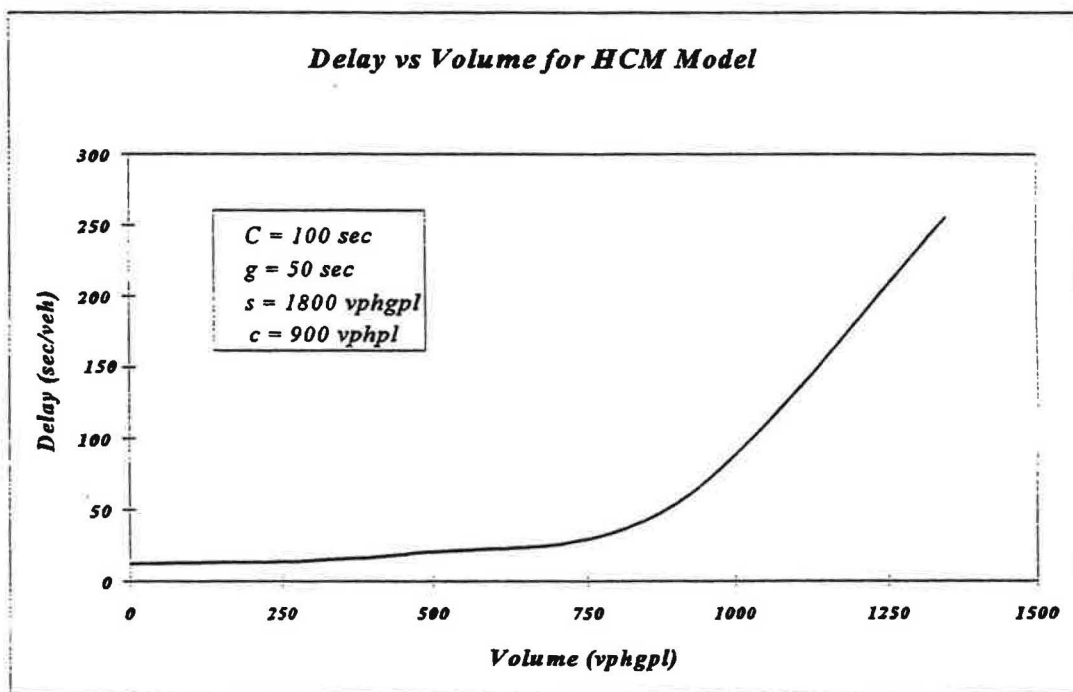


Figure 38. Proposed HCM Model for Estimating Vehicular Delay for Random Arrivals (11).

3.4 DESCRIPTION OF SOFTWARE NEEDED FOR IMPLEMENTATION

Various operational software exist which have the capability of analyzing different interchange forms; however, the analysis requires that input be manually re-entered for each interchange evaluated. Making comparisons among various interchange alternatives becomes a tedious task as the data must be manually calculated and input for every interchange type evaluated. Outputs must be saved to different files for evaluation. The development of the INTERCHANGE prototype software is seen as a solution to these problems since the software can evaluate various interchanges using only a single configuration input. Two design concepts were developed and are shown in Figure 39. One concept involves using INTERCHANGE as a data input and conversion software to be used in conjunction with existing software. The other procedural design proposes a standalone program capable of analyzing all operational aspects of the various interchanges being compared.

The left side of the diagram depicts the software design option which uses existing software to perform the analysis. In the first step, the input for one interchange form is entered which includes the turning movement volumes for one interchange form, the geometric and signalization conditions, the type of analysis requested, and all the interchange types to analyze. Once these data are entered the program performs a conversion analysis using the database conversion algorithm shown earlier in Figure 36. For each alternative form, the program produces data unique to that interchange, such as turning movement volumes and required geometry. This converted data could be then automatically input into existing software, such as HCS, PASSER-II, and TRANSYT-7F, for an operational analysis of each interchange being considered. The output from these existing software can then be viewed through INTERCHANGE to create a common output screen for ease of viewing and analyzing the performance measures. A common database of output values can be stored within INTERCHANGE to facilitate the comparisons. As a last step, the output could be re-entered as input to re-evaluate other design options.

As a standalone program, the software would perform the same input and conversion analyses; however, the analyses would be performed within the program itself. Much of the operational procedure shown in Figure 39 would now be evaluated internally. The volume adjustment module, saturation flow rate module, effective green module, capacity analysis module, and level of service would all be computed within INTERCHANGE. After performing the analysis on each interchange selected, the program could produce outputs displaying performance measures for each selected configuration individually or together for easy comparisons. The output could be tailored to produce useful graphical results for different users.

The prototype currently has the capability of converting turning movement volumes among 10 interchange configurations and one at-grade intersection. A graphical interface has also been developed to facilitate interchange comparisons and communicate the information to the user. Appendix F describes the specific features of INTERCHANGE and demonstrates its use (15).

INTERCHANGE Procedural Design

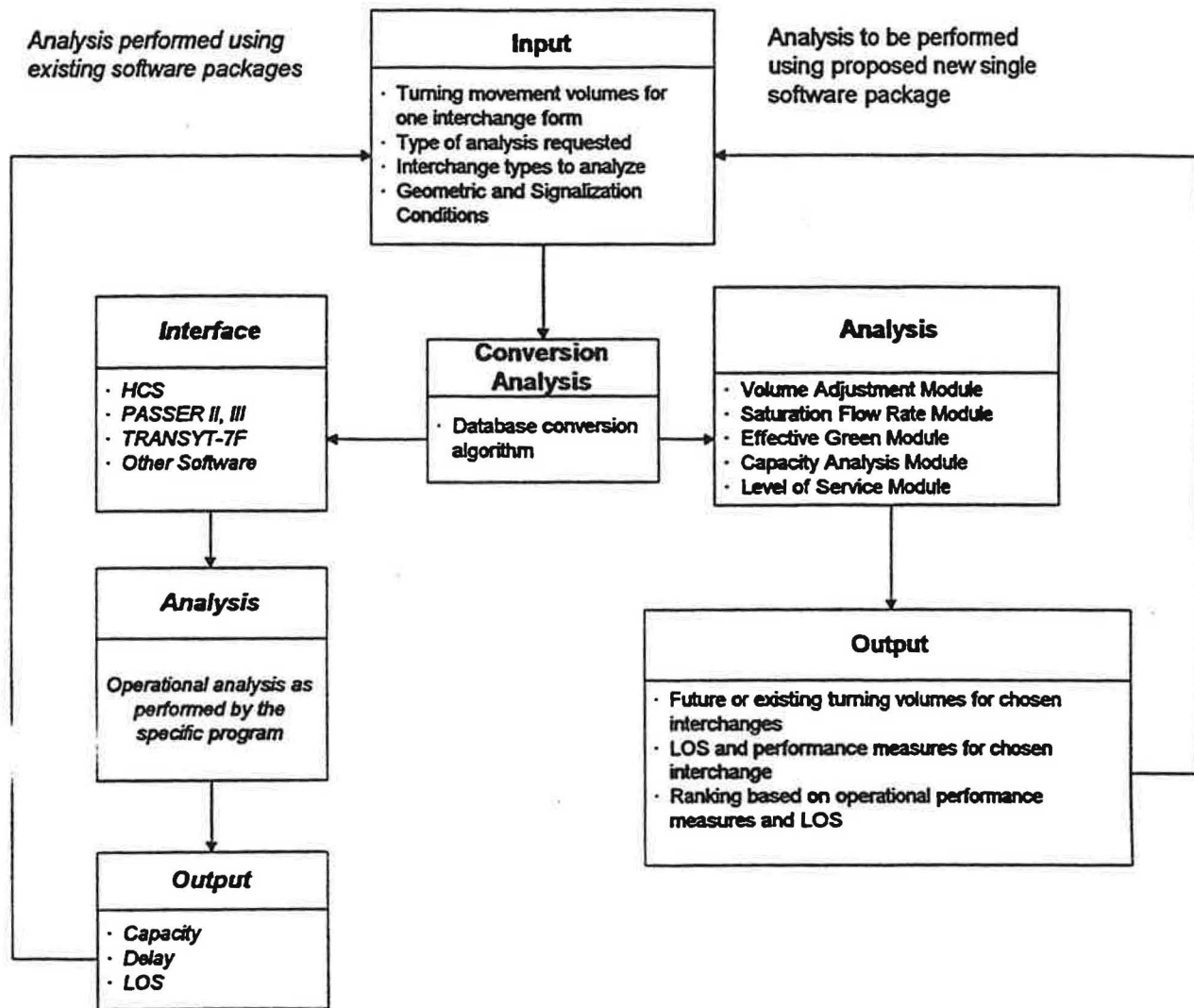


Figure 39. Flow diagram of optional procedural design for INTERCHANGE.

3.5 SIGNAL TIMING IMPROVEMENTS FOR CLOSELY SPACED INTERSECTIONS

Developing optimal, or even good, signal timing plans for signalized interchanges that connect to a high-volume signalized crossing arterial is a complex task and current technology is limited. Unintentionally poor signal timings may produce oversaturation and avoidable congestion. Even if oversaturation is unavoidable due to excessive traffic volumes, improved operations can be provided along the arterial and interchange ramps by avoiding signal plans that will be inefficient for the given situation. All aspects of the operations should be examined. Short links, say less than 150 m, are especially troublesome. Links that have a predominate input flow are subject to demand starvation, and much of the downstream green may be wasted if a poor signal plan is implemented. On the other hand, links having nearly balanced input flows, as can often happen around interchanges with high local access demands, are insensitive to some operational control measures. No unproductive constraints should be placed on the overall system to try to treat such overloaded links for all other links would probably only get worse and no appreciable benefits would likely accrue. The following models are provided as guidelines toward improving traffic signal control during these problematic situations.

Signal timing, traffic pattern, and queue storage length are principal factors of flow control during such high-volume conditions. Effects of these factors on flow were illustrated in Chapter Two. Within signal timing, signal offset is a robust and critical control variable. In this section, signal offset, θ_{ij} , is defined as the elapsed time from the start of green of the major phase at the upstream intersection carrying the highest volume of input traffic onto the link $i-j$ to the start of green of the arterial through phase at the downstream intersection. The upstream major phase will usually be the arterial input phase, but not always in cases of high-volume turning traffic. The offset described below may become "negative" meaning an early start of the downstream green, but the offset used in the signal plan must be the equivalent "modulo C" offset. Two offset controls will be described: (1) queue clearance offset and (2) storage offset. The larger of the two offsets should not be exceeded in most cases currently envisioned if peak flow efficiency is to be maintained.

3.5.1 Queue Clearance Offset

In special arterial traffic signal coordination cases where all the arterial traffic is 100% through traffic (arterial dominance), the ideal signal offset between the upstream signal and the downstream signal would be equal to the travel time over the link for the running speed of the traffic. As non-arterial traffic (assumed to be non-coordinated) becomes a larger portion of the link flow, queuing during red increases and the downstream offset should be reduced to "clear the cross-street queue" before arrival of the major platoon. One offset selection model of this process is:

$$\theta_{ij} = \frac{L}{u} - \frac{r v_r}{s} \quad (93)$$

where:

θ_{ij} = relative offset between major phases on link $i-j$, sec;
 L = length of link $i-j$, meters;

u	=	running speed of arterial link flow, mps;
r	=	effective red of downstream signal, sec;
v_r	=	downstream arrival volume on red, vph; and
s	=	saturation flow of downstream phase, vphg.

As the upstream turning volumes become a larger percentage of the total downstream volume, then v_r increases, the downstream queue grows, and the relative offset θ_{ij} should be reduced from the ideal progression offset to reduce spillback.

Using the Highway Capacity Manual's (3) concept of determining the proportion of arterial traffic arriving on the downstream green, P , then the queue clearance offset would be

$$\theta_{ij} = \frac{L}{u} - \frac{(1 - P) v C}{s} \quad (94)$$

where:

θ_{ij}	=	relative offset between major phases on link $i-j$, sec;
L	=	length of link $i-j$, meters;
u	=	running speed of arterial traffic flow, mps;
P	=	proportion of arterial traffic arriving on downstream green;
v	=	total downstream arterial arrival volume on phase, vph; and
C	=	cycle length, sec.

Offsets less than that given above would be expected produce some demand starvation. Any reduction in offset from the ideal should not exceed the green time necessary to saturate the phase. During undersaturated conditions, offsets greater than this value will produce some queue spillback on the major (arterial) phase which may result in queue spillback into the intersection and oversaturation on the upstream link if the downstream link is too short (See Equation D-15.).

3.5.2 Maximum Storage Offset

Another limit placed on signal offset adjustment that should be considered during high-volume conditions is related to the length of the link. To minimize the chances of demand starvation, the link offset should not be less than

$$\theta_{ij} = \frac{L}{u} - \frac{3600 NL}{sl} \approx \frac{L}{u} - \frac{L}{3.5} \quad (95)$$

where:

θ_{ij}	=	relative offset between major phases on link $i-j$, sec;
L	=	length of link $i-j$, meters;
u	=	running speed of arterial traffic flow, mps;
N	=	number of lanes on link $i-j$;
s	=	saturation flow of downstream phase, vphg; and
l	=	queue storage length per vehicle, about 7 m/veh.

The final selection of offset should consider many factors, including the volume level, traffic pattern and degree of saturation of the downstream signal. However, to enhance the throughput efficiency during high-volume conditions, the link offset should not be less than the larger of Equations 94 and 95 given above.

3.6 ARTERIAL WEAVING SPEED ANALYSIS METHODOLOGY

The procedure for computing the speeds of vehicles in the arterial weaving section is outlined in the flow chart shown in Figure 37. This flow chart illustrates the sequence of computations required to compute the average running speed of arterial through and weaving vehicles in the weaving section. The arterial weaving section of interest is the one created by the combination of a signalized interchange off-ramp and an adjacent, closely-spaced signalized intersection.

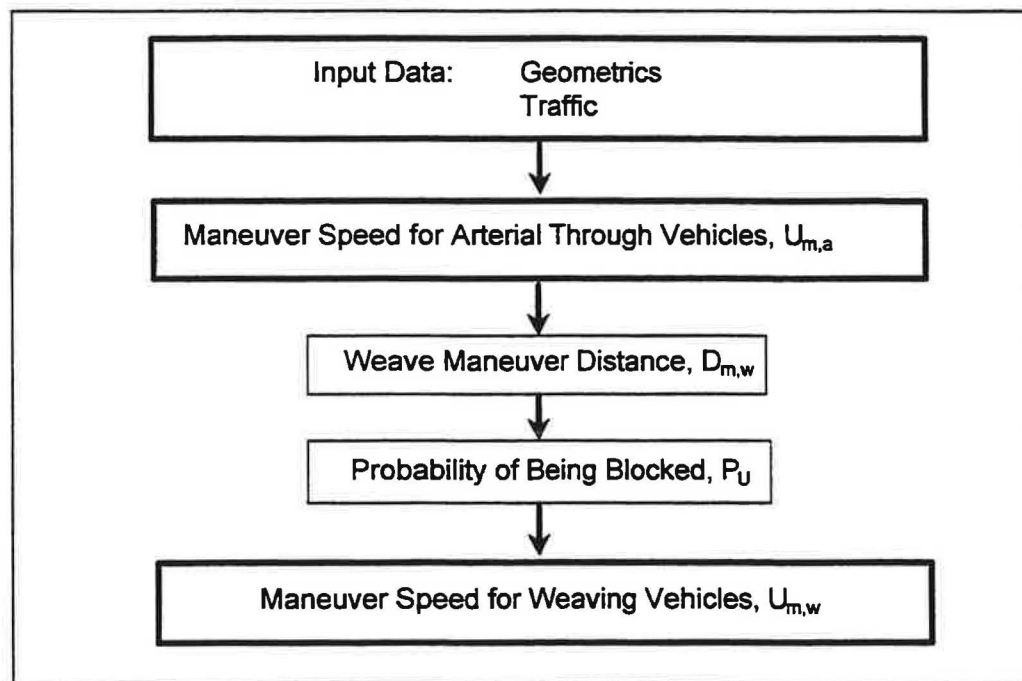


Figure 37. Flow chart to determine speeds in an arterial weaving section.

The weaving maneuver considered is the off-ramp right-turn movement that weaves across the arterial to make a left-turn at the downstream signalized intersection. The terminology "arterial weaving" is adopted in this methodology to clearly indicate that the weaving occurs on streets whose traffic flow is periodically interrupted by traffic signals, as compared to the more extensively studied weaving that occurs on uninterrupted flow facilities.

3.6.1 Maneuver Speed for Arterial Through Vehicles

Input data needed to compute the maneuver speed for arterial through vehicles includes the average arterial speed and flow rate entering the weaving section. This average speed represents the desired speed of the traffic stream when there is no weaving activity present. The arterial maneuver speed can be computed with the following equation:

$$U_{m,a} = 1.986 U_a^{0.717} e^{(-0.634 V_a/3600)} \quad (96)$$

where:

- $U_{m,a}$ = average maneuver speed for arterial through vehicles, m/s;
- U_a = average arterial speed entering the weaving section, m/s; and
- V_a = average arterial flow rate entering the weaving section, vph.

The average maneuver speed is defined as the average running speed of arterial vehicles through the weaving section. It represents the ratio of travel distance to travel time within this section. The arterial and weaving maneuver speeds are computed in a similar manner; however, they do not tend to converge for low volume conditions. This trend is due to the fact that the weaving maneuver generally has to accelerate from a stopped (or slowed) condition whereas the arterial maneuver generally enters the weaving section at speed.

3.6.2 Maneuver Speed for Weaving Vehicles

The average maneuver speed for weaving vehicles is dependent on a variety of traffic conditions and events. Traffic conditions include the average arterial speed entering the weaving section, volume of weaving vehicles, and available weaving maneuver distance. The primary traffic event is the probability of the weaving vehicle not being blocked from entering the weaving section or delayed during its maneuver.

The average arterial speed was discussed in the preceding section. The average weaving flow rate V_w represents the number of vehicles that complete the off-ramp-right-to-downstream-left-turn maneuver during the analysis period; it is a subset of the number of vehicles that enter the arterial via the off-ramp. The remaining variables require computation; equations for this purpose are described in the following paragraphs.

The available weaving maneuver distance $L_{q,w}$ represents the average length of arterial available for a weaving vehicle to complete its weave maneuver. This distance can be estimated as:

$$L_{q,w} = \frac{V_a r}{1 - \frac{V_a}{s_l N_l}} \times \frac{L_v}{3600} \quad (97)$$

where:

- $L_{q,w}$ = average length of queue joined by weaving vehicles, m.
- s_l = saturation flow rate per lane under prevailing conditions ($\approx 1,800$), vphpl;
- N_l = number of arterial through lanes in the subject direction, lanes;
- r = effective red time for the downstream intersection through movement; sec; and
- L_v = average lane length occupied by a queued vehicle (See Equation C-45.), m/veh.

The "probability of a weaving vehicle being unblocked" P_U represents the portion of time that the end of the off-ramp (i.e., the beginning of the weaving section) is not blocked by the passing of the arterial traffic stream or the spillback of a downstream queue. This probability is dependent on the length of the weaving section. A "long" weaving section is defined as a section that has sufficient length to allow a weaving driver to weave across the arterial one lane at a time. If this minimum length is not available, then the section is referred to as "short." The weave maneuver in a short section requires the simultaneous crossing of all arterial lanes (in the subject direction) using more of a "crossing" than a "lane-changing" action. The equation for estimating P_U is:

$$P_U = \left(1 - \frac{V_a}{s_l N_l} \right)^{N_l} (1 - I_L) + \left(1 - \frac{V_a}{s_l N_l} \right) I_L \geq 0 \quad (98)$$

where:

- P_U = probability of a weaving vehicle being unblocked;
- I_L = indicator variable (1.0 if $D_{m,w} > 90 (N_l - 1)$, 0.0 otherwise);
- $D_{m,w}$ = average maneuver distance for weaving vehicles ($= L_w - L_{q,w}$), m; and
- L_w = distance between the off-ramp entry point and the stop line of the downstream intersection (i.e., the length of the weaving section), m.

It should be noted that P_U is equal to 0.0 when the average queue length equals the length of the weaving section (i.e., when $D_{m,w}$ is effectively zero).

Using the probability of a weaving vehicle being unblocked and the weaving volume, the average weaving maneuver speed can be computed as:

$$U_{m,w} = 3.741 U_a^{0.408} e^{(-11.045 (1 - P_U) V_w / 3600)} \quad (99)$$

where:

- $U_{m,w}$ = average maneuver speed for weaving vehicles, m/s; and
- V_w = average weaving flow rate, vph.

Like the arterial maneuver speed, the weaving maneuver speed is defined as the average running speed of weaving vehicles through the weaving section.

3.6.3 Example Application

The following example is used to demonstrate the application of the arterial weaving speed analysis methodology. The arterial weaving section has the following attributes:

Geometric Data:

Number of arterial through lanes in the subject direction, N_t = 2 lanes
Length of the weaving section, L_w = 200 meters

Traffic Data:

Average arterial speed entering the weaving section, U_a = 12.5 m/s (45 km/h)
Average arterial flow rate entering the weaving section, V_a = 1,000 vph
Average weaving flow rate, V_w = 170 vph
Saturation flow rate per lane, s_l = 1,800 vphgpl
Effective red time for the downstream intersection, r = 50 seconds
Average lane length occupied by a queued vehicle, L_v = 7.0 m/veh

Analysis:

1. Using the average arterial speed and flow rate, the average maneuver speed for arterial through vehicles $U_{m,a}$ can be computed as 10.2 m/s.
2. Using the arterial flow rate, effective red time, saturation flow rate, and number of lanes, the average length of queue joined by weaving vehicles $L_{q,w}$ can be computed as 135 meters.
3. The average length of queue, combined with the weaving section length, yields a maneuver distance $D_{m,w}$ of 65 meters.
4. Based on the maneuver distance of 65 meters, it can be concluded that the weaving section is too short for a driver to weave by changing lanes one at a time. The driver is more likely to cross all lanes at the same time (i.e., $I_L = 0.0$)
5. For short weaving sections, the probability of a weaving vehicle being unblocked P_U is computed as 0.52.
6. The weaving flow rate and probability of blockage can be used to compute the average weaving maneuver speed $U_{m,w}$ as 8.2 m/s.

If the weaving section length were 300 meters long then it would be characterized as a "long" section. In this case, weaving through the section would likely occur one lane at a time and the weaving speed would increase to 9.1 m/s. However,, the predicted arterial through movement speed $U_{m,a}$ is unchanged (i.e., it equals 10.2 m/s).

3.7 RAMP WEAVING CAPACITY ANALYSIS METHODOLOGY

The procedure to determine the interchange ramp weaving capacity across the connecting arterial street for a specific progression factor and arterial through volume is shown in Figure 40. The flow chart illustrates the methodology to determine the ramp weaving capacity for crossing the arterial and maneuvering into a downstream left turn bay given progressed flow along the arterial. The procedures are as follows.

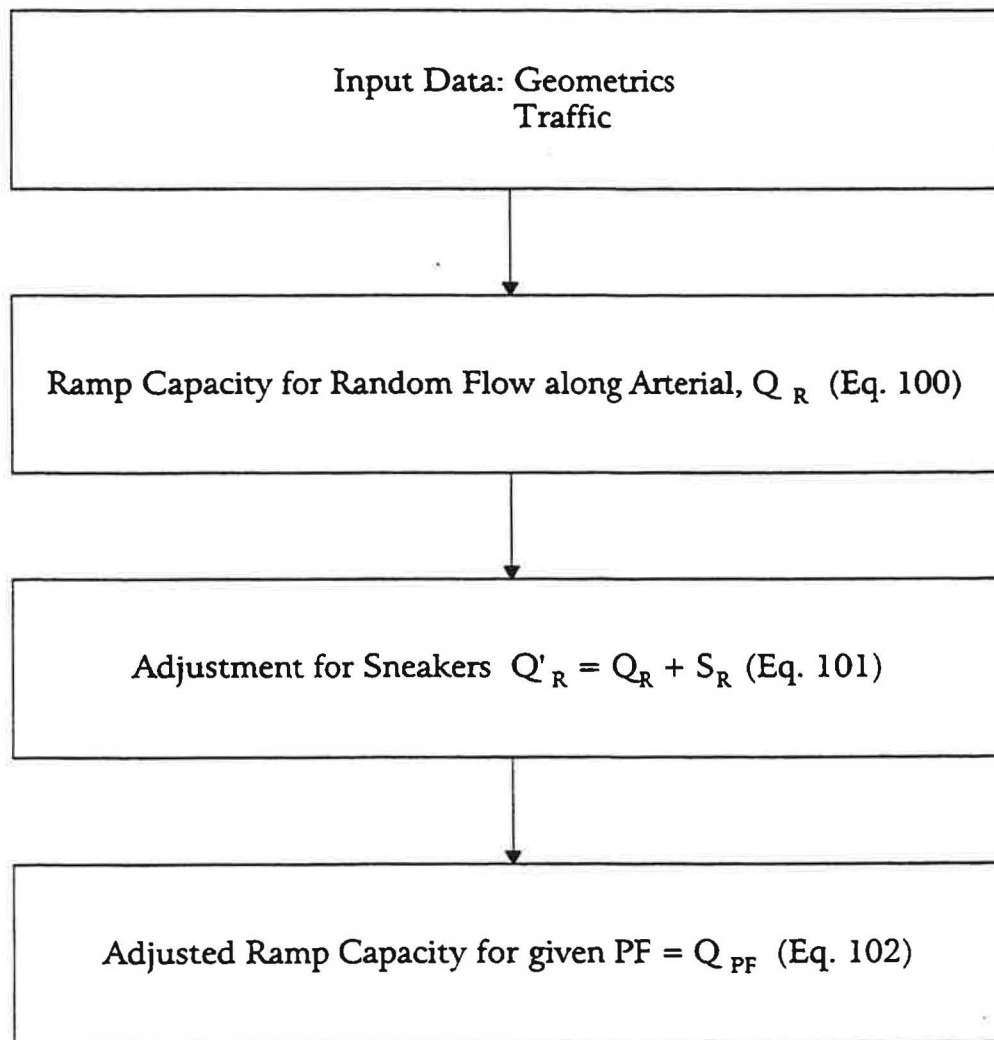


Figure 40. Flow chart to determine ramp weaving capacity.

3.7.1 Ramp Capacity During Random Flow

Data necessary to compute ramp capacity for random flow are total arterial volume, number of lanes on the arterial in the direction of merging operations, and the coefficients α , β . The data required to determine the ramp capacity for progressed flow include the number of lanes on the arterial, the required PF and v/c ratio. The number of lanes of the arterial are used to select the α and β coefficients from Table 26 which are employed in the exponential regression equation for estimating ramp capacity during random flow conditions given below.

$$Q_R = \frac{V_t e^{-\alpha V_t}}{1 - e^{-\beta V_t}} \quad (100)$$

where:

- Q_R = ramp crossing capacity, vph;
- V_t = arterial through volume, vph;
- α = coefficient of the model = $T_c / 3600$; and
- β = coefficient of the model = $H_s / 3600$.

Table 26. Coefficients of the exponential regression model

Thru Lanes	1 lane		2 lanes		3 lanes	
Coefficients	α	β	α	β	α	β
Exponential	0.00195	0.000657	0.00118	0.000574	0.00088	0.000565
R ² Value	0.9977		0.9995		0.9989	
Conversion of Coefficients	T_c	H_s	T_c	H_s	T_c	H_s
Values (sec.)	7.02	2.36	4.26	2.06	3.17	2.03

Table 26 shows the coefficients α and β of the exponential model computed for one, two and three thru lanes of arterials. The coefficients in the proposed exponential equation are accurate estimations of the TRAF-NETSIM imulated operations in terms of standard errors and their variances.

3.7.2 Adjustment for Sneakers

Additional ramp capacity may be obtained by ramp vehicles weaving across the arterial flow during the phase change intervals of the upstream signal. Simulation studies suggest that as many as three "sneakers" per phase change may cross during capacity conditions. Thus, the ramp crossing capacity adjusted for sneakers would be

$$Q'_R = Q_R + \frac{3600 S_\phi n_\phi}{C} \quad (101)$$

where:

$$\begin{aligned} Q'_R &= \text{ramp weaving capacity adjusted for sneakers, vph;} \\ Q_R &= \text{ramp weaving capacity for random flow, vph; and} \\ S_\phi &= \text{sneakers per phase change, vehicles;} \\ n_\phi &= \text{number of phases per cycle; and} \\ C &= \text{cycle length.} \end{aligned}$$

3.7.3 Adjustment for Progression

Ramp weaving capacity is affected by the quality of progression on the arterial street. The quality of progression is identified in the Highway Capacity Manual by progression factors. In order to obtain the ramp weaving capacity for different progression factors, weaving adjustment factors for progression, f_{PF} are determined and multiplied by the ramp capacity which has been adjusted for sneakers. The adjustment procedure is

$$Q_{PF} = Q'_R * f_{PF} \quad (102)$$

where:

$$\begin{aligned} Q_{PF} &= \text{ramp weaving capacity adjusted for progression, vph;} \\ Q'_R &= \text{ramp weaving capacity for random flow, vph; and} \\ f_{PF} &= \text{ramp weaving adjustment factor for progression.} \end{aligned}$$

The adjustment factors for progression range from 0.1 to 1.0 and are determined from

$$f_{PF} = 1 + 0.015 * e^{[0.0044 * V_l - 3.05 * PF]} \quad (103)$$

where:

$$\begin{aligned} f_{PF} &= \text{ramp weaving capacity adjustment factor;} \\ PF &= \text{HCM progression factor; and} \\ V_l &= \text{arterial volume per hour per lane (vphpl).} \end{aligned}$$

Note that HCM PF values exceeding 1.0 should be replaced by their complimentary value to 1.0. That is, a PF of 1.3 should be replaced by a value of 0.7 in the above equation. The progression factors are determined from Table 9.13 in the 1994 Highway Capacity Manual (3).

3.7.4 Example Application

Geometric Data:

Number of lanes on arterial = 3 lanes
Number of lanes on ramp = 1 lane
Link length between ramp and downstream intersection = 183 meters (600 feet)

Traffic Data:

Arterial Through Volume = 1500 vph
Cycle Length = 100 seconds
Lost time per Phase = 4 seconds

Analysis:

Ramp Capacity = Q_R
 α (for three lane arterial) = 0.00088 [From Table 26]
 β (for three lane arterial) = 0.000565 [From Table 26]
Sneaker Volume = [3 veh/phase * 2 phase change intervals * 36 cycles/hr]
= 216 vph

Assume the ramp capacity for a PF of 0.2 and v/c of 0.6 is to be determined.

Ramp capacity for random flow = Q_R
Substituting values in the above equation yields $Q_R = 701$ vph

Accounting for ramp vehicles weaving during the phase change interval

$$\begin{aligned} Q'_R &= Q_R + \text{Sneaker Volume} \\ &= 701 + 216 \\ &= 917 \text{ vph} \end{aligned}$$

For progressed flow along the arterial having a PF of 0.2 and a v/c of 0.6, f_{PF} can be determined from Table 20 as 1.074. Therefore,

Ramp Capacity for a PF of 0.2 and v/c of 0.6: $Q_{PF} = Q'_R * f_{PF} = 917 * 1.074 = 984$ vph

CHAPTER 4

CONCLUSIONS AND RECOMMENDATIONS

4.1 CONCLUSIONS

Mobility at a signalized (service) interchange is dependent on many factors that influence capacity. The importance of the prevailing factors may change depending on the general volume level and degree of existing congestion. Traditional prevailing factors affecting capacity include the interchange geometry, traffic mix, and signal green splits of the serving phases. During periods of oversaturation where the storage links are filled, additional non-traditional factors come into play. During oversaturation, the upstream input capacity becomes highly dependent on downstream signal timings and capacity to keep the output link clear. Upstream input capacity cannot exceed the total downstream service capacity available, and may be less than this capacity if demand starvation on short links occurs. In addition to phase capacity, signal offset during oversaturation is the most important factor in the allocation of downstream capacity to upstream phases. Also, this offset is not the offset that provides optimal progression during uncongested flow periods.

Signal timing and coordination together with signal spacing (link length) are important variables in determining the ability of the cross arterial to move traffic. Moreover, in the mesosaturation (near capacity) range of traffic conditions, the ultimate determination of the question of "oversaturation?" depends on whether potential flow impediments due to queue spillback can be sufficiently mitigated to the extent that oversaturation can be avoided.

Capacity analysis methods of signalized interchanges should be able to determine whether oversaturation will occur. However, this determination is a complex task in the mesosaturation range near capacity. These conditions not only depend on nominal volumes and phase capacities, but conditions also depend on the interactions of the variables with queue storage, link storage capacity and signal timing. Criteria have been provided that indicate the likelihood of queue spillback affecting nominal output capacity. These criteria can address probably 80 percent of the operational problems. However, close-calls can only be solved now using computer simulation models, like NETSIM, or possibly using models like the PDX Model developed in this research.

Phase capacity is dependent on the prevailing saturation flow rate and effective green time when motion can occur at the stopline. Based on this research, it is concluded that the distance to the downstream queue, the radius of the turn path, and traffic pressure have a significant effect on the saturation flow rate of a traffic movement. Specifically, saturation flow rate decreases when the distance to the downstream queue decreases and is relatively short. This effect is amplified when the signal timing relationship between the two intersections allows queue spillback to occur. As the distance-to-queue variable is bounded to a maximum value equaling the length of the downstream street segment, the effect of distance-to-queue also includes the effect of spacing between interchange ramp terminals or between a ramp terminal and a closely-spaced intersection.

Turn radius has a significant effect on the saturation flow rate of a turn-related traffic movement. Saturation flow rates are lower for turn movements with small radii than they are for turn movements with large radii. In the context of junction type (e.g., single-point urban diamond, diamond interchange, etc.), the saturation flow rates for the left-turn movements at single-point urban interchanges are more nearly equal to those of through movements because of the large turn radii associated with this interchange type.

Traffic pressure, as quantified by traffic volume per cycle per lane, has a significant effect on saturation flow rate. Traffic pressure relates to the presence of aggressive, commuter drivers in the traffic stream. Traffic volume is used as a surrogate measure of the number of these aggressive drivers in the traffic stream. Saturation flow rates of low-volume movements are much lower than those of high-volume movements because the low-volume movements have less traffic pressure.

Other factors were examined for their potential effect on saturation flow rate. These factors include: g/C ratio, junction type, downstream signal indication at the start of the upstream phase, and dual versus single left-turn lane. Of these factors, only g/C ratio was found to be correlated with saturation flow rate in a statistically significant manner. Specifically, the saturation flow rate for left-turn movements with low g/C ratios was found to be higher than the rates of similar movements with larger g/C ratios. This effect was also found in the through movements studied; however, it was much smaller in magnitude and not statistically significant. Therefore, it was determined that more research is needed to verify the significance of this trend and its magnitude before an adjustment factor for g/C effect can be recommended.

The definition of effective green time should be changed to be only the time when saturation flow can occur at the stopline for existing conditions. This definition is more robust and can be used in all operating conditions, including periods of oversaturation. Moreover, delay estimates will be improved using this new definition of effective green even during undersaturated conditions because the queue profile will be estimated more accurately than using the current HCM methodology.

The estimation of phase lost time should be improved. Start-up lost time is not a constant value; rather, it is statistically dependent on the prevailing saturation flow rate. Specifically, start-up lost time increases with increasing saturation flow rate. This paradoxical increase is due to the increased acceleration time the discharging queue requires to attain the higher speed associated with a higher saturation flow rate. The term "lost" is a bit of a misnomer in this case, as more lost time occurs as saturation flow gets better. Start-up lost times typically range from 0.61 to 3.18 seconds for prevailing saturation flow rates of 1,400 to 2,100 pcphgpl, respectively.

The average yellow warning interval used by drivers clearing the intersection at the end of the phase is termed "green extension, or end use." In the context of phase capacity, end use is equivalent to an extension of the effective green period into the yellow. The study of end use indicates that it is a relatively constant value for intersections and interchanges and that it averages about 2.5 seconds for most undersaturated conditions. Thus, this quantity can be subtracted from the signal change interval duration to estimate the lost time at the end of a signal phase.

Lane use is almost always uneven (or unbalanced) in intersection lane groups. The degree of this imbalance is expressed in terms of the lane utilization factor. The lane utilization factor varies depending on the nature of drivers' lane-choice decisions (i.e., to minimize travel time or to preposition for a downstream turn). Lane utilization factors based on travel time minimization tend to be subject to randomness in the lane-choice decision process. The factors stemming from this process range from 1.1 to 2.0, depending on the number of lanes in the lane group and its corresponding traffic volume. Lane utilization factors based on driver desire to preposition can vary widely, depending on the volume of traffic that is prepositioning in the subject lane group.

Neither the signal capacity of various interchange types nor their relative capacity per lane was specifically determined within this research. Some examples of this form of analysis are illustrated in a related NCHRP publication (2). However, examination of parclo as compared to diamonds reveals obviously different traffic volume input patterns that may result in one design being more efficient than another for a given case. The software INTERCHANGE described in Appendix F can readily examine the patterns provided by each interchange type. Parclo versus diamonds also have more right-turn capacity per input lane due to their normal signal overlaps, but this feature may tend to overflow downstream closely-spaced links more than diamonds. Moreover, single-point diamonds are known (2) to have more arterial right-turn capacity per input lane than on its exit ramps because of unbalanced right-turn signal overlaps using three-phase signal operations.

4.2 RECOMMENDATIONS

The ideal saturation flow rate recommended for signalized (service) interchanges is 2,000 pcphgpl. In the context of the factors studied for this research, this ideal flow rate applies to through traffic movements that have an infinite distance to the back of downstream queues, operate under non-spillback conditions, discharge along tangent (i.e., straight) and level alignments, and have traffic volumes that are relatively high, reflecting those found during peak demand periods.

It is recommended that the equations provided in Chapter 3 be used to estimate the saturation flow rate, start-up lost time, green extension, end lost time, and lane utilization factor. In recognition of the relationship between saturation flow rate and start-up lost time, it is recommended that the equations provided in Chapter 3 be used to estimate *all* necessary phase capacity characteristics. In other words, selective use of only some of the equations in Chapter 3 for a capacity analysis is not recommended.

The recommended green extension value is 2.5 seconds for most undersaturated conditions. Other values are possible if the approach speed is outside the range of 64 to 76 km/h or when the volume-to-capacity ratio for the analysis period is above 0.88. An equation is provided in Chapter 3 for these situations.

The definition of effective green should be changed slightly from that used in the 1994 HCM. The new definition should be "effective green is that time during the subject phase when saturation flow at the stopline can occur under prevailing conditions." All "lost" times should be removed

from the phase, including: start-up, opposing queue blockage, output blockage due to spillback, and phase clearance lost times. This definition is very robust and covers all operating conditions, movements, and phases, including protected-plus-permitted. The PDX Model should be considered for estimating the output Clear Period and effective green time of the subject phase when oversaturation is likely. High-volume links which are nominally oversaturated or less than 200 meters long should be analyzed for queue spillback blockages using the features provided in the PDX Model.

Implementation of the PDX Model features into internationally recognized computer signal timing optimization programs, such as the PASSER programs, TRANSYT and SIDRA, is highly recommended. None currently handle oversaturated conditions very well, and the addition of the PDX Model features would give the programs the capability to reliably estimate queue spillback effects on saturation flow and effective green time. Some work toward this objective is known to be already underway (17,18,19).

Ramp weaving speeds and crossing capacity can be estimated using the methodology presented in Chapter 3 and Appendix E. Adequate travel distance to the back of the downstream receiving queue must be available for this capacity to be attained.

The *Highway Capacity Manual* should contain a chapter on Interchanges which emphasizes the unique forms and features of interchanges together with the special challenges associated with urban interchange traffic operations in general. Two-level signalized interchanges operating within a crossing arterial system should be presented together with freeway system integration issues associated with freeway traffic management. Special design and operational issues dealing with continuous one-way frontage roads should be presented. Unsignalized rotary interchanges could be identified as an alternative design concept. Moreover, the selective application of signalized interchanges to upgrade the capacity of a major urban arterial corridor should be noted, as illustrated in Appendix F of this report.

A major development of computer software should be funded in support of the *Highway Capacity Manual* effort. However, this software should not be limited to just being a processor to the HCM Interchange chapter that probably would focus just on "operational analysis of existing conditions." This new software should promote "options analysis" as needed to expediently conduct operational impact analyses for preliminary planning and design activities. The new program INTERCHANGE described in Appendix F illustrates some of the analysis concepts and software features recommended. However, significant funding is still needed to complete a professional-level software package for interchanges.

REFERENCES

1. Turner, J.M. and Messer, C.J., "Ramp Distance Requirements for Frontage-Road Ramps to Cross Street: Urban Freeway Design." *Transportation Research Record 682*, Transportation Research Board, Washington, D.C. (1978) pp. 58-64.
2. Messer, C.J., Bonneson, J.A., Anderson, S.D., and McFarland, W.F. *NCHRP Report 345: Single Point Urban Interchange Design and Operations Analysis*. Transportation Research Board, National Research Council, Washington, D.C. (1991).
3. *Special Report 209: Highway Capacity Manual*. Transportation Research Board, National Research Council, Washington, D.C. (1994).
4. *SAS/STAT User's Guide*. Version 6, 4th ed. SAS Institute, Inc., Cary, North Carolina (1990).
5. Yedlin, M., Leiberman, E., Phlegar, A., Kanaan, A., and Santiago, A.J., "The New TRAF-NETSIM Version 5.0." *Transportation Research Record 1440*. Transportation Research Board, Washington, D.C. (1994).
6. Bonneson, J.A. "Study of Headway and Lost Time at Single-Point Urban Interchanges." *Transportation Research Record 1365*. Transportation Research Board, Washington, D.C. (1992) pp. 30-39.
7. Stokes, R.W., Messer, C.J., and Stover, V.G. "Saturation Flows of Exclusive Double Left-Turn Lanes." *Transportation Research Record 1091*. Transportation Research Board, Washington, D.C. (1986) pp. 86-95.
8. Kimber, R.M., McDonald, M., and Hounsell, N.B. *TRRL Research Report RR 67: The Prediction of Saturation Flows for Road Junctions Controlled by Traffic Signals*. Transport and Road Research Laboratory, Department of Transport, Berkshire, England (1986).
9. Fambro, D.B., Messer, C.J., and Andersen, D.A. "Estimation of Unprotected Left-Turn Capacity at Signalized Intersections." *Transportation Research Record 644*. Transportation Research Board, Washington, D.C. (1977) pp. 113-119.
10. Messer, C.J. and Fambro, D.B. "Critical Lane Analysis for Intersection Design." *Transportation Research Record 644*. Transportation Research Board, Washington, D.C. (1977) pp. 26-35.
11. Fambro, D.B., Rouphail, N.M., Sloup, P.R., Daniel, J.R., Anwar, M., and R.J. Englebrecht. "Highway Capacity Revisions for Chapters 9 and 11." Report No. FHWA-RD-96-088, Federal Highway Administration, Washington, D.C., (1996).
12. Munjal, P.K., "An Analysis of Diamond Interchanges." *Transportation Research Record 349*. Transportation Research Board, Washington, D.C. (1971) pp. 47-64.

13. Messer, C.J., Fambro, D.B., and Richards, S.H. "Optimization of Pretimed Signalized Diamond Interchanges." *Transportation Research Record 644*. Transportation Research Board, Washington, D.C. (1977) pp. 78-84.
14. Messer, C.J., and Berry, D.J. "Effects of Design Alternatives on Quality of Service at Signalized Diamond Interchanges." *Transportation Research Record 538*. Transportation Research Board, Washington, D.C. (1975).
15. Van Arendonk, J. "Development of an Interchange Analysis Software." M.S. Thesis. Texas A&M University, College Station, May 1997.
16. Lloyd, M. D. "Assessment of Traf-NETSIM for Analyzing Arterial Weaving Between Ramp Terminals and Cross Streets." M.S. Thesis. Texas A&M University, College Station, May 1995.
17. Rouphail, N.M. and Akcelik, R. "Paired Intersections: Initial Development of Platooned Arrival and Queue Interaction Models." Australian Road Research Board. Working Paper WD TE91/010, Vermont South, Australia (1991).
18. Akcelik, R., Besley, M., and Shepherd, R. "SIDRA (Windows) Input Redesign for Paired Intersection Modelling." Discussion Note to WD 96/008. Australian Road Research Board, Vermont South, Australia (1996).
19. Chaudhary, N.A. and Messer, C.J. "PASSER IV-96, Version 2.1, User/Reference Manual." Texas Transportation Institute, College Station, Texas. (1996).

APPENDIX A

STATE-OF-THE-ART

A.1 SURVEY OF CURRENT PRACTICE

A comprehensive evaluation of the state-of-the-art in areas related to interchange design and traffic operations was conducted as part of this research. The focus of this evaluation was on issues underlying the design and operation of interchanges in urban or suburban areas. More specifically, the issues related to the signal-controlled ramp terminals and traffic flow along the cross street through the interchange. Consideration was also given to the relationship between the interchange terminals and any adjacent, closely-spaced signalized intersections.

One aspect of the evaluation involved a survey of transportation engineers. The intent of this survey was to gain insight into the current practices and concerns of engineers who evaluate interchange traffic operations. The survey was conducted in two stages. The first stage consisted of a one-page questionnaire. This questionnaire was intended to obtain basic types of information such as:

1. Common interchange types (geometric configurations)
2. Common operational problems
3. Common interchange operations analysis techniques
4. Common measures of effectiveness used for evaluation
5. Willingness of the respondent to participate in the second-stage survey.

The second-stage survey was designed to obtain more detailed information about interchange operations. This survey asked the respondent to select one interchange that they were familiar with and then respond to detailed questions about its operation and any steps taken to alleviate flow problems at this interchange. The respondent was also asked to describe the analysis techniques (or computer models) that they had successfully used to evaluate interchange operations. The findings from these two surveys are described in the next section.

A.1.1 First-Stage Survey

Distribution. The first-stage survey was sent during the first week of February, 1994. More than 2,400 surveys were sent out to engineers in the U.S. and abroad. The members of the following groups were specifically targeted:

- AASHTO Subcommittee on Traffic Engineering;
- AASHTO Subcommittee on Design;
- AASHTO Special Committee on Transportation Systems Operation;
- ITE Urban Traffic Engineers Council; and
- ITE Consultants Council.

Individuals in these groups include engineers responsible for planning, design, and operations of transportation facilities in the United States. In addition, several hundred surveys were sent to selected members of the Institute of Transportation Engineers (ITE).

After a thorough review of each returned questionnaire, a finalized total of 350 first-stage questionnaires were deemed completely responsive and valid. This group represents a 15-percent response rate, which is within the 10 to 20 percent rate expected prior to the survey. Overall, there were 146 responses from the public sector, including 68 from state DOTs, 63 from cities in 16 states, and 15 from counties in 8 states. Seventeen responses were received from outside of the United States (i.e., Canada - 11, Germany - 2, South Africa - 4). Responses were also received from 187 consultants in 23 states. The geographical distribution of the responses is summarized in Table A-1.

Results. The first-stage questionnaire consisted of six questions, primarily requesting but not limited to multiple-choice replies. The results for Questions 1, 3, 4, 5, and 6 are provided in Table A-2. The response format for Question 2 is somewhat different from the other questions and will be discussed separately.

Question 1 inquired about the frequency of interchange operations analysis. An analysis of the survey responses shown in Table A-2 indicates that cities/counties and consultants evaluate interchanges about "3 to 6 times per year." In contrast, most state DOTs evaluate interchanges "less than 3 times per year." However, a relatively large percentage of state DOTs indicated that they evaluate interchanges as frequently as "once per week."

Question 3 asked the respondents to rank the operational problems listed in terms of their frequency of occurrence at interchanges that the respondent is personally aware of through their work experiences. As indicated in Table A-2, "inadequate capacity" was given the highest ranking signifying it as the most frequently occurring problem. This problem was followed by "queue spillback" and then "weaving" in terms of their frequency of occurrence. Operational problems other than the three listed were described by twenty-two respondents. A common theme in these problems was a lack of effective signal coordination between the ramp terminals (or between the ramp terminal and adjacent signalized intersection).

Question 4 inquired about the types of analysis methods used to evaluate interchange traffic operations. In general, software methods were more frequently used than manual methods. The most commonly used software method is the Highway Capacity Software (HCS). PASSER II and TRANSYT-7F were also found to be frequently used.

Question 5 inquired about the most useful measure of effectiveness (MOE) for evaluating interchange traffic operations, particularly at ramp terminals. The most commonly selected MOE was delay, followed by spillback frequency, and volume-to-capacity ratio. The "other" category was infrequently used. Those that did use it indicated that speed or travel time measured along the cross street through the interchange would be most helpful.

Table A-1. Geographical distribution of responses to first-stage questionnaire.

State	Number of Responses					State	Number of Responses				
	State DOT	Cities	Counties	Consultants	Total		State DOT	Cities	Counties	Consultants	Total
Alabama	2				2	Montana					0
Alaska	1				1	Nebraska	1	1		2	4
Arizona	1	1	1	6	9	Nevada					0
Arkansas	1				1	New Hampshire					0
California	5	25	4	35	69	New Jersey	1			10	11
Colorado	2	3		6	11	New Mexico	1				1
Connecticut	1				1	New York	1	2	1	7	11
Delaware					0	North Carolina		5		3	8
Florida	9	3	3	26	41	North Dakota	1				1
Georgia	5		3	6	14	Ohio	1	5		4	10
Hawaii	1				1	Oklahoma		1		2	3
Idaho	1				1	Oregon	2		1	9	12
Illinois	2	2		11	15	Pennsylvania				14	14
Indiana	1				1	Rhode Island					0
Iowa	1				1	South Carolina	1				1
Kansas	2			1	3	South Dakota	1				1
Kentucky	1				1	Tennessee	1				1
Louisiana	1				1	Texas	2	9		10	21
Maine	1				1	Utah	1				1
Maryland	3		1	3	7	Vermont	1				1
Massachusetts	1				1	Virginia	3	2	1	5	11
Michigan	1	1		4	6	Washington	1	1		10	12
Minnesota					0	West Virginia	1				1
Mississippi	1				1	Wisconsin	3	1		4	8
Missouri		1		4	5	Wyoming	1				1
Number of States Represented:							40	16	8	22	45
U.S. Responses:							68	63	15	182	328
Responses from Outside U.S.:							4	4	0	9	17
Unknown:											5
Total:							72	67	15	191	350

Table A-2. Summary of responses to Questions 1, 3, 4, 5, and 6.

Response	Number (Percent) of Responses		
	City & County	State DOT	Consultant
Question 1. How often do you analyze some aspect of traffic operations at signalized intersections?			
Once per Week	11 (14%)	18 (26%)	32 (17%)
Once per Month	13 (17%)	14 (21%)	50 (27%)
3 to 6 Times per Year	28 (36%)	11 (16%)	70 (37%)
Less than 3 Times per Year	24 (31%)	24 (35%)	35 (19%)
No Response	2 (3%)	1 (1%)	0 (0%)
Question 3. Please rank the operational problem that you have encountered on the cross street at the most frequently used "Existing" interchange type. (1- no problem; 5 - serious problem)			
	Average Rank		
Inadequate Capacity	3.7	3.7	3.8
Weaving	2.2	2.3	2.6
Queue Spillback	3.5	3.5	3.7
Question 4. What analysis techniques do you use to evaluate traffic operations at the interchange ramp terminals? (mark all that apply)			
	Number (Percent) of Responses ¹		
Highway Capacity Manual	25 (32%)	20 (29%)	77 (41%)
Other Manual Methods	9 (12%)	9 (13%)	15 (8%)
Highway Capacity Software	38 (49%)	48 (71%)	161 (86%)
TRANSYT-7F	37 (47%)	30 (44%)	95 (51%)
PASSER II	35 (45%)	31 (46%)	89 (48%)
PASSER III	22 (28%)	20 (29%)	57 (30%)
TRAF-NETSIM	8 (10%)	15 (22%)	35 (19%)
Other Software Methods	10 (13%)	9 (13%)	35 (19%)
Question 5. What measure of effectiveness would be the most useful in evaluating traffic operations at the ramp terminals? (Including the effects of and nearby intersections on the minor street)			
Volume-To-Capacity Ratio	30 (38%)	33 (49%)	80 (43%)
Queue Spillback Frequency	34 (44%)	39 (57%)	99 (53%)
Delay per Vehicle	52 (67%)	44 (65%)	115 (61%)
Thru Movement Bandwidth	30 (38%)	21 (31%)	34 (18%)
Stops per Vehicle	17 (22%)	10 (15%)	29 (16%)
Other	5 (6%)	1 (1%)	10 (5%)
Question 6. Would you be willing to respond to a more detailed questionnaire concerning the details of interchange ramp junction operation?			
Yes	36 (46%)	38 (56%)	105 (56%)
No	34 (44%)	28 (41%)	70 (37%)
No Response	8 (10%)	2 (3%)	12 (6%)

Notes:

1 - Percentages for Questions 4 and 5 do not sum to 100% due to the "mark all that apply" nature of the questions.

Question 6 inquired about the willingness of the respondent to participate in the Second-Stage Survey. All total, 179 U.S. respondents indicated that they would be willing to participate. This represents about 51 percent of the 350 U.S. responses received.

Question 2 was used to identify the interchange configurations that were most commonly being evaluated by engineers. The frequency of evaluation was further categorized by "existing interchanges," "interchanges in design," and "interchanges in planning." This categorization was helpful in identifying current trends in interchange design. Variations of the question were prepared for engineers employed in the public or private sectors. Engineers employed in the public sector were asked to assess the percentage of each type of signalized interchange in their jurisdiction that are "existing," "in design," or "in planning." Engineers employed in the private sector were asked to assess the percentage of interchanges that they typically evaluate as part of their consulting activities.

The response to Question 2 is illustrated in Table A-3. As the data in this table suggest, the most commonly evaluated interchange configuration is the Compressed Diamond. However, all of the interchange forms were selected with sufficient frequency as to suggest that none should be excluded from consideration in the development of methods to evaluate interchange operations.

Table A-3. Distribution of interchange type by agency and stage of project development.

Interchange Type ¹	Existing Interchanges			Design or Construction			Planning		
	City or County	State DOT	Consultant	City or County	State DOT	Consultant	City or County	State DOT	Consultant
CD	30%	59%	37%	50%	45%	38%	37%	52%	40%
TD	25%	13%	25%	8%	13%	17%	9%	4%	16%
TDw/F	13%	8%	10%	17%	5%	9%	27%	10%	6%
PC	25%	10%	25%	17%	21%	30%	9%	24%	32%
Other	7%	10%	3%	8%	16%	6%	18%	10%	6%
Responses	47	52	124	12	32	76	11	21	104

Notes:

1 - Interchange Type Descriptions: CD - Compressed diamond (ramps 120 to 240 m); TD - Tight diamond (ramps less than 120 m); TDw/F - Tight diamond with frontage roads; PC - Partial cloverleaf of several variations.

Discussion of Results. The first-stage survey results show that practicing engineers are concerned with the effective operation of interchanges. Questions 1, 2, and 3 were asked to determine how much and what is being done on interchange design and operations. Based on the replies given for Question 3, engineers frequently encounter operational problems at interchanges in urban areas. However, it appears that neither the reasons for the problems (e.g., lack of capacity, queue spillback, weaving, etc.) are well understood nor are the solutions (e.g., interchange-specific analysis techniques) readily available. These limitations hinder an engineer's ability to analyze interchange traffic operations.

As a means of examining the operational problems at interchanges in more detail, the second-stage questionnaire was developed and distributed to the interested first-stage respondents. The results from the second-stage questionnaire can be found in the next section.

Question 2 verified that the diamond interchange (either compressed or tight urban) was the most common existing, designed, and planned interchange. This fact is likely due to the reduced right-of-way costs associated with interchanges of the diamond family, relative to those of the partial cloverleaf family.

Question 4 was asked to determine which traffic models are being used by practicing engineers to evaluate interchange operations. The most common type of analysis used by the respondents is computer software models and, most often, the Highway Capacity Software (HCS). This may be due to its widespread acceptance, consistency with the *Highway Capacity Manual*, or the relative ease with which it can be used.

As the current HCS is relegated to worksheet-based procedures that are sufficiently simple that they can be used in a manual fashion, it tends to be limited in its ability to evaluate traffic flow problems in interchange areas. As a result, several computer-based simulation models were often cited by the respondents. Specifically, TRANSYT-7F was cited by nearly half of all the respondents. This may be due to the fact that TRANSYT-7F is sensitive to the proximity of adjacent ramp terminals or signalized intersections in its signal timing optimization routine. Another software model, PASSER-II was also cited by 40 to 50 percent of the respondents as being a useful tool to analyze arterial traffic flow through interchange ramp terminals. This large response may be due to the fact that PASSER-II optimizes signal timing based on progression analysis. In order to better understand the strengths and weaknesses of these models, as perceived by the users, the second-stage questionnaire requested that the respondents expand upon their reasons for selecting a specific analysis tool. The results of the second-stage survey can be found in a later section.

Question 5 asked the respondents to identify an MOE that they felt would be useful in evaluating traffic operations at an interchange. In order to better explain traffic operations at an interchange, MOEs must be selected that are comprehensible and practical. Thus, it is important that the MOEs selected to evaluate an interchange be those that are easy to observe and to comprehend (not something abstract in nature). Delay per vehicle was the MOE most often selected by respondents. This finding is probably due in part to the fact that the HCM uses delay to describe the level of service provided to motorists at intersections. It would appear to be a logical extension on the part of the respondents as a diamond interchange has the appearance of two arterial intersections rather than two closely-spaced ramp terminals whose individual operation is highly dependent on the signal operation of the other ramp terminal.

After delay, queue spillback frequency was the next most frequently cited MOE by the respondents. This is consistent with the findings regarding operation problems, as requested in Question 3. Queue spillback is recognized by many engineers as a significant problem at urban interchanges. It is likely that the length of the queues formed between the ramp terminals and the frequency that they spillback into the upstream ramp terminal (or closely-spaced adjacent

signalized intersection) could be used as a primary indicator of the quality of flow within the interchange area.

Question 6 showed a willingness to respond to the second-stage questionnaire. The large positive response received in this regard is believed to represent the engineering community's overall level of interest in the topic of this research.

A.1.2 Second-Stage Survey

The findings from the first-stage survey provided important information regarding the extent of operational problems at urban interchanges and the general thoughts of the practicing engineering community regarding techniques for evaluation of these problems. These findings were used to develop the format and content of the second-stage survey. This survey sought specific details of operational problems occurring at specific types of interchanges. The second-stage survey also inquired about the strengths and weaknesses of specific analysis techniques.

Distribution. The second-stage survey was sent during the last week of March 1994. This survey was sent to 179 individuals who indicated a willingness to respond to it from the first survey. A total of 31 completed surveys were returned representing a 17 percent response rate, a rate that was somewhat lower than anticipated.

The findings from the second-stage survey were generally consistent with those from the first-stage survey. Therefore, it was concluded that the information obtained from the second-stage survey would be more representative than the small sample size would otherwise suggest. Possible reasons for the small sample size could include a combination of the following: (1) the survey may have been conducted during a busy time of the year for the respondents, (2) respondents may have believed that the time required to complete the survey was excessive, and (3) the return date may not have allowed the respondent enough time to adequately respond.

Of the 31 surveys returned, 29 were determined to be valid responses in the context that they addressed the interchange types and issues described in the survey. Valid responses were returned from 10 state DOTs, 8 cities in 6 states, 9 consultants in 8 states, and 2 cities in Canada. Overall, 21 states are represented among the 29 valid returned surveys. The response rate was about 29 percent for the DOTs, 25 percent for the cities, 0 percent for the counties, 9 percent for the consultants, and 18 percent for international replies.

Results. The findings from the second-stage survey are described in the following paragraphs. These findings are presented in the following format: the individual question is repeated (in *italics*); then, the response to each question is summarized; finally, some observations and insights are provided to put the findings in the proper context.

In general, each respondent was asked to identify one interchange of the diamond or partial cloverleaf family and answer the survey questions as they relate to this interchange. The interchange that they selected was to have attributes that were consistent with the objectives of this research and that were otherwise not unusual or geometrically constrained. Specifically, the

selected interchange was to be located in an urban or suburban setting, have signalized ramp terminals, and a distance between ramp terminals of 275 meters or less. The respondents were encouraged to complete additional survey forms for a second or third interchange, if time permitted.

1. Please sketch the interchange.

The types of interchanges sketched (and described in subsequent questions) ranged from the partial cloverleaf to the single-point urban interchange. In three instances, the respondent submitted a second survey describing a different interchange. As a result, descriptions of 32 interchanges were received; however, one interchange was described twice by two different respondents. As a result, only 31 unique interchanges are described in the summary statistics. These interchanges are distributed among the seven interchange types listed below.

- | | |
|--|----|
| 1. Tight Urban Diamond (less than 120 m between ramps): | 10 |
| 2. Tight Urban Diamond with frontage roads: | 2 |
| 3. Compressed Diamond (120 to 240 m between ramps): | 11 |
| 4. Conventional Diamond (more than 240 m between ramps): | 2 |
| 5. Single Point Urban Diamond: | 2 |
| 6. Partial Cloverleaf (Type A): | 3 |
| 7. Partial Cloverleaf (Type AB): | 1 |

2. What is the distance between the two ramp terminals (as measured along the cross street from stop line to stop line)?

Average:	150 meters,	Standard Deviation:	90 meters
Minimum:	60 meters,	Maximum:	410 meters

These distances are not representative of all interchanges because the survey specifically requested information on interchanges whose ramp-to-ramp separation distance was less than 275 meters. However, they are representative of urban interchanges that tend to experience traffic operational problems because of short ramp separation distances

3. What is the distance between the ramp terminal and the nearest downstream signalized intersection (as measured along the cross street from stop line to stop line)?

Average:	180 meters,	Standard Deviation:	90 meters
Minimum:	50 meters,	Maximum:	440 meters

As with Question 2, these distances should not be taken as typical of all interchange locations; just those interchanges in urban areas with relatively close ramp spacings. The respondent was informed (in the survey) that one objective of the project was to address the operational impact of closely-spaced intersections. As a result, the respondents, tended to include

interchanges with closely-spaced intersections. These closely-spaced intersections often lead to problems such as queue spillback between ramp terminals and left-turn bay overflow.

4. Could you provide a block diagram illustrating the phase sequence for one signal cycle?

Twenty-four respondents provided phase sequence information. The open nature of this question led to a wide range of response formats. As a result, it was difficult to generalize the types of phasing based on the descriptions provided by the respondents. The problems with interpretation were grouped into three categories. First, very few of the respondents used the block diagram format requested; many provided signal timing information from plan sets or from manufacturer-specific controller printouts that could not be translated with any real certainty. Second, it was apparent that many respondents were only guessing at the phase sequence based on their observation rather than obtaining the actual sequence from the appropriate authority. Finally, many respondents described the phase sequence for each ramp terminal but did not convey the manner in which they were coordinated.

After reviewing the phase sequences provided, the following generalizations were made. First, only 2 of the 25 diamond interchanges appear to be using the four-phase-with-overlap phasing. It was expected that this type of phasing would be more prevalent due to its ability to deal with high-volume left-turns and narrow ramp separation distances. Second, it appeared that most of the interchanges with two controllers used three-phase operation at each ramp terminal with, presumably, some type of signal offset timing used to coordinate the two major street through movements at the ramp terminals.

5. What type of signal control is used to implement the phasing described in Question 4?

About 59 percent of the respondents indicated that two controllers were used at the interchange (one controller for each ramp terminal). Another 31 percent of the respondents indicated that one controller was used for both terminals; the remaining 10 percent did not know the controller type. As diamond interchanges were the most common interchange type cited in Question 1, it is somewhat surprising that so many sites had two controllers at the interchange. One common controller for both diamond interchange ramp terminals is generally best able to maintain the type of two-way traffic progression necessary to eliminate queues on the street segment internal to the ramp terminals. The trend of using two controllers (with presumably signal offset timing) may possibly contribute to the queue spillback that many of the interchanges exhibit because of the lower level of coordination it affords to the left-turn movements.

6. What control mode does the controller provide?

About 75 percent of the respondents indicated that semiactuated control was used at their interchange. Thirteen percent indicated that pretimed control was used and 9 percent indicated that fully-actuated control was used. Comparison of the responses among Questions 5 and 6 indicate that there is no correlation between the number of controllers and the type of control mode.

7. *Is the interchange controller(s) coordinated with the cross street signal system?*

As semi-actuated control implies coordination, it is logical that coordination was found at the same percentage of interchanges as those having semi-actuated control. In fact, this was the case, 75 percent of the interchanges described had semi-actuated control. The high percentage of coordinated interchanges suggests that, while efforts should be made elsewhere in improving interchange traffic operations, impacts on coordination should not be forgotten.

8. *Describe the traffic flow problem which tends to be most disruptive to smooth traffic flow.*

Although this question asked about the most disruptive problem, most respondents chose to describe more than one problem. In general, they selected one or more of the traffic flow problems that were described in the survey. These problems are restated below along with the percentage of responses that identified a particular problem as being the most disruptive.

- 41% a. Capacity restriction due to queue spillback between ramp terminals.
- 34% b. Capacity restriction due to queue spillback from a ramp terminal into the upstream signalized intersection.
- 31% c. Capacity restriction due to cross street left-turn bay queue overflow into the through lanes.
- 25% d. Unbalanced lane volumes on the cross street approaches to the ramp terminals due to high-volume downstream turn movement.
- 25% e. Flow turbulence between a ramp terminal and an adjacent signalized intersection due to high-volume lane changing (i.e., right-turn at terminal followed by left-turn at intersection, or vice versa).
- 22% f. Capacity restriction due to queue spillback from a signalized intersection into the upstream ramp terminal.
- 22% g. Capacity restriction due to queue spillback from the off-ramp signal into the freeway main lanes.
- 19% h. Poor signal coordination between the two ramp terminals due to complex signal phasing, variability in hourly turning movement volumes, or minimal interior queue storage space.
- 16% i. Capacity restriction due to queue spillback from a ramp meter into the upstream ramp terminal.
- 6% j. Poor or nonexistent signal coordination between the ramp terminals and adjacent intersections due to jurisdictional policies (i.e., City control of the intersection and State control of the interchange).
- 0% k. Poor or nonexistent signal coordination between the ramp terminal and ramp meter.

Based on the percentages listed above, it appears that "queue spillback between ramp terminals" is the most frequently found problem at interchanges in narrow-rights-of-way. When combined with "left-turn bay overflow," it would appear that traffic flow problems at interchanges are most frequently found between the ramp terminals, where the volume of left-turns is highest.

One response from a consultant in Portland, Oregon reported a lack of capacity between the terminals of a compressed diamond interchange. This interchange has one controller for both terminals and operates in a three-phase sequence. The respondent indicated that the restricted capacity "results in a queue spillback into the adjacent cross street signalized intersection." This spillback, in turn, "results in little or no capacity for local circulation" at the adjacent intersection. Another response from a consultant in New York identified problems associated with turning movements. The respondent reported that the tight urban diamond exhibited left-turn bay queue overflow at one of the ramp terminals and severe turbulence associated with high-volume weaving on the cross street between the ramp and adjacent signalized intersection. The maneuver that caused most of this turbulence was the off-ramp right-turn movement becoming a left-turn movement at the downstream intersection.

Further examination of the responses to this question revealed that all of the reported flow problems related to queue spillback between the ramp terminals were associated with tight or compressed diamond interchanges. Single point diamond interchanges, conventional (wide) diamond interchanges, and partial cloverleaf interchanges were not associated with queue spillback-related flow problems. The single point diamonds do not experience queue spillback because they combine the two ramp terminals into one intersection. The conventional and partial cloverleaf interchanges do not experience spillback because of the relatively large distances separating the two ramp terminals.

9. What treatments have you applied (or would apply) to alleviate the traffic flow problem described in Question 8?

A wide range of treatments were described by the respondents. There were no definitive trends although it appeared that geometric changes were commonly seen as the only available treatment. Typical geometric treatments included adding a second left-turn lane or an additional through lane to the cross street. In some instances, the respondent recognized the difficulty of adding lanes to (i.e., widening) an existing bridge. One of the more interesting signal timing treatments was the use of signal phasing at the adjacent intersection to separate the traffic movements accessing the on-ramp so as to prevent the congestion associated with a high-volume of weaving vehicles. Many respondents indicated that improved or updated signal timing and coordination helped mitigate some traffic problems.

10. If you were asked to evaluate and quantify the problem described in Question 8, what analysis technique (or techniques) would you presently use?

The analysis techniques cited by most (60 percent) of the respondents can be described as those developed for isolated signalized intersections. These techniques were used for the analysis of the individual ramp terminals and adjacent intersection. Of those techniques identified, that described in Chapter 9 of the 1985 *Highway Capacity Manual* (HCM) was cited as being most frequently used. PASSER II was identified by 33 percent of the respondents as being helpful in coordinating the two ramp terminals and the adjacent signalized intersection. Other, less frequently noted techniques included the use of the NETSIM and TRANSYT-7F computer models.

11. *With regard to the analysis technique described in Question 10:*

a) *What is its main technical strength?*

The most frequently cited strength of the "Chapter 9" HCM technique was that it is easier to use than multiple-intersection software programs (e.g., PASSER II, TRANSYT-7F, NETSIM, etc.). In general, the HCM technique was used to evaluate the individual ramp terminals with appropriate calibration of the progression adjustment factors to account for nearby intersections. The PASSER II program was credited with being the easiest multiple-intersection program to use. This program was used when a through traffic progression solution and/or queue length estimate was desired. NETSIM was noted to be the only program that accurately modeled queue spillback and congested flow conditions. TRANSYT-7F was noted to consider upstream queue length and left-turn demand when determining the "optimum" traffic progression solution.

b) *What is its main technical weakness?*

The weaknesses cited for the HCM technique were that it did not accurately model the effect of closely-spaced upstream intersections and that it did not yield queue length estimates. The weaknesses cited for the PASSER II program were that it did not provide progression solutions for left-turn movements, did not consider upstream queue length when determining the progression solution, did not allow the user to enter some types of interchange phasing, and did not consider right-turn demand. NETSIM was noted to be very time consuming to use due to its microscopic simulation formulation. A couple of respondents noted that none of the techniques dealt with the coordination of a ramp meter with the ramp terminal.

c) *Describe how you have overcome any weakness described in Question 11-b.*

In general, the respondents indicated that they used engineering judgement and field observation to manually adjust the signal offset or timings to optimize traffic progression and minimize queue lengths. A few respondents indicated that they used a second analysis technique; however, they did not elaborate on which supplemental techniques were used and under what conditions.

Summary. The results of the second-stage survey indicated that queue spillback, left-turn bay overflow, and weaving between the off-ramp and downstream intersection were significant operational problems at interchanges with closely-spaced ramps or adjacent signalized intersections. Of these problems, queue spillback tends to degrade the smooth flow of many interchange traffic movements and thereby, aggravate mild inefficiencies into significant capacity constraints.

Thus, the indirect solution to many interchange-related problems appears to be related to devising analysis techniques that are sensitive to the proximity to downstream queues, the propensity of these queues to spillback, and the relationship between queue-clearance-time and the signalization of the interchange ramp terminals and adjacent intersection. The implementation of the findings of this research could be facilitated by their incorporation into one or more of the existing capacity analysis techniques (e.g., Highway Capacity Software, PASSER II, etc.).

A.1.3 Researchers Field Observations

The research team studied a dozen interchanges during the field studies and spent many hours observing traffic operations at the sites. Traffic congestion was routinely observed at all of these interchanges. Comparisons were rapidly made among interchange types, types of operational problems observed, and the probable cause of these problems. Our summary of these field observations are noted below:

1. Design life of older interchanges usually exceeded so traffic demand exceeded interchange capacity during rush hours.
2. Many older interchanges noted above have a predominant number of single-lane left turn lanes within the interchange and/or have single lanes at ramp terminals assigned to serve heavy left and/or right turning movements. Deficient turning capacity exists.
3. Due to urban growth, four-lane crossing arterials need to have six lanes. Cross street has functionally become a major urban arterial.
4. Traffic management of queueing and spillback is difficult at some interchanges due to high volumes and high percentages of turning traffic having typical lane distribution problems. Some approaches along the crossing arterial and within the interchange have almost constant demand within the cycle so queueing can not be mitigated using traditional arterial signal coordination techniques. Random flow should be assumed, as a minimum, for queueing analysis. Parclo A's seem to be more susceptible to constant demand conditions. However, all off ramp terminals having free right turning operations will be more prone to overloading downstream arterial storage areas.
5. Many arterial links connecting the freeway interchange with the "next" downstream signalized intersection experience high traffic demands to/from the freeway (interchange) and the flows are frequently nearly constant over the cycle. These adjacent intersections often have four-phase signals that provide less arterial capacity than the three-phase signals at the interchanges. For these conditions, many of these connecting links appear to be too short to provide good storage and operating conditions. Longer intersection spacings and better design policies are needed for interchange planning and design.
6. Traffic control strategies employed appear to be based on undersaturated flow conditions which may lose efficiency during oversaturated conditions now being more commonly experienced. Better management of queue spillback to mitigate the onset of congestion is needed together with the need to transition to downstream bottleneck control strategies once oversaturation has occurred.

A.2 EXISTING CAPACITY AND LEVEL OF SERVICE MODELS

A.2.1 Overview

This section presents a state-of-the-art summary of current traffic models for assessing the capacity, delay, and level of service of traffic operations at the signalized interchanges shown in Figure A-1. Most interchange forms have two signalized intersections per interchange. The primary focus of previous research has been on diamond interchanges because they are the predominant signalized interchange form (1). Partial cloverleaf (parclos) interchanges have similar phasing strategies and can be modeled using the same general capacity analysis methodology provided for diamond interchanges.

Current interchange capacity analysis essentially treats each intersection within the interchange as a separate entity, with minimal consideration given progression effects and spillback. The significant number of users of the Highway Capacity Software (HCS) for interchange studies, based on the field survey of practicing engineers previously described, shows this technology limitation since there is no generally accepted standard analysis methodology for interchanges. The Arterials (Chapter 11) methodology of the 1994 Highway Capacity Manual (HCM) (2) is sometimes used for interchanges, but it assumes that the intersections are widely spaced and traffic operations are undersaturated. Because the signalized ramp junctions at an urban interchange are usually less than 300 meters apart, and most urban diamonds are less than 200 meters, the effects of closely-spaced signals should be identified and modeled in interchange analysis.

Current capacity analysis for intersections also assumes that the output (saturation) flow from a signal is independent of downstream traffic conditions. This is a major deficiency for interchange analysis due to the high traffic volumes and closely-spaced signals. Even the most highly utilized macroscopic computer-based signal timing optimization models (PASSER II, PASSER III and TRANSYT 7F) presently fail to reliably address oversaturation issues at signalized intersections.

A.2.2 Signal Phase

A signal phase is a period of time provided by the signal controller unit to an approach permitting legal entry of vehicles into the intersection. The entry may be described as being protected (from conflicting vehicular and pedestrian traffic), permitted (to legally enter but exposed to other potential conflicting movements), or combinations of the two (protected-plus-permitted). Other descriptive terms are used such as exclusive/permissive phasing or combined phasing. A basic protected through or left turn phase would have the following signal interval times

$$\phi = G + Y + R_c \quad (A-1)$$

DIAMOND INTERCHANGES

PARCLO INTERCHANGES

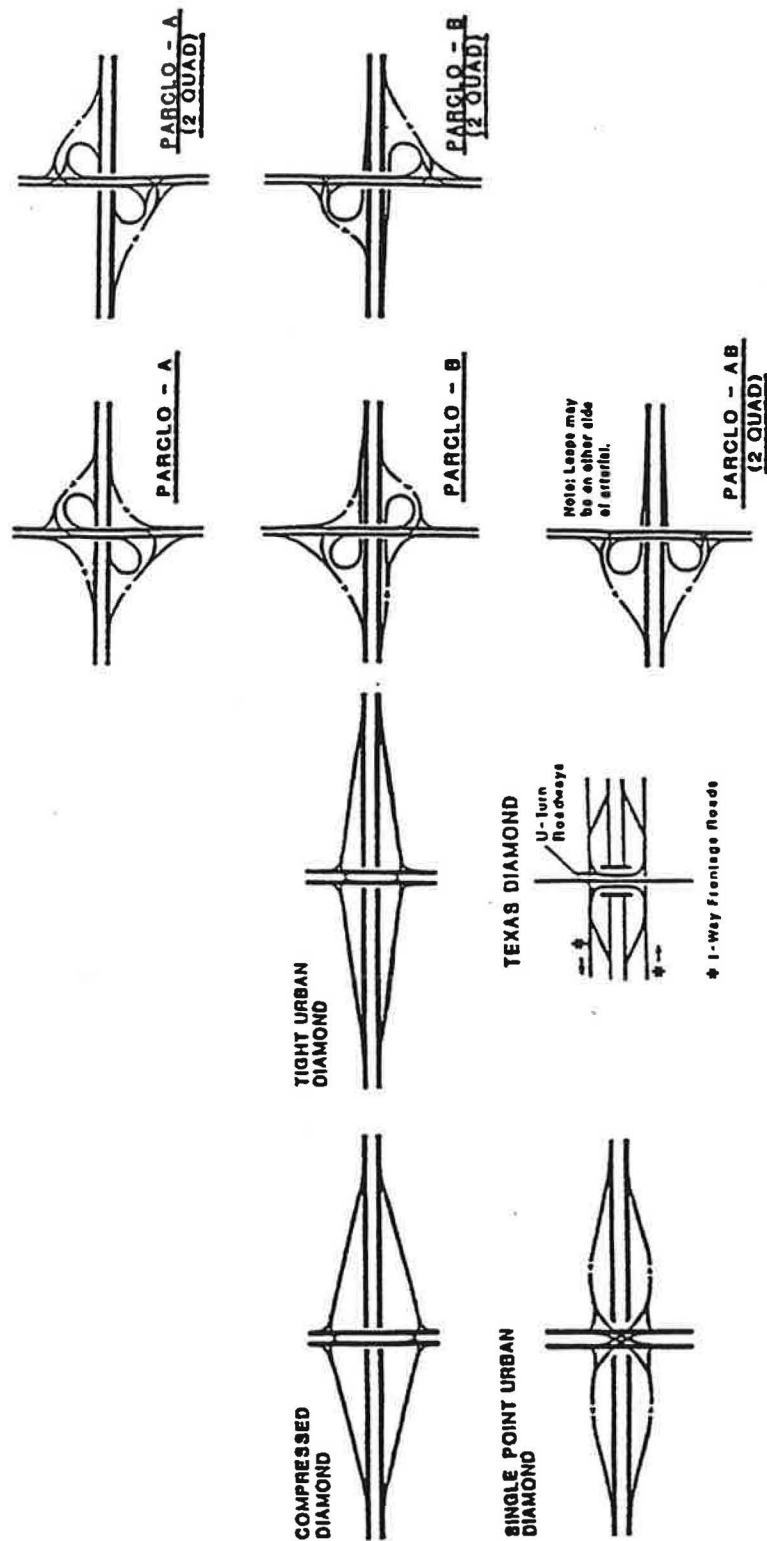


Figure A-1. Common Two-Level Signalized Interchanges.

where:

ϕ	=	total duration of the signal phase, sec;
G	=	green signal interval, sec;
Y	=	yellow warning interval, sec, and
R_c	=	red clearance interval, sec.

Assuming that a long line of cars is trying to use the phase, the flow profile measured at the stopline shown in Figure A-2 would be expected, assuming that the phase is protected from conflicting traffic and unimpeded by downstream queue spillback. Following onset of green, flow reaches a maximum or saturation flow rate and would be expected to maintain this flow until the green interval ends. Some usage of the yellow change interval occurs, perhaps as much as 2.5 seconds on the average. The 1994 HCM assumes through phases are never blocked or impeded by downstream storage conditions.

A.2.3 Phase Capacity

The capacity of a traffic signal phase is the maximum number of vehicles that can be expected to enter the intersection per cycle (assuming one phase per cycle) from the lane group being analyzed under prevailing roadway, traffic and control conditions. The Highway Capacity Manual (HCM) expresses this concept of phase capacity at maximum flow as being the total area under the saturation flow curve per phase

$$n = \int_0^{\phi} s(t) dt \quad (A-2)$$

Assuming the saturation flow $s(t)$ is a known constant; s_g (vps), during the effective green portion, g , of the phase and there is one phase per cycle, then the phase capacity per cycle then becomes

$$n = g \cdot s_g \quad (A-3)$$

It is assumed that the phase is protected and unblocked during the effective green period, g , while serving a waiting queue. A representative capacity flow profile of the signal was shown in Figure A-2 for unimpeded/unblocked saturation flow.

Phase capacity is usually expressed in equivalent flow rate units of vehicles per hour consistent with traditional volume counting practice. Assuming there is only one phase of interest per cycle and noting that the numbers of cycles per hour are

$$k(C) = \frac{3600}{C} \quad (A-4)$$

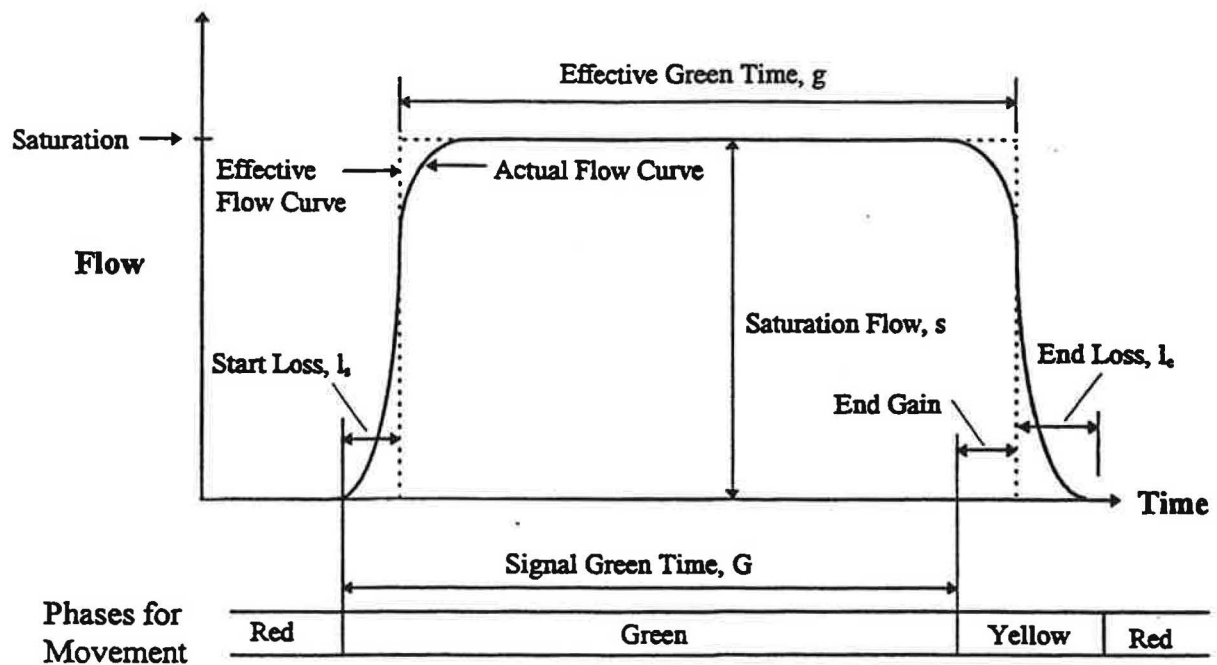


Figure A-2. Basic Saturation Flow Model for Unimpeded Traffic Conditions.

then the hourly phase capacity is the product of the capacity per phase times the number of phases (cycles) per hour, or

$$c = n \cdot k(C) \quad (\text{A-5})$$

substituting for n from Equation A-3 yields

$$c = g s_g \frac{3600}{C} \quad (\text{A-6})$$

Rearranging terms yields

$$c = \frac{g}{C} s_g \cdot 3600 \quad (\text{A-7})$$

$$c = \frac{g}{C} s_g \cdot 3600 \quad (\text{A-7})$$

Letting the saturation flow be expressed in vehicles per hour, as $s = 3600 s_g$, results in the more traditional hourly-based phase capacity formula used in the 1994 HCM (2), of

$$c = \frac{g}{C} s \quad (\text{A-8})$$

where:

c	=	phase capacity for the subject lane group, vph;
g	=	effective green time of phase, sec;
C	=	cycle length, sec;
n	=	phase capacity, veh;
$k(C)$	=	number of phases (cycles) per hour;
s_g	=	average saturation flow during phase, vpsg; and
s	=	average saturation flow during phase, vphg.

A.2.4 Saturation Flow

The 1994 HCM provides a module for calculating saturation flow for signalized intersections which is summarized below (2). The HCM defines saturation flow rate as the flow in vehicles per hour that could be accommodated by the lane group assuming that the green phase was always available to the lane group, that is, the green ratio (g/C) was 1.0. Computations begin by selecting an "ideal" saturation flow rate, usually 1,900 passenger cars per hour of display green time per lane (pcphgpl), and then adjusting this value to prevailing conditions which may not be ideal. Following Eq. 9-12 of the 1994 HCM for signalized intersections, the adjusted saturation flow is

$$s = s_o N f_w f_{HV} f_g f_p f_{bb} f_a f_{RT} f_{LT} \quad (\text{A-9})$$

where:

s	=	saturation flow rate for the subject lane group, expressed as a total for all lanes in the lane group under prevailing conditions, vphg;
s_o	=	ideal saturation flow rate per lane, usually 1,900 pcphgpl;
N	=	number of lanes in the lane group;
f_w	=	adjustment factor for lane width (12-ft lanes are standard), given in Table A-4;

f_{HV}	=	adjustment factor for heavy vehicles in the traffic stream, given in Table A-5;
f_g	=	adjustment factor for approach grade, given in Table A-6;
f_p	=	adjustment factor for the existence of a parking lane adjacent to the lane group and the parking activity in that lane, given in Table A-7;
f_{bb}	=	adjustment factor for the blocking effect of local buses that stop within the intersection area, given in Table A-8;
f_a	=	adjustment factor for area type, given in Table A-9;
f_{RT}	=	adjustment factor for right turns in the lane group, given in Table A-10; and
f_{LT}	=	adjustment factor for left turns in the lane group, assumed to be 0.95.

A.2.5 Saturation Flow Adjustment Factors

The use of adjustment factors is a common feature throughout the HCM. Each factor accounts for the impact of one or several prevailing conditions that are different from the ideal conditions for which the ideal saturation flow rate applies. The factor represents the average adjustment needed over the entire duration of the displayed effective green time.

Lane Width. The lane width adjustment factor, f_w , accounts for the deleterious impact of narrow lanes on saturation flow rate and allows for an increased flow on wide lanes. Twelve-foot lanes are the standard. The lane width factor may be calculated with caution for lane widths greater than 5 m (16 ft), or an analysis using two narrow lanes may be conducted. Note that the use of two lanes will always result in a higher saturation flow rate than a single wide lane, but in either case the analysis should reflect the way in which the width is actually used or expected to be used. In no case should the lane width factor be calculated for lane widths less than 2.5 m (8 ft).

Table A-4. Adjustment Factor For Average Lane Width (f_w)

Average Lane Width, W(ft)	Lane Width Factor, f_w
8	0.867
9	0.900
10	0.933
11	0.967
12	1.000
13	1.033
14	1.067
15	1.100
16	1.133

NOTE: $f_w = 1 + W - \frac{12}{30}$ $W \geq 8$ (if $W > 16$, consider two-lane analysis)

Table A-5. Adjustment Factor For Heavy Vehicles (f_{HV})

Percent Heavy Vehicles, % HV	Heavy Vehicle Factor, f_{HV}
0	1.000
2	0.980
4	0.962
6	0.943
8	0.926
10	0.909
15	0.870
20	0.833
25	0.800
30	0.769
35	0.741
40	0.714
45	0.690
50	0.667
75	0.571
100	0.500

NOTE: $f_{HV} = \frac{100}{100 + \%HV (E_T - 1)}$ $0 \leq \%HV \leq 100$
 where $E_T = 2.0$ passenger cars per heavy vehicle.

Table A-6. Adjustment Factor For Grade (f_g)

Grade %G		
Type	Percent	Grade Factor (f_g)
Downhill	-6 or less	1.030
	-4	1.020
	-2	1.010
Level	0	1.000
Uphill	+2	0.990
	+4	0.980
	+6	0.970
	+8	0.960
	+10 or more	0.950

NOTE: $f_g = 1 - \% \frac{G}{200}$ $-6 \leq \%G \leq +10$

Table A-7. Adjustment Factor For Parking (f_p)

No. of Lanes In Lane Group, N	No. of Parking Maneuvers Per Hour N_m					
	No Parking	0	10	20	30	40
1	1.000	0.900	0.850	0.800	0.750	0.700
2	1.000	0.950	0.925	0.900	0.875	0.850
3 ^a	1.000	0.967	0.950	0.933	0.917	0.900

NOTE: $f_p = \frac{N - 0.1 - 18N_m/3600}{N}$ $0 \leq N_m \leq 180$ $f_p \geq 0.05$ ^aUse formula for more than 3 lanes, or more than 40 maneuvers per hour.

Table A-8. Adjustment Factor For Bus Blockage (f_{bb})

No. or Lanes In Lane Group, N	No. of Buses Stopping Per Hour, N_B				
	0	10	20	30	40
1	1.000	0.960	0.920	0.880	0.840
2	1.000	0.980	0.960	0.940	0.920
3 ^a	1.000	0.987	0.973	0.960	0.947

NOTE: $f_{bb} = \frac{N - 14.4N_B/3600}{N}$ $0 \leq N_B \leq 250$ $f_{bb} \geq 0.05$ ^aUse formula for more than 3 lanes or more than 40 buses stopping per hour

Table A-9. Adjustment Factor For Area Type (f_a)

Type of Area	Area Type Factor, (f_a)
CBD	0.90
All other areas	1.00

Heavy Vehicle and Grade. The effects of heavy vehicles and grades are treated by separate factors, f_g and f_{HV} , respectively. Their separate treatment recognizes that passenger cars are affected by approach grades, as are heavy vehicles. The heavy vehicle factor accounts for the additional space occupied by these vehicles and for the negative differential in the operating capabilities of heavy vehicles with respect to passenger cars. The passenger car equivalent (E_T) used for each heavy vehicle is 2.0 passenger car units (pcus) and is reflected in the formula. This value was increased from 1.5 used in the 1985 HCM. The grade factor accounts for the effects of grades on the operation of all vehicles.

Parking. The parking adjustment factor, f_p , accounts for the frictional effect of a parking lane on flow in an adjacent lane group, as well as for the occasional blocking of an adjacent lane by vehicles moving into and out of parking spaces. Each maneuver (either in or out) is assumed to block traffic in the lane next to the parking maneuver for an average of 18 sec. The number of parking maneuvers used is the number of maneuvers per hour in parking areas directly adjacent to the lane group and within 76 m (250 ft) upstream from the stop line. If more than 180 maneuvers per hour exist, a practical limit of 180 should be used. If the parking is adjacent to an exclusive-turn-lane group, the factor only applies to that lane group. On a one-way street, parking on the left side will affect the left most lane group, as in a one-way street with no exclusive-turn lanes, the number of maneuvers used is the total for both sides of the lane group. Note that parking conditions with zero maneuvers are not the same as no parking.

Bus Blockage. The bus blockage adjustment factor, f_{bb} , accounts for the impacts of local transit buses that stop to discharge or pick up passengers at a near-side or far-side bus stop within 76 m (250 ft) of the stop line (upstream or downstream). This factor should only be used when stopping buses block traffic flow in the subject lane group. If more than 250 buses per hour exist, a practical limit of 250 should be used. When local transit buses are believed to be a major factor in intersection performance, more precise methods may be needed. The factor used here assumes as average bus blockage time of 14.4 sec. during a green indication.

Area Type. The area type adjustment factor, f_a , accounts for the relative inefficiency of business area intersections in comparison with those in other locations, primarily because of the complexity and general congestion in the business environment.

Right-Turn. Turning factors depend upon a number of parameters. The most important characteristic is the manner in which turns are accommodated in the intersection. Turns may operate out of exclusive or shared lanes, with protected or permitted signal phasing, or with some combination of these conditions. The impact of turns on saturation flow rates is very much dependent upon the mode of turning operations.

Table A-10. Adjustment Factor For Right Turns (f_{RT})

Cases 1-6: Exclusive/Shared Lanes and Protected/Permitted Phasing				
$f_{RT} = 1.0 - P_{RT} [0.15 + (PEDS/2100) (1 - P_{RTA})]$ $0.0 \leq P_{RT} \leq 1.0$ $0.0 \leq P_{RTA} \leq 1.0$ $0 \leq PEDS \leq 1700$ $f_{RT} \geq 0.05$				
Proportion of RT in lane group = 1.00 for excl. RT lane (Cases 1-3); <1.00 for shared lane (Cases 4-6) Proportion of RT using protected phase = 1.00 Volume (peds/hr) of peds conflicting with RT (if PEDS > 1700, use 1700)				
Case 7: Single-Lane Approach (all traffic on approach in a single lane).				
$f_{RT} = 0.90 - P_{RT} [0.135 + (PEDS/2100)]$ $0 < P_{RT} \leq 1.0$ $0 < PEDS \leq 1700$ $F_{RT} = 1.00$ if $P_{RT} = 0.0$ $F_{RT} \geq 0.05$				
Proportion of RT in lane group. Volume (peds/hr) of peds conflicting with RT (use 0 if RT is completely protected).				
Case	Range of Variable Values			Simplified Formula
	P_{RT}	P_{RTA}	PEDS	
1. Excl. RT lane; prot. RT phase	1.0	1.0	0	0.85
2. Excl. RT lane; perm. RT phase	1.0	0.0	0-1700	$0.85 - (PEDS/2100)$
3. Excl. RT lane; prot. + perm. RT phase	1.0	0-1.0	0-1700	$0.85 - (PEDS/2100) (1 - P_{RTA})$
4. Shared RT lane; prot. RT phase	0-1.0	1.0	0	$1.0 - P_{RT}[0.15]$
5. Shared RT lane; perm. RT phase	0-1.0	0.0	0-1700	$1.0 - P_{RT}[0.15 + (PEDS/2100)]$
6. Shared RT lane; Prot.+ perm. RT phase	0-1.0	0-1.0	0-1700	$1.0 - P_{RT}[0.15 + (PEDS/2100)]$
7. Single-lane approach	0-1.0	-----	0-1700	$0.9 - P_{RT}[0.135 + (PEDS/2100)]$

The right-turn adjustment factor, f_{RT} , depends upon a number of variables, including

1. Whether the right turn is made from an exclusive or shared lane;
2. Type of signal phasing (protected, permitted, or protected plus permitted) a protected right-turn phase has no conflicting pedestrian movements and a permitted phase has conflicting pedestrian movements;
3. Volume of pedestrians using the conflicting crosswalk;
4. Proportion of right-turning vehicles in the shared lane, and
5. Proportion of right turns using the protected part of protected-plus-permitted phase.

Item 5 should be determined by field observation but can be grossly estimated from the signal timing. This is done by assuming that the proportion of right-turning vehicles using the protected phase is approximately equal to the proportion of the turning phase that is protected. If $P_{RTA} = 1.0$ that is, the right turn is completely protected from conflicting pedestrians, a pedestrian volume of zero should be used. The right-turn factor is 1.0 if no right turns are present on the lane group. When RTOR is permitted, the right-turn volume may be reduced as described in the discussion of the Volume Adjustment Module of the HCM.

Left-Turn. The left-turn adjustment factor, f_{LT} , is based on similar variables to those for the right-turn adjustment factor, including

1. Whether left turns are made from exclusive or shared lanes;
2. Type of phasing (protected, permitted, or protected plus permitted);
3. Proportion of left-turning vehicles using a shared lane group; and
4. Opposing flow rate when permitted left turns are made.

The left-turn adjustment factor is 1.0 if the lane group does not include any left turns. When a left turn is not opposed at any time by through vehicles but encounters conflicting pedestrian movements, the left turn should be treated using the adjustment procedure for right turns. If no conflicting pedestrian movements are present, a normal protected left-turn adjustment is performed.

Basically, turn factors account for the fact that these movements are not made at the same speeds and saturation flow rates as through movements. They consume more of the available green time and consequently more of the lane group's available capacity. The turn adjustment factors reflect seven different conditions under which turns may be made, as follows:

- Case 1: Exclusive lane with protected phasing, $f_{lt} = 0.95$;
- Case 2: Exclusive lane with permitted phasing, see below;
- Case 3: Exclusive lane with protected-plus-permitted phasing;
- Case 4: Shared lane with protected phasing;
- Case 5: Shared lane with permitted phasing;
- Case 6: Shared lane with protected-plus-permitted phasing; and
- Case 7: Single-lane approaches (right-turn factors only).

A.2.6 Assessment of HCM Capacity Methodology

A critical assessment of the 1994 HCM capacity analysis methodology for signalized intersections should consider all aspects and issues related to its application to interchanges. Some factors appear to be directly usable as is for interchanges. Other factors would generally appear to be unneeded. The need to consider turning radius at interchanges seems apparent due to the high volumes of turning traffic and range of turning radius that might be encountered. Also, cost of some interchange designs are sensitive to turning radius. Because interchanges usually have closely-spaced intersections, and often are oversaturated with queue spillback blocking upstream flow, queue impediments should be considered. The definition and resulting application of effective green, g , is judged a major problem for general use. In addition, technology issues arise as to how the HCM uses capacity analysis in the estimation of performance measures and level of service using vehicular traffic delay.

Adjustment Factors. Geometric, location and environmental aspects are more identifiable as to transferability from intersections to interchanges. These aspects are considered first for all of the adjustment factors noted in Equation A-9 for saturation flow.

<u>Factor</u>	<u>Recommendation</u>
f_w =	same lane width factors should be used, although larger vehicles on the average may be found at interchanges;
f_{HV} =	same heavy vehicle factor and E_T (pcu) should be used, although larger and heavier vehicles may use some interchanges, and if so, larger E_T should be applied based on field observations;
f_g =	adjustment factor for grades should be similar for interchanges;
f_p =	parking will not likely be permitted in/around interchanges, so this factor is not needed;
f_{bb} =	buses may be stopping on the cross street, so this factor should be retained;
f_a =	area-type factor is not needed for interchange environments;
f_{RT} =	right-turn factor should be retained, but give additional consideration to turning radius, number of turning lanes, and the fact that very few (if any) phases are permissive from shared lanes. Moreover, this factor should only adjust for time periods when the traffic actually moves, unlike in the 1994 HCM, so that delays can be better estimated;
f_{LT} =	left-turn factor should be retained only for protected phasing as a base case. Further adjustments based on turning radius should be made ; and
f_D, f_v =	new adjustment factors for downstream queue spillback impediments and traffic pressure, as described further in this report.

Permissive Left Turn Operations. Permissive/permitted left turn operations occur at signalized interchanges when left turns, after yielding to opposing queues, subsequently find acceptable gaps in the opposing traffic flow and then complete their intended left turn maneuvers. The period of time the opposing queue blocks the left turn during green, g_{bq} , must be known and can be estimated from

$$g_{bq} = \frac{rv_{ro}}{s_{go} - v_{go}} \quad (A-10)$$

Assuming undersaturated conditions with no queue spillback, the unblocked green time, g_u , is the time the permitted left turns can safely maneuver across the opposing flow during the latter portion of the phase for a time of

$$g_u = g_p - g_{bq} \quad (A-11)$$

where:

g_u	=	unblocked green time of the phase serving the lane group, sec;
g_p	=	original effective green time of the phase, sec;
g_{bq}	=	time opposing queue blocks the permitted phase from serving the permitted left turns, sec;
r	=	red time of opposing lane group per cycle, sec;
v_{ro}	=	opposing arrival rate on red for lane group, vph;
v_{go}	=	opposing arrival rate on green for lane group; vph; and
s_{go}	=	nominal saturation flow for opposing lane group; vphg.

The left turn saturation flow possible during the permitted green interval, g_u , should be calculated from the following relationship:

$$s_L = v_{go} \frac{e^{-v_{go} T_L}}{1 - e^{-v_{go} H_L}} \quad (A-12)$$

where:

v_{go}	=	opposing lane group volume during permitted green, vph;
T_L	=	left turn critical gap, sec/3600; and
H_L	=	left turn minimum headway, sec/3600.

Recommended values for T_L and H_L depend on several factors, including whether the left turns are made from a dedicated lane or a shared lane. Table A-11 provides recommended values of the factors T_L , H_L based on the 1994 HCM applications and other references (2, 3). The resulting relationships of saturation flow for permitted left turns versus opposing volume are estimated by Equation A-12 are shown in Figure A-3 for two types of left turn lane use.

Table A-11. Parameters for Permitted Left Turn Phases

Factor	Left-Turn Lane	Shared-Turn Lane
T_L (sec)	4.5 (sec)	4.6 (sec)
T_L	0.00125	0.00128
H_L (sec)	2.4 (sec)	4.5 (sec)
H_L	0.00067	0.00125

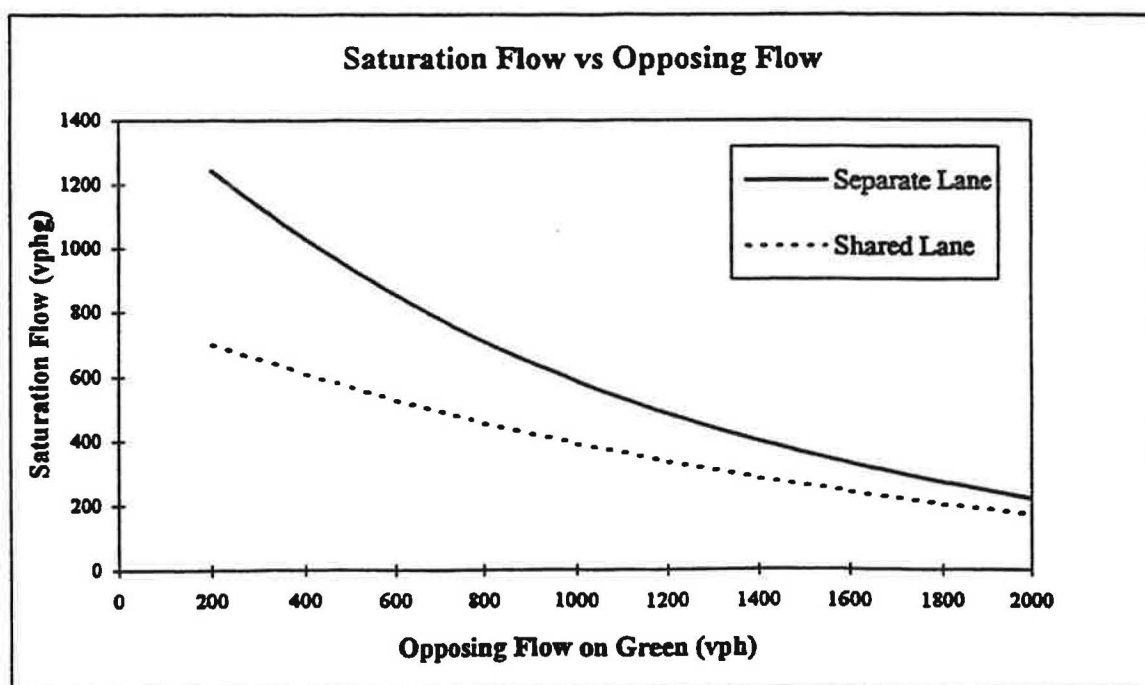


Figure A-3. Permitted Left Turn Saturation Flow as Related to Opposing Volumes.

The modeling of permissive left turns raises a major technological issue with the 1994 HCM. The question is "How is it best to estimate capacity and delay using the basic HCM methodology?" The 1994 HCM uses the following general methodology to estimate phase capacity, c , as Equations A-8 and A-9 have noted

$$c = \frac{g}{C} s_a = \frac{g}{C} \cdot s_o \cdot N \cdot f_{\pi} \quad (\text{A-13})$$

where g is the phase's effective green, including times during the phase when flow may be blocked, e.g. by opposing queues for permitted left turns, s_a is the adjusted saturation flow, and f_{π} is the product of all relevant adjustment factors to existing conditions during the phase.

The recommended method for estimating phase capacity per cycle essentially sums component maximum allowable flows over the cycle according to

$$c = \frac{\sum g_f s_f}{C} \quad (\text{A-14})$$

for differing flow conditions (f) over the cycle, C . For permitted left turn phases, this can be implemented in either of two ways, both of which yield the same estimate of capacity:

HCM Method:

$$c = g s_a = g \cdot N \left[\frac{g_f}{g} \cdot \frac{s_f}{s_p} \right] s_p \quad (\text{A-15})$$

$$c = g N f_{pm} \cdot s_p \quad (\text{A-16})$$

where f_{pm} is the resulting HCM capacity adjustment component for permitted left turns due to opposing queue blockages and turn flow reductions due to opposing traffic flow from

$$f_{pm} = \frac{g_f}{g} \cdot \frac{s_f}{s_p} \quad (\text{A-17})$$

and s_p is the adjusted saturation flow for protected left turns for existing conditions.

and s_p is the adjusted saturation flow for protected left turns for existing conditions.

Recommended Method:

$$c = g_b \cdot 0 + g_f s_f \quad (\text{A-18})$$

$$c = g_f s_f \quad (\text{A-19})$$

$$c = g_f \cdot f_o \cdot s_p \quad (\text{A-20})$$

$$f_o = \frac{s_L}{s_p} \quad (\text{A-21})$$

where:

- s_L = equivalent permitted left turn saturation flow rate through opposed flow from Equation A-12 ;
- s_p = saturation flow rate for protected left turn operations for otherwise existing conditions; and
- s_o = base ideal saturation flow for signalized intersections, pcphgpl.

Both the HCM and proposed flow modeling methods yield the same phase capacity, but they produce significant differences in delay when used in traditional HCM delay models as noted below.

The major difference between the two methods is in delay estimation. These differences arise due to how the first term of the HCM delay equation calculates queuing delay for left turning vehicles arriving on red and opposed green. Consider the following analysis for permitted left turn operations from a separate left turn bay. It has been shown that the phase capacity of the two methods would be equal to

$$c = g s_a = g_f s_f \quad (\text{A-22})$$

The HCM effective green definition exceeds the actual green period when left turn flow occurs

but the saturation flow during turning is greater than the HCM adjusted saturation flow

$$s_f > s_a \quad (A-24)$$

Since the capacities of the two methods are the same, for convenience let

$$s_a = \frac{g_f}{g} s_f = \alpha s_f \quad \alpha = \frac{g_f}{g} \leq 1 \quad (A-25)$$

HCM Delay Estimation Method:

$$d_H = \frac{C}{2} \frac{(1 - g/C)^2}{(1 - v/s_a)}$$

Recommended Delay Estimation Method:

$$d_R = \frac{C}{2} \frac{(1 - g_f/C)^2}{(1 - v/s_f)}$$

Define the delay ratio, e , to be the ratio of the HCM delay to the recommended method for calculating the first term of the delay equation such that

$$e = \frac{d_H}{d_F} = \frac{(1 - g/C)^2}{(1 - \alpha g/C)^2} \cdot \frac{(1 - \alpha v/s)}{(1 - v/s)} \quad (A-26)$$

Let $\lambda = g/c$; $y = v/s$, and $x = y/\lambda$ to simplify the comparisons

$$e = \frac{(1 - \lambda)^2}{(1 - \alpha \lambda)^2} \cdot \frac{(1 - \alpha y)}{(1 - y)} = \frac{(1 - \lambda)^2}{(1 - \alpha \lambda)^2} \cdot \frac{(1 - \lambda \alpha x)}{(1 - \lambda x)} \quad (A-27)$$

Consider the following two volume cases, low and high volumes, for three green splits, with $\alpha = 0.5$ for assumed moderate opposing volumes.

Case I. Low Left Turn Volumes: $y \rightarrow 0, v \rightarrow 0, x \rightarrow 0$

$$e_I = \frac{(1 - \lambda)^2}{(1 - \alpha \lambda)^2} = \beta \quad (\text{A-28})$$

Case II. High Left Turn Volumes: $x\phi = 1$ (degree of saturation equals 1.0)

$$e_{II} = \beta \cdot \frac{(1 - \lambda \alpha)}{(1 - \lambda)} \quad (\text{A-29})$$

These delay comparisons are presented in Table A-12. At extremely low volumes and high green splits, the HCM method would underestimate the average delay by over 55 percent (a delay ratio of 0.444). At high volumes, the error in delay would be less, but still practically significant. The HCM Method (the combination of effective green definition and resulting saturation flow adjustment method) will consistently underestimate the signal delay incurred for permitted left turns even when calculated turn capacities are the same as the proposed method.

Table A-12. Comparison of First-Term Errors in HCM Delay Estimation

Green Ratio $\lambda = g/C$	Low Volumes $X = 0$	High Volumes $X = 1.0$
1/4 0.250	0.735	0.858
1/3 0.333	0.641	0.801
1/2 0.500	0.444	0.667

Blockage. Output flow from the stop line may be blocked and otherwise impeded by several conditions. Permitted left turns are blocked by opposing queues from using a portion of the displayed green interval. All turning movements may also be blocked by queue spillback from other movements storing behind the downstream signal, even though the downstream signal is undersaturated. This research has identified that queue blockage is a major consideration during oversaturated conditions, and spillback blockages must be identified for meaningful capacity and delay analyses to be conducted. Thus, the true "effective green" should be used.

oversaturated conditions, and spillback blockages must be identified for meaningful capacity and delay analyses to be conducted. Thus, the true "effective green" should be used.

A.2.7 HCM Vehicular Delay Methodology

Vehicular delay is recognized by the 1994 HCM as being a significant traffic performance measure and, consequently, is used as the sole criterion for the level of service provided, for isolated intersections. Federal Highway Administration has sponsored a research project in coordination with the Highway Capacity Committee of TRB to provide a recommended update for HCM chapters on isolated signalized intersections (Ch.9) and on coordinated signalized arterials (Ch.11). This research has just recently been completed and is being reviewed by HCQS committee of TRB. The following is a summary of the current arterial recommendations provided by the cited researchers (4).

Generalized Delay Model. The proposed generalized delay model for signalized intersections and arterial streets (interrupted traffic flow facilities) for a subject lane group (phase) is (4).

$$d = d_1 + d_2 + d_3 \quad (\text{A-30})$$

where:

- d = average total delay per vehicle for vehicles arriving during the analysis period, sec/veh;
- d_1 = uniform delay, sec/veh;
- d_2 = incremental delay due to random arrivals and overflow queues, sec/veh; and
- d_3 = incremental delay due to non-zero queues at the start of the analysis period, sec/veh.

where the uniform, or so-called first-term, delay is

$$d_1 = 0.5 C \frac{(1 - g/C)^2}{1 - (g/C) \min(X, 1.0)} PF f_{pp} \quad (\text{A-31})$$

and the second-term of delay is

$$d_2 = 900 T \left[(X-1) + \sqrt{(X-1)^2 + \frac{8 k I X}{T c}} \right] \quad (\text{A-32})$$

If $X < 1$, a residual queue of size n_i exist at the start of the analysis period, and a zero queue exists at the end of the analysis period, then

$$d_3 = \left(3600 \frac{n_i}{c} \right) \frac{0.5 n_i}{T c (1 - X)} \quad (\text{A-33})$$

If $X < 1$ and non-zero queue exists at the end of the analysis period, then

$$d_3 = \left(3600 \frac{n_i}{c} \right) - 1800 T (1 - X) \quad (\text{A-34})$$

If $X \geq 1$, or oversaturation is present, then

$$d_3 = \left(3600 \frac{n_i}{c} \right) \quad (\text{A-35})$$

If $X < 1$ and zero queue exists at the start of the analysis period, then

$$d_3 = 0 \quad (\text{A-36})$$

where:

C	=	average cycle length, sec;
g	=	average effective green time, sec;
X	=	degree of saturation for subject lane group;
PF	=	progression adjustment factor;
f_{pp}	=	early/late arrival adjustment factor;
T	=	analysis period in hours, in which the model parameters are fixed;
k	=	delay parameter for given arrival and service distributions;
I	=	variance-to-mean ratio of arrivals/cycle at a point;
c	=	capacity of the lane group, veh/hr; and
n_i	=	queue at the start of the analysis period.

Isolated Intersections. Considering isolated signalized intersections for a 15-minute analysis period under pretimed control and no initial queue, which essentially describes the nominal analysis conditions of Chapter 9 of the 1994 HCM, the following default values would be used: $PF = 1.0$, $f_p = 1.0$, $T = 0.25$ hours, $k = 0.50$, $I = 1.0$ and $d_3 = 0$. The resulting total delay equation for a lane group for an isolated approach (or one having random arrivals) would be

$$d = d_1 + d_2 \quad (A-37)$$

$$d = 0.5C \frac{(1 - g/C)^2}{1 - (g/C) \min(x, 1.0)} + 225 \left[(X - 1) + \sqrt{(X - 1)^2 + \frac{16X}{c}} \right] \quad (A-38)$$

which is almost the same model for pretimed control as has been used in the HCM since 1985. Only the X^2 term in d_2 has been dropped from the overall delay equation.

Signalized Arterials. The generalized delay model for interrupted flow facilities can be used for coordinated approaches by selecting appropriate values for coordinated conditions and type of traffic control. Table A-13 provides the progression adjustment factor (PF) for the first term of the delay equation based on arrival type (AT) together with the early/late arrival factor (f_{pp}) which also depends on the degree of saturation of the lane group.

For coordinated intersections, the following equation is proposed for 15-minute analysis periods

$$d = 0.5C \frac{(1 - g/C)^2}{1 - (g/C) \min(X, 1.0)} PF f_{pp} + 225 \left[(X - 1) + \sqrt{(X - 1)^2 + \frac{(32 kIX)}{c}} \right] \quad (A-39)$$

Table A-13. Uniform Delay Adjustment Factors

Progression Adjustment Factor (PF)						
Green Ratio (g/C)	Arrival Type (AT)					
	AT-1	AT-2	AT-3	AT-4	AT-5	AT-6
0.20	1.167	1.083	1.000	0.917	0.833	0.750
0.30	1.268	1.143	1.000	0.857	0.714	0.571
0.40	1.445	1.222	1.000	0.778	0.555	0.333
0.50	1.667	1.333	1.000	0.667	0.333	0.000
0.60	2.001	1.500	1.000	0.500	0.000	0.000
0.70	2.556	1.778	1.000	0.222	0.000	0.000
Default R_p	0.333	0.667	1.000	1.333	1.667	2.000
$PF = (1 - P)/(1 - g/C); R_p = R/(g/C)$						
Early/Late Arrival Factor (f_{pp})						
Degree of Saturation (X)	Arrival Type (AT)					
	AT-1	AT-2	AT-3	AT-4	AT-5	AT-6
0.2	1.000	0.880	1.000	1.240	1.000	1.000
0.4	1.000	0.910	1.000	1.180	1.000	1.000
0.6	1.000	0.940	1.000	1.120	1.000	1.000
0.8	1.000	0.970	1.000	1.060	1.000	1.000
1.0	1.000	1.000	1.000	1.000	1.000	1.000
$f_{pp(x=0)}$	1.000	0.850	1.000	1.300	1.000	1.000
$f_{pp} = f_{pp(x=0)} + (1 - f_{pp(x=0)})x$						

Source: Reference 4.

Delay Results. For isolated approaches having random arrival flow the model (Eq. A-38) predicts that vehicular delay increases as volumes and resulting v/c ratio (degree of saturation) increase, as depicted in Figure A-4 for an isolated intersection approach. At low volumes (v/c ratio), delays are caused primarily by vehicles arriving on red. At higher volumes nearing saturation ($X = 1.0$), some cycles fail to completely clear the number of vehicles arriving (randomly) per cycle and these vehicles, which would previously not be delayed hardly at all (because they arrived on late green), are now being delayed a full red duration plus adding to the delay of subsequent arrivals. At even higher arrival volumes that routinely exceed capacity, queueing and delay continue to increase for each cycle that oversaturation exists with theoretical delays being limited only by the ending of the oversaturation period, T . Extremely large delays are theoretically possible if adequate queue storage is available. A practical maximum X value of 1.2 is employed by the HCS software to limit the delay calculated, but no maximum delay value has been selected.

A.2.8 HCM Capacity Analysis Recommendations

The general HCM capacity, delay and level of service methodology used for isolated signalized intersections and arterial street systems should be generally applicable for interchanges. However, the effects of signal coordination and limited queue storage available due to the closely-spaced signals having high turning traffic should be more precisely identified. Calculations of queueing delay and spillback should be specifically related to signal offset and available queue storage. Effects of downstream blockage should be identified and assessed.

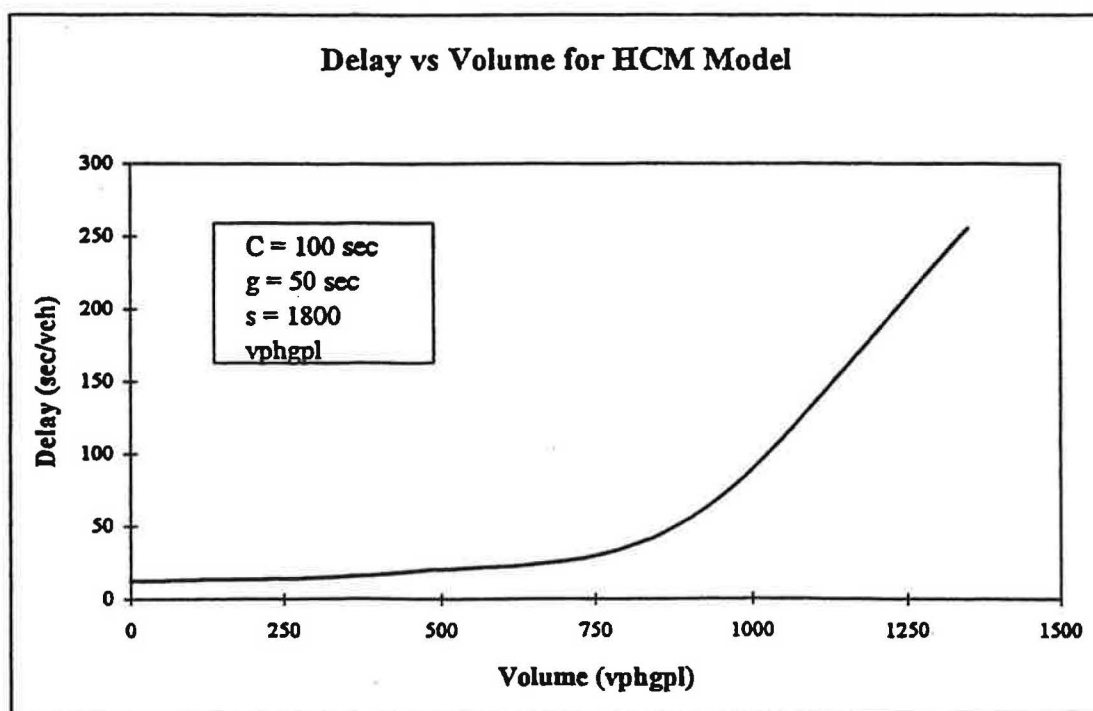


Figure A-4. Proposed HCM Model for Estimating Vehicular Delay for Random Arrivals.

The operational dichotomy between capacity and delay should be recognized. Capacity of a phase fundamentally depends on the forward motion of a platoon. For capacity to be provided during a cycle, right-of-way must be provided, flow through the intersection must be unblocked, and downstream storage/processing must be available. Delay, on the other hand, is the antithesis of motion. Delay occurs because vehicles have arrived that can not continue their forward motion at their desired speed. Normally, delay occurs because vehicles arrive on red, not green. Capacity depends on motion; delay depends on restriction to motion.

The preferred set of analysis variables should be the optimal choice when considering both capacity and delay. There are several ways to precisely calculate signal capacity. There are only two ways to precisely calculate delay: integration of vehicle travel through the queue, or integration of vehicular queue over time (the cycle). All other delay models are approximations.

The precise estimation of traffic performance at interchanges and other closely-spaced intersection depends on several important assumptions and estimates being true and accurate. Central to this issue is that of phase capacity per cycle. It is recommended that the phase capacity per cycle, n , be estimated from

$$n = \sum_C g_f s_f(z) \quad (\text{A-40})$$

where f signifies flow (motion) occurring. Both g_f and s_f must be correct for the period and prevailing conditions. The preferred definitions for g and s are as follows:

$$\begin{aligned} g &\equiv g_f &= &\text{effective green time during phase (cycle) when platoon flow can occur} \\ & & &\text{at rate } s_f \text{ sec; and} \\ s &\equiv s_f(z) &= &\text{maximum average platoon flow that can occur during } g_f \text{ considering} \\ & & &\text{the distance } z \text{ to the back of the downstream queue at start of green,} \\ & & &\text{vpsg/vphg.} \end{aligned}$$

Note that only when there is one motion period per cycle/phase does

$$n = g_f s_f = g_\phi s_\phi \quad (\text{A-41})$$

and only during a fully protected, unimpeded, unblocked phase does

$$n = g s \quad (\text{A-42})$$

as assumed in the 1994 HCM.

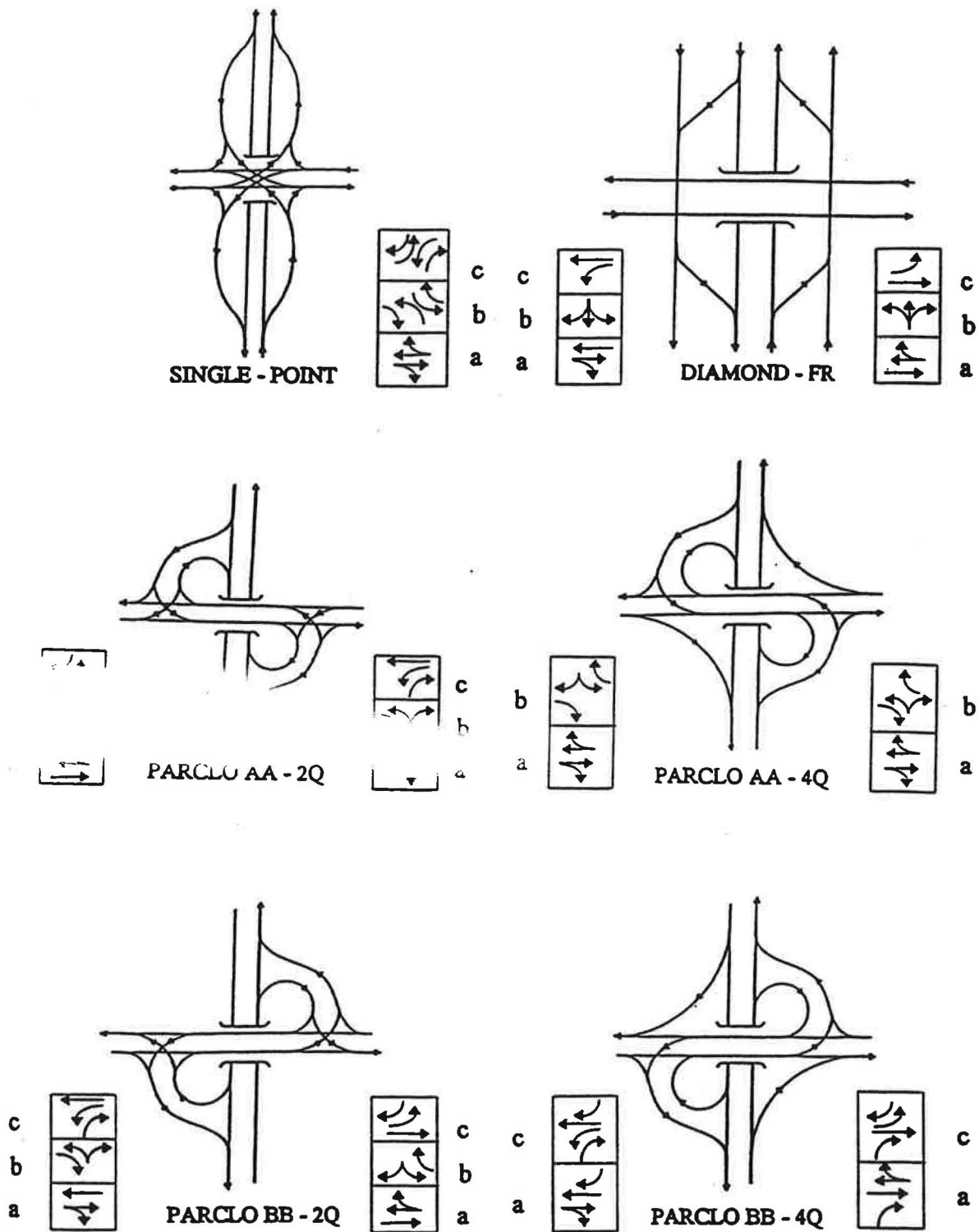


Figure A-5. Interchange Signal Phasing.

Table A-14. Basic Signal Phase Sequences at Interchanges

Phase Combination	Left-Side Intersection	Right-Side Intersection	Signal Sequence
1	<i>a:b:c</i>	<i>a:b:c</i>	Lead-Lead
2	<i>a:b:c</i>	<i>a:c:b</i>	Lead-Lag
3	<i>a:c:b</i>	<i>a:b:c</i>	Lag-Lead
4	<i>a:c:b</i>	<i>a:c:b</i>	Lag-Lag

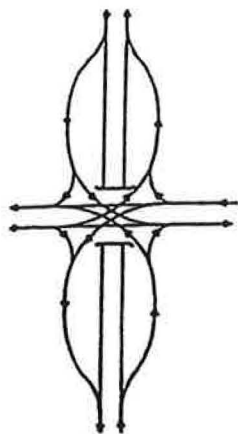
Figure A-6 presents typical phasing sequences for some common interchange types for illustrative purposes. The four-phase strategy is depicted for diamond interchanges. As can be seen, partial cloverleafs (parclos), in contrast to traditional diamond interchanges, provide some application variation but do not change the basic concepts. Two-quad parclos have three phases per intersection; whereas, four-quad parclos have only two phases per side, deleting Phase *b*. The two-quad parclos may have Phase *c* in the outbound direction (Parclo BB) or in the inbound direction (Parclo AA). Moreover, the ramp phases (Phase *b*) may be on the approach side (Parclo AA) as in a conventional diamond, or on the opposite side (Parclo BB). Parclos provide one distinguishing phasing difference to diamonds in that ramp right-turns (sometimes free) signal overlaps are common. Single-point urban (diamond) interchanges (SPUI) basically employ a conventional intersection phasing sequence, using dual-ring operations on the crossing arterial and a single phase for the ramps. Longer phase lost times arise due to the longer clearance intervals generally employed to clear the large intersection area. Arterial right-turn capacities at SPUI's are typically larger than for the off-ramps due to the significant increase in protected (overlap) turning time.

Traffic flow through an interchange depends on the signal timing, traffic mix and geometric features. Throughput depends on the capability to first enter the interchange and then exit the facility. The more restrictive case is usually entry into the interchange, although cross street left turn lanes can become overloaded under some traffic patterns. Seldom is the arterial outbound phase a restraint unless it is blocked or impeded by queue spillback from a closely-spaced downstream intersection. Operations at some two-quadrant parclos may be an exception.

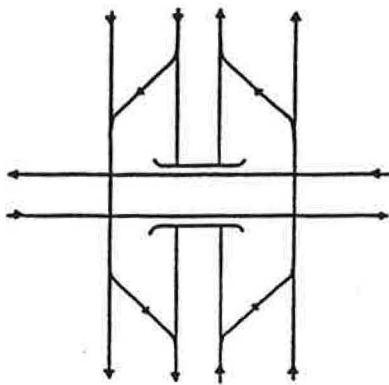
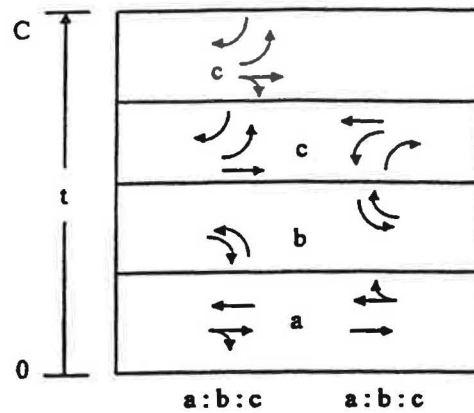
Independent Operations. At low traffic volumes and/or wide intersection spacings, the two intersections (*u, d*) could operate independently. Let ϕ be the total phase duration (green + yellow + red clearance intervals) such that the effective green is $g = \phi - l$ where l is the total phase lost time. At one intersection, *u*, the three conflicting phases (*a, b, c*) must add to one cycle

$$\phi_{ua} + \phi_{ub} + \phi_{uc} = C_u \quad (A-44)$$

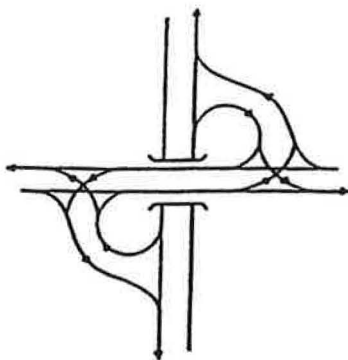
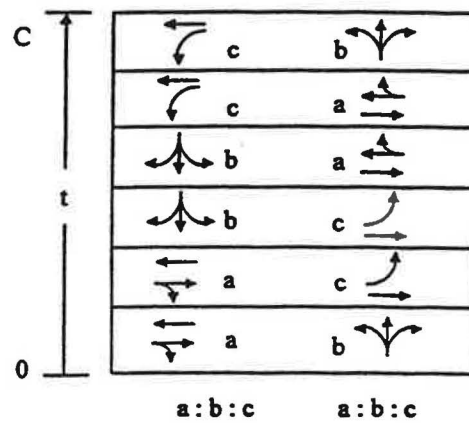
and at the second intersection, *d*, the sum of three conflicting phase times may add to another cycle



SINGLE - POINT



DIAMOND - FR



PARCLO BB - 2Q

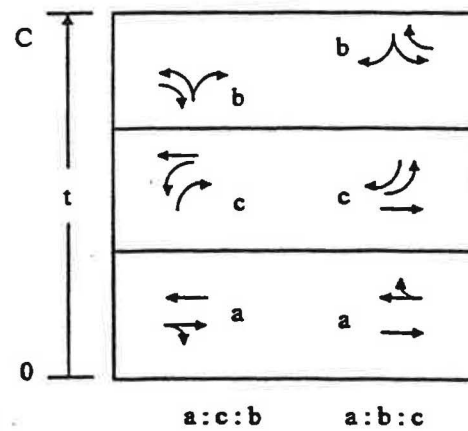


Figure A-6. Interchange Signal Phase Sequences.

length of

$$\phi_{da} + \phi_{db} + \phi_{dc} = C_d \quad (A-45)$$

Subject to each phase, m , satisfying its minimum green requirements $\phi_m \geq \min \phi_m$.

As long as the two signals operate independently, cycle times can be different to accommodate variable traffic demands, green splits can be provided without constraints to better satisfy those demands, and capacities are at a maximum. Capacity principles described in Chapter 9 - Signalized Intersections of the HCM would apply, as follows. For a representative intersection and cycle, C_i , the sum of critical phases is

$$\phi_{ia} + \phi_{ib} + \phi_{ic} = C_i \quad (A-46)$$

which is equivalent to

$$(g + l)_{ia} + (g + l)_{ib} + (g + l)_{ic} = C_i \quad (A-47)$$

Letting L_3 equal the sum of the three lost times per phase, the total effective green time per cycle is

$$g_{ia} + g_{ib} + g_{ic} = C_i - L_3 \quad (A-48)$$

And since the v/c ratio for a phase or related lane group "m" is given by

$$X_{im} = \left(\frac{vC}{sg} \right)_m \quad (A-49)$$

Solving for the effective green, g_m , and substituting into Equation A-48 for the available effective green time, yields a more general expression of variables

$$\left(\frac{vC}{sX} \right)_{ia} + \left(\frac{vC}{sX} \right)_{ib} + \left(\frac{vC}{sX} \right)_{ic} = C_i - L_3 \quad (A-50)$$

Dividing by the cycle results in the fundamental capacity equation for intersections of

$$\left(\frac{v}{sX}\right)_{ia} + \left(\frac{v}{sX}\right)_{ib} + \left(\frac{v}{sX}\right)_{ic} = \frac{C_i - L_3}{C_i} \quad (\text{A-51})$$

Letting $Y = v/s$ be the flow ratio of demand flow to saturation flow, then

$$\left(\frac{Y}{X}\right)_{ia} + \left(\frac{Y}{X}\right)_{ib} + \left(\frac{Y}{X}\right)_{ic} = \frac{C_i - L_3}{C_i} \quad (\text{A-52})$$

Defining a new term, the pie ratio for a critical phase m , to be

$$\pi_{im} = \left(\frac{Y}{X}\right)_{im} = \left(\frac{v}{sX}\right)_{im} \quad (\text{A-53})$$

which is the proportion of the available green time needed (used) to serve the demand volume for a given geometry and degree of saturation, then for each intersection, i , the pie ratio must equal the proportion of the cycle time available for moving traffic

$$\pi_i = \sum_{m=1}^{m=\infty} \pi_{im} = \frac{C_i - L_3}{C_i} = 1 - \frac{L_3}{C_i} \quad (\text{A-54})$$

For planning and design purposes, the fundamental capacity equation (A-51) may be solved for critical service volumes for given average/design assumptions. First, a signal timing strategy is usually assumed that provides equal v/c ratios for all critical phases such that

$$X_{ia} = X_{ib} = X_{ic} = X_{im} = X_i \quad (\text{A-55})$$

following the well-known Webster strategy which also tends to minimize intersection delay. Thus, for equal degrees of saturation X_{im} for all critical movements m , $X_{im} = X_i$, Equation A-51 becomes

$$\frac{1}{X_i} \left(\frac{v}{s}\right)_{ia} + \frac{1}{X_i} \left(\frac{v}{s}\right)_{ib} + \frac{1}{X_i} \left(\frac{v}{s}\right)_{ic} = 1 - \frac{L_3}{C_i} \quad (\text{A-56})$$

so that for the three critical phases at intersection i , the total intersection flow ratio is

$$Y_i = \sum_{m=1}^{m=3} Y_{im} = \sum_{m=1}^{m=3} \left(\frac{v}{s} \right)_{im} = X_i \left(1 - \frac{L_3}{C_i} \right) \quad (\text{A-57})$$

For planning and design purposes, it is also convenient to work with locally adjusted average saturation flow values, and per lane volumes, to estimate the resulting flow ratios;

$$Y_{im} = \left(\frac{v}{s} \right)_{im} = \frac{1}{s_o \prod_j (f_j)} \left(\frac{v}{N} \right)_{im} \quad (\text{A-58})$$

or on a per lane basis

$$Y_{im} = \frac{1}{s_{ilm}} \left(\frac{v}{N} \right)_{im} = \frac{v_{ilm}}{s_{ilm}} \quad (\text{A-59})$$

such that the equivalent critical lane results are

$$\frac{v_{ila}}{s_{ila}} + \frac{v_{ilb}}{s_{ilb}} + \frac{v_{ilc}}{s_{ilc}} = X_i \left(1 - \frac{L_3}{C_i} \right) \quad (\text{A-60})$$

Assuming that $s_{ila} = s_{ilb} = s_{ilc}$ and that v_{ilb} and v_{ilc} have been adjusted slightly to equivalent through volumes (CV_{ilb} , CV_{ilc}) to keep the flow ratios the same, then the sum of the "equivalent through vehicle" critical lane service volumes for a given X_i would be

$$CV_{il} = \sum_{m=1}^{m=3} CV_{ilm} = s_{ila} X_i \left(1 - \frac{L_3}{C_i} \right) \quad (\text{A-61})$$

The average allowable service flow per critical phase, P_{il} , would be

$$P_{il} = \frac{CV_{il}}{3} = s_{ila} X_i \left(\frac{1}{3} - \frac{l}{C_i} \right) \quad (\text{A-62})$$

where l is the average lost time per critical phase. The equivalent through vehicle can represent passenger cars only, or the average fleet mix for the locale by adjusting s_{ila} accordingly. Assuming that $s_{ila} = 1800$ vphgpl for the local traffic mix, $X_i = 1.0$ for capacity flow, $l = 4.0$ sec/phase, and $C_i = 100$ sec; then the total critical lane capacity flow, CV_{il} , for intersection i from Equation A-61 would be

$$CV_{il} = 1800 \times 1.0 \left(1.0 - \frac{3(4)}{100} \right) = 1584 \text{ vphpl} \quad (\text{A-63})$$

and the resulting average critical lane capacity per phase from Equation A-62 is

$$P_{il} = 1800 \times 1.0 \left(0.333 - \frac{4}{100} \right) = 528 \text{ vphpl} \quad (\text{A-64})$$

Other critical lane service volumes can be calculated for the degree of saturation selected. In addition, an operational analysis could be conducted to determine the overall degree of saturation (X_i) on the intersection produced by the critical volume loading present.

Coordinated Intersections. Traffic signals at most interchanges are coordinated to improve overall traffic operations because the intersections are closely spaced and traffic volumes are often high. Under these conditions, coordination generally improves operational reliability and reduces internal queueing, queue spillback into upstream intersections, and the threat of operational gridlock. At lower traffic volumes, cross arterial progression can also be provided in most cases to reduce queueing. The overall quality of operations depends on the features of the signal system deployed, the progression provided, and other factors.

Coordinated signal operation implies that the two intersections no longer can operate independently. The first coordination constraint applied usually is that the two cycles, C_u and C_d , must be the same length at a particular point in time. This cycle constraint is true whether the interchange is controlled by one controller unit or by two. Thus

$$C = C_i = C_u = C_d \quad (\text{A-65})$$

Delays and queueing incurred on the two external approaches to the lower-volume intersection will increase due to operations at a suboptimal (higher) cycle, assuming that the interchange would operate at the cycle of the higher-volume intersection. Signal coordination within the interchange can almost eliminate outbound delays, but arterial coordination is required to mitigate arterial approach delays incurred on the inbound Phase a .

Cycle timing is critical to interchange operations for many complex and interrelated reasons. Capacity analysis should recognize this central control parameter. The first consideration in establishing cycle time is to determine whether interchange control is coordinated with the cross street arterial or not. If yes, then cycle times will be fixed for various coordination time periods. The interchange's local control may be coordinated pretimed or coordinated (semi) actuated, but the system cycle length is the same in either case.

The second consideration in establishing cycle time, given that cross arterial coordination is not effected, is whether the interchange's controller(s) is pretimed or coordinated (semi) actuated. Local interchange coordination is presumed. Pretimed systems are fixed to prescribed durations regardless of current traffic demands and local capacity provided. Presumably forecasted traffic demands and estimated roadway capacity were considered in the initial selection of cycle times, but conditions may have changed. For locally coordinated actuated control, field measurements are highly recommended over unproven analytical models or engineering judgement. Actuated control is reactive to queues and tends to be unstable in congested operational environments and, therefore, may generate extremely long cycles and undesirable green splits based only on maximum green settings, which may further exacerbate the congestion.

Another critical operational feature affecting interchange traffic operations is whether the green splits at the two intersections depend on one another. If two separate controllers are used (or independent rings are provided using one controller), then capacity estimates are best given by Equation A-51. If phasing is dependent (e.g. only one controller is used with dependent phases), then more complex demand analysis is required consistent with the type of interchange and phasing used. Operations at "four-phase" diamonds is a classic example, but others abound.

Coordinated Diamond Interchanges. A popular diamond interchange signal timing strategy is "four-phase with two overlaps." In addition to having many of the above timing features (e.g., same cycle, a fixed sequence, and a fixed offset), this strategy also provides quality platoon progression for the arterial traffic in both directions of flow through the interchange by either special controller design, or by judicious signal timing. By either method, the following signal timing relationship must occur (6) for four-phase with two overlaps signalization to result

$$G_{ua} + G_{ub} + G_{da} + G_{db} = C_i + \Phi \quad (A-66)$$

where it is presumed that G_{ua} and G_{ub} are the thru and ramp phases, respectively, and Φ is the total interchange "overlap" for both directions of flow through the interchange (5). This operational requirement provides great progression for the arterial traffic passing through the interchange. However, the sum of the four external phases serving traffic input to the interchange is fixed for a constant cycle and does not have full flexibility to optimally adjust to all possible traffic patterns that might arise at the interchange. In addition, this constraint (Equation A-66) further implies that the sum of the two internal left turns within the interchange is also fixed (6) at

$$G_{uc} + G_{dc} = C_i - \Phi \quad (A-67)$$

because the sum of the conflicting phases must equal two cycles (Equations A-66 and A-67). This fact may be a significant constraint on the optimal solution depending on the cycle time, minimum green times required, and traffic pattern being serviced (6). The computer signal timing program PASSER III, developed at Texas Transportation Institute, contains these strategies for developing optimal signal timing for all forms of two-level signalized diamond interchanges. PASSER III also contains delay/difference-in-relative offset algorithms, somewhat like TRANSYT 7F, to evaluate traffic performance for given signal timings (5).

A.2.10 Interchange Capacity Studies

Most operational studies reported in the literature have been conducted on signalized tight urban diamond interchanges (TUDI), probably because of operational problems due to the closely spaced ramp terminal signals. These signal spacings in many urban areas are often on the order of 70-100 meters. Capelle and Pinnell wrote in the early 1960's (7) that:

"After studying the problem of evaluating the capacity of diamond interchanges, it was determined that it would be necessary to consider the two signalized intersections as a single unit. This is due primarily to the requirements of signalization which should perform two basic functions. These functions are as follows: (a) all highway volume conflicting movements at both intersections must be separated, and (b) storing of vehicles between the two intersections must be kept to a minimum due to limited distance between them."

The above study was conducted at diamonds that were "locking up" using three-phase operation. Capelle and Pinnell (7) then tested a new phasing plan that has since become known as "four-phase with two overlaps" to improve operations. They proposed a method for calculating interchange capacity for this strategy which they termed the critical lane capacity, CV_c , as being

$$CV_c = \frac{(C + 4 + 4D + 8)}{H} \frac{3600}{C} \quad (A-68)$$

where C represents the cycle length, D the starting delay, and H the saturation headway. The starting delay used for their calculations was the time required for the first two vehicles in a lane to enter the intersection. The critical lane capacity, CV_c , represents the maximum sum of the four critical lane approach volumes from the four external approaches to the interchange. Capelle and Pinnell (7) computed critical lane capacity using $H = 2.1$ seconds/vehicle and $D = 5.8$ sec. A 2.1 seconds saturation headway is equivalent to a saturation flow of 1,714 vphgpl in the HCM.

In a recent HCQSC literature review, Lee (8) updated the Capelle/Pinnell Method to include recent interchange studies in the Phoenix area using Hook's data (9). The equivalent Hook's value for starting delay and average saturation headway, weighted by the volumes of the movements were

7.1 and 1.89 seconds, respectively. Capelle and Pinnell assumed that the starting delay was incurred by the first two vehicles, while Hook assumed the third vehicle in queue. Following some additional modifications, Lee developed the equivalent sum of critical lane volumes shown in column 3 of Table A-15. Column 4 was derived by TTI from work on NCHRP 3-40 (10) described below.

TABLE A-15. Critical Lane Capacity of Diamond Interchanges by Capelle/Pinnell, Lee, and NCHRP 3-40

Cycle Length	Original Critical Lane Capacity ⁽¹⁾	Updated Critical Lane Capacity ⁽²⁾	NCHRP 3-40 Critical Lane Capacity ⁽³⁾
40	1,611	1,821	1,980
50	1,635	1,838	1,984
60	1,650	1,849	1,987
70	1,660	1,857	1,989
80	1,668	1,863	1,990
100	1,674	1,872	1,992
180	1,692	1,886	1,996
⁽¹⁾ Based on 1961 values for starting delay, D, of 5.8 seconds and average headway, H, of 2.1 seconds (7). ⁽²⁾ Based on current values of D = 5.2 seconds and H = 1.9 seconds, (8). ⁽³⁾ Based on studies published in References 9,10,11,12, and 13.			

About the time Lee was performing his analysis, Messer and Bonneson were publishing a more updated version of these formulations in NCHRP 345 (10) based on the Phoenix data (9,11) plus additional data collected in Florida and Texas by Bonneson (12,13,14). The basic "saturation flow" model followed the earlier work by Messer in 1975/76 (15) described in the previous section. The critical movement model for interchange capacity is, however, basically a reciprocal formulation of the headway model presented by Capelle and Pinnell in 1961, but with slightly different saturation flow and start-up delay calibration factors. Messer, et al, showed in NCHRP 3-40 (10) that the sum of the four critical external inputs to a diamond interchange operating with four-phase with two external overlap signal timing has a critical lane capacity of

$$CV_{Ic} = CV_{ua} + CV_{ub} + CV_{da} + CV_{db} = s_{ila} \left[1 + \frac{\Phi}{C} - \frac{nl}{C} \right] \quad (A-69)$$

where:

CV_{ic}	=	sum of interchange critical input volumes, vphpl;
s_{ila}	=	adjusted saturation flow for all critical input phases, vphgpl; (= 3600/H above)
Φ	=	total interchange overlap, sec;
n	=	number of critical input phases, $n = 4$;
l	=	average phase lost time, sec; and
C	=	cycle length, sec.

The total interchange overlap is a function of the center-to-center spacing between the ramps, arterial grade, and quality of cross street coordination. An equation for estimating Φ for nominal conditions of no grade nor coordination is given in the PASSER III users manual as

$$\Phi = 2 [0.50 + \sqrt{0.137L}] \quad (A-70)$$

for an intersection spacing of L meters subject to a maximum speed of 50 km/h (30 mph). Higher maximum speeds and lower overlaps may be appropriate under well coordinated cross street operations.

Messer's updated interchange/intersection capacity model (10) can be applied to other types of two-level signalized interchanges operating in a coordinated system/common cycle mode. The model estimates the sum of critical lane volumes, CV_{ic} , that can be input into the interchange from the four external approaches according to the following equation

$$CV_{ic} = s_{ila} [1 + \Phi/C - n(l_s + l_e)/C] \quad (A-71)$$

where s_{ila} equals the mean saturation flow rate for all critical phases, ($s = 3,600/H$ above), Φ is the total phase overlap in the interchange, n is the number of critical phases, l_s is the platoon startup lost time (about 2.0 seconds) per phase, and l_e is the end lost time associated with ending the phase as related to the width of the interchange and the duration of the signal change intervals (the yellow warning and (all?) red clearance intervals).

The average phase capacity, P_{μ} , can be estimated by dividing by the number of critical input phases, n , on the external approaches to yield

$$P_{\mu} = s_{ila} [1/n + \Phi/nC - (l_s + l_e)/C] \quad (A-72)$$

Equations developed in the earlier research projects to estimate the above parameters are given in the following section.

The capacity of an interchange depends on several operational parameters, as Equation A-71 illustrated. Signalized interchanges are characterized by the number of critical phases that may exist (usually 2 for 4-quad parclo, 3 for 2-quad parclo and many diamonds, or 4 for many TUDIs), by the total phase overlap Φ that may be present (usually 0.0 except for four-phase overlap systems), and by the total amount of clearance time used to safely terminate phases.

Operational Parameters. TTI used the above referenced reports (12, 13, 14) to develop several useful relationships to study interchange capacity presented in NCHRP 345 (10). Saturation flow for protected left turns at interchanges was estimated from:

$$s_{lt} = 3,600 / (1.50 + 1.11/r^{0.245}) \quad (A-73)$$

where s_{lt} is the saturation flow per lane for left turns, pcphgpl, and r is the average left turn radius of the turning maneuver, ft. A s_{lt} value of 2,000 pcphgpl is predicted at a turning radius of 200 feet (60 m). This was the maximum value of saturation flow recommended, which was also the assumed value for through movements under ideal conditions (10).

The application of Equation A-73 to signalized interchanges could easily follow HCM procedures where an adjustment factor for left turns, based on the radius of turn, is applied to the saturation flow for "ideal conditions" of 2,000 pcphgpl for interchanges. A protected left turn factor, f_{lt} , as derived from Equation A-73, would then be (10)

$$f_{lt} = 1.0 / (0.833 + 0.617/r^{0.245}) \quad (A-74)$$

Quantifying the effects of trucks and other heavy vehicles on saturation flow at traffic signals is another matter. The passenger car equivalency (PCE) used in the 1985 HCM was 1.5 for through traffic only, but Molina (16) showed that the PCE for typical through moving urban truck traffic averages about 2.7, with a range from 1.7 for small trucks to 3.7 for five-axle trucks. The 1994 HCM provided an small increase in the average PCE of heavy vehicles at signalized intersections to 2.0. Heavy truck volumes turning left from off-ramps under tight geometric conditions could have an even larger impact on interchange capacity.

Estimation of phase clearance lost times for interchanges is somewhat complicated by the diverse internal clearance paths taken by vehicles traveling through the interchanges and the different forms of interchanges. Poppe (11) estimated the clearance lost time, for a SPUI as related to the signal change interval, CI, of the phase as

$$l_c = 0.95 CI - 2.3 \quad (A-75)$$

with an R-square of 0.97, suggesting an "end use" of the initial portion of the yellow interval of 2.3 seconds. These results compared well with Bonneson's Florida data (12) and with those collected within this study and reported in Appendix C.

REFERENCES

1. Munjal, P.K. "An Analysis of Diamond Interchanges." *Transportation Research Record 349*, Transportation Research Board, Washington, D.C. (1971) pp. 47-64.
2. Highway Capacity Manual." *Special Report 209*, Third Edition, Transportation Research Board, Washington, D.C. (1994).
3. Fambro, D. B., Messer C.J., and Andersen, D.A. "Estimation of Unprotected Left-Turn Capacity at Signalized Intersections." *Transportation Research Record 644*, Transportation Research Board, Washington, D.C. (1977), pp.113-119.
4. Fambro, D.B., Rouphail, N.M., Sloup, P.R., Daniel, J.R., Li, J., Anwar, M., and Engelbrecht, R.J. "Highway Capacity Revisions for Chapters 9 and 11." Federal Highway Administration, FHWA-RD-96-088, Washington, D.C. (1996).
5. Messer, C.J., Fambro, D.B., and Richards, S.H. "Optimization of Pretimed Signalized Diamond Interchanges." *Transportation Research Record 644*, Transportation Research Board, Washington, D.C. (1977) pp. 78-84.
6. Messer, C.J., Whitson, R.H., and Carvell, J.D. "A Real-Time Frontage Road Progression Analysis and Control Strategy." *Transportation Research Record 503*, Transportation Research Board, Washington, D.C. (1974) pp. 1-12.
7. Capelle, D.G. and Pinnell, C. "Capacity Study of Signalized Diamond Interchanges." *Highway Research Board Bulletin 291*, Highway Research Board, Washington, D.C. (1961).
8. Lee, J.C. "Review of Diamond Interchange Analysis Techniques: Past and Present." Synthesis Paper for Committee A3A10. Transportation Research Board, Washington, D.C. (1992).
9. Hook, D.J. and Upchurch, J. "Comparison of Operational Parameters for Conventional Diamond Interchanges and Single-Point Diamond Interchanges," *Transportation Research Record 1356*, Transportation Research Board, Washington, D.C. (1992).
10. Messer, C.J., Bonneson, J.A., Anderson, S.D., and McFarland, W.F. "Single-Point Urban Interchange Design and Operational Analysis." NCHRP 345, Transportation Research Board, Washington, D.C. (1991).
11. Poppe, M.J., Radwan, A.E., and Matthias, J.S. "Some Traffic Parameters for the Evaluation of the Single-Point Diamond Interchange." *Transportation Research Record 1303*, Washington, D.C. (1991).

12. Bonneson, J.A. "A Study of Headway and Lost Time at Single-Point Urban Interchanges." *Transportation Research Record 1365*, Transportation Research Board, Washington, D.C. (1992).
13. Bonneson, J.A. "Factors Affecting Bridge Size and Clearance Time of the Single-Point Urban Interchange." *Journal of Transportation Engineering*, Vol. 119, No. 1, American Society of Civil Engineers, New York (1993).
14. Bonneson, J. A., "Operational Efficiency of the Single-Point Urban Interchange." *ITE Journal*, Vol 62, No. 6. Institute of Transportation Engineers, Washington, D.C. (1992).
15. Messer, C.J., and D.J. Berry. "Effects of Design Alternatives on Quality of Service at Signalized Diamond Interchanges." *Transportation Research Record 538*, Transportation Research Board, Washington, D.C. (1975).
16. Molina, C.A. "Passenger Car Equivalents of Trucks at Signalized Intersections." *ITE Journal*, Vol 54, No. 9. Institute of Transportation Engineers, Washington, D.C. (1988).

APPENDIX B

FIELD STUDY DATA COLLECTION AND REDUCTION

Several analytic models were developed during this study to facilitate the evaluation of interchange ramp terminal capacity and level of service. This appendix provides a description of the database used to calibrate these models. Specifically, it describes the composition of the database, the field study sites, the methods used to collect the data, and finally, some summary statistics of the reduced data. Subsequent appendices describe the development and calibration of the analytic models as well as their application to the interchange evaluation process.

B.1 DATABASE COMPOSITION

This section describes the traffic flow problems associated with interchange areas, as they relate to the objectives of this research. The problems of primary interest are those occurring on the arterial cross street at or between the interchange ramp terminals and any adjacent, closely-spaced intersections. Initially, these flow problems are described in the context of models needed to describe the problem's effect on arterial performance. Then, the variables included in these models are identified in the context of defining a data collection plan.

B.1.1 Traffic Flow Problems Associated With Interchange Ramp Terminals

The findings from the survey of practitioners (see Appendix A) indicated that there were several types of traffic flow problems associated with signalized interchange ramp terminals. The impact of these flow problems was typically amplified by relatively short distances (as measured along the arterial cross street) between the terminals or between a terminal and an adjacent signalized intersection. These flow problems were broadly categorized as: (1) midblock turbulence (i.e., weaving) and unbalanced lane volumes that stem from high-volume turn movements in the interchange vicinity; and (2) flow restriction or impediment to discharging queues due to a relatively near downstream traffic queue. Four models were developed to facilitate the evaluation of these flow problems. The variables included in these models were used to identify the data needed for model calibration. The four models include:

1. **Capacity Model.** This model quantifies the effect of downstream traffic conditions on the traffic characteristics used to estimate the capacity of left-turn and through movements at interchange ramp terminals and adjacent intersections. These characteristics include start-up lost time, saturation flow rate, and clearance lost time. The capacity of an upstream signal phase has been found to be adversely affected by the close proximity of a downstream queue, particularly when the queue spills back into the upstream intersection. The form of this model is described in Appendix C.
2. **Approach Lane Utilization Model.** This model quantifies the extent of unbalanced lane use in multi-lane lane groups. On a cycle-by-cycle basis, many drivers in the interchange area tend

to use one lane of a multi-lane lane group more than the others; they rarely choose the lane with the fewest vehicles in it. One possible reason for this clustering in interchange areas may be driver desire to "preposition" for a downstream turn. In this situation, drivers in interchange areas position their vehicles in the appropriate inside (or outside) lane at an upstream ramp terminal or adjacent signalized intersection in anticipation of a turn at the next terminal or intersection. The frequency of this behavior is increased when turn volumes are high or the distance between terminals is too short for "comfortable" lane changing. The form of the lane utilization model is described in Appendix C.

3. **Queue Length Model.** This model can be used to convert a predicted queue length from the number of equivalent passenger cars to units of distance (e.g., meters). This queue length conversion model was found to be an essential component of the capacity model. This model was also incorporated into the queue interaction model developed for this research, as described in Appendix D.
4. **Arterial Weaving Model.** This model quantifies the effect of weaving activity on the efficiency of arterial traffic flow. The weaving maneuver that is predominate in interchange areas is the off-ramp right-turn movement that weaves across the arterial to make a left-turn at the downstream signalized intersection. This maneuver has been observed to cause significant turbulence in the arterial traffic flow resulting in significant increases in travel time and, in some cases, lengthy queues on the off-ramp. The form of this model is described in Appendix E.

B.1.2 Database Elements

The data needed to calibrate the aforementioned models can be categorized as: (1) basic traffic characteristics (model inputs), (2) traffic performance measures (model outputs), (3) signal controller settings, (4) traffic control features, and (5) geometric data. The elements that comprise the first two categories are dynamic and were collected continuously during the field studies. These elements are listed in Table B-1.

The latter three categories represent static data types. They were often measured prior to the start of the study. Elements of each of these three categories are listed below.

- **Signal Controller Settings.** This category includes the traffic signal controller settings and operation. In particular, cycle times, coordination offsets, signal phase sequence, change interval, and phase splits were collected during each study. If the controller had one or more actuated phases, then the individual actuated phase durations were recorded with a computer connected to the signal controller.
- **Traffic Control Features.** This category includes speed limit, traffic control signs, and pavement markings. This information was of a general nature and was used to describe the character of the interchange area.

Table B-1. Traffic-related database elements

Category	Data Type	Model			
		Phase Capacity	Lane Utilization	Queue Length	Weaving
Traffic Characteristics	Discharge Headway	C ¹			
	Discharge Speed	C			
	Start-Up Lost Time	C			
	Count & Location of Cars on the Downstream Link at the Start of Green	V			
	Queued Driver Starting Reaction Time			V	
	Queued Vehicle Storage Length			V	
	Traffic Demands by Lane and Movement		V		
	Travel Path Matrix (O-D's)		V		V
	Weave & Non-Weave Vehicle Travel Time				V
	Weave & Non-Weave Flow Rate per Lane				V
Performance Measures	Saturation Flow Rate	C			
	Lane Utilization Percentage		V		
	Queue Length			V	
	Weaving Vehicle Speed				V

Notes:

1 - Data collection method: C - computer-monitored tape switches; V - video tape records.

- **Geometric Data.** This category includes geometric information along the arterial and at each intersection. Arterial information includes cross section, distance between intersections, turn bay lengths, and lane assignments. Intersection information includes approach grade, skew angle, and turn radii.

B.2 STUDY SITE DESCRIPTION

B.2.1 Study Site Characteristics

A list of desirable characteristics for the field study sites was prepared based on information obtained from the survey of practitioners and the insights obtained while formulating the aforementioned models. These site characteristics are described in the following paragraphs.

Interchange Types. The study sites were selected to collectively include the two basic forms of service interchange commonly used in suburban and urban areas: the diamond and the partial cloverleaf (or parclo) interchanges. Variations of these two interchange forms stem from variations in the distance between the ramp terminals and the routing of the left and right-turn movements (i.e., through the signal or via a loop ramp). Further assessment of the correlation between interchange type, the extent of its operational problems, and its frequency of application in

urban areas led to the following six interchange types being identified as the most appropriate candidates for the field studies:

Diamond Interchange

1. Compressed Diamond
2. Tight Urban Diamond (without frontage roads)
3. Tight Urban Diamond (with frontage roads)
4. Single Point Urban Diamond

Partial Cloverleaf (Parclo)

5. Parclo B (2-quad)
6. Parclo AB (2-quad)

Study Types. The development of the data collection procedure was strongly influenced by: (1) the need to optimize the quality and quantity of traffic-related characteristics (see Table B-1), and (2) the capabilities of the computer and video data collection equipment available to the study team. Based on a thorough examination of all feasible collection procedures, it was concluded that two different types of field study were needed.

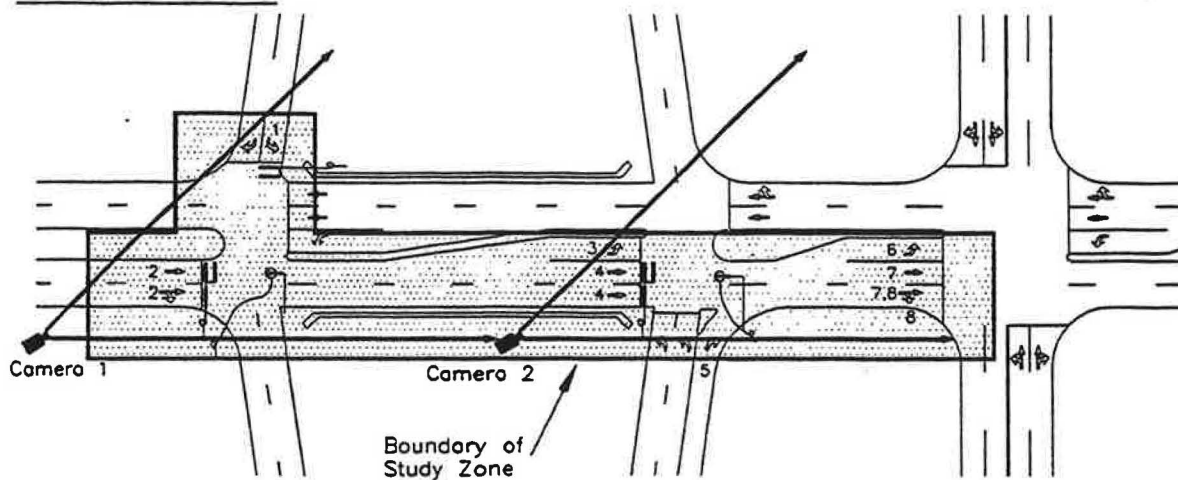
One type of study required the equipment to be deployed for the purpose of collecting the capacity and lane utilization model data. The data needs of these models required the concurrent study of one direction of travel along two successive, arterial street segments. These segments included the section between the two ramp terminals and the section between the ramp terminal and adjacent intersection. This study type was referred to as a *Capacity Study*. A typical data collection setup for this study is shown in Figure B-1. This figure shows the setup for the diamond interchange form, a similar setup was used for the parclo form.

As shown in Figure B-1, there are two possible study "cases" at an interchange. These cases are named the "downstream" and "upstream" cases. The name of each case identifies the orientation of the adjacent intersection with respect to the travel direction studied. The field studies were designed to include a mixture of both cases. This approach permitted an examination of the effects of the adjacent intersection on both interchange inflow and outflow.

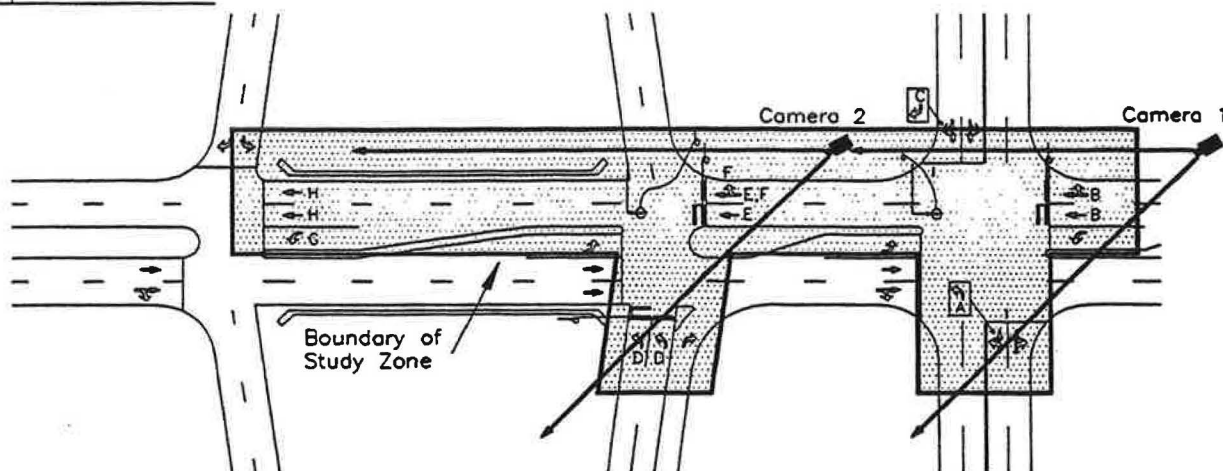
The second type of study required the equipment to be deployed for the purpose of collecting the weaving data. This study focused on the off-ramp right-turn movement weaving across the through traffic to make a left-turn at the adjacent, downstream intersection. To collect this data, the equipment was deployed at each end of the arterial weaving segment. This study type was referred to as a *Weaving Study*. A typical data collection setup for this study is shown in Figure B-2.

Queue length and starting reaction time data, needed for the Queue Length Model, were collected with both study types. The data needs of this model required only a view of the front and rear of the traffic queue. As this type of view was generally available from either study type, a third study type was not developed to collect this data.

Downstream Case



Upstream Case



LEGEND

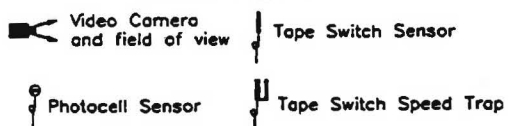


Figure B-1. Capacity study data collection setup for a diamond interchange.

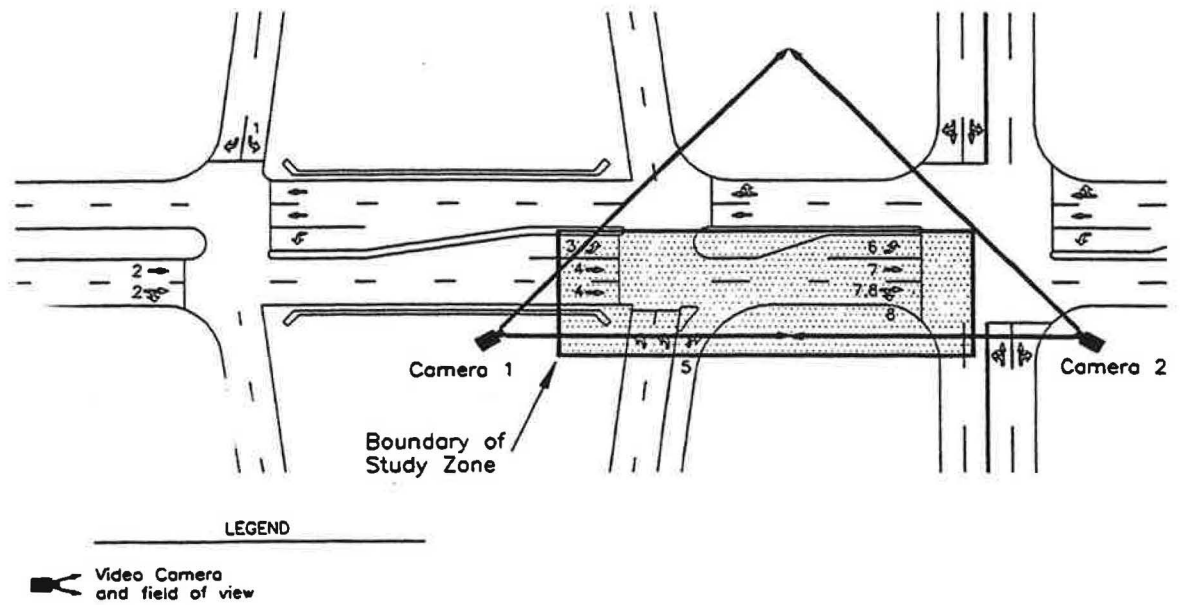


Figure B-2. Weaving study data collection setup.

The road segments and corresponding traffic movements considered during each study are shown (by number or letter) in Figures B-1 and B-2. These figures also show the locations of the tape switch sensors, photocell sensors, and video cameras. The sensors were monitored by a computer, which also served as the data processing and recording device. This computer processed the sensor input and automatically converted it into more desirable forms such as phase duration and discharge headway. The video cameras were used to record the vehicle locations on the downstream segments, travel paths, and weaving/non-weaving vehicle characteristics. The actual data were manually extracted from videotapes during playback in the laboratory, subsequent to the field studies, and processed into more useful forms at that time. The specific types of data collected by each collection system are listed in Table B-1.

Geometric and Traffic Demand Criteria. The selection of specific field study sites (i.e., interchanges) was based on their degree of compliance with the following geometric and traffic demand criteria.

Geometric Criteria

- | | |
|-----------------------------------|--|
| 1. Cross Section: | 4 or 6 through traffic lanes |
| 2. Ramp to Ramp Distance: | 60 to 275 meters (stop line to stop line) |
| 3. Ramp to Intersection Distance: | 60 to 275 meters (stop line to stop line) |
| 4. Adjacent Land Access: | no parking and preferably no driveways on arterial |

Traffic Demand Criteria

- | | |
|-----------------------------|------------------------------|
| 1. Arterial AADT: | 20,000 or more |
| 2. Ramp Terminal v/c Ratio: | 0.7 to 1.0 during peak hours |

Traffic Control Criteria

- | | |
|---------------------------------------|---|
| 1. Arterial Speed Limit: | 48 to 80 km/h |
| 2. Ramp & Intersection Signalization: | pretimed or semi-actuated/coordinated |
| 3. Ramp & Intersection Cycle Length: | preferably the same at both |
| 4. Arterial Left-Turn Phasing: | protected |
| 5. Arterial alignment: | preferably less than 2° horizontal curvature, less than 2% grade, and negligible skew |

In addition to these criteria, the study sites had to have frequent and recurring traffic queues on the arterial during the peak traffic periods. The intent of these criteria was to insure that the database would be representative of interchanges in urban areas with fairly typical geometrics. The requirement of "frequent and recurring queues" was a recognized deviation from the characterization of "being representative" and "typical;" however, it was a necessary extension as it produced the number of observations necessary to yield meaningful and statistically sound models.

B.2.2 Study Site Locations

In addition to the aforementioned criteria, there was a need for geographic diversity in the collective list of study sites. Study sites were identified in six geographic regions of the U.S. These regions included the Northwest, Southwest, Upper Midwest, Lower Midwest, Northeast, and

Southeast. Within these regions, highway agencies in the states with large metropolitan areas were contacted and inquiry was made as to potential study locations. Interchanges that most nearly complied with the desired criteria were identified as candidates for a preliminary site visit.

Based on the results of the preliminary visit to the candidate sites, twelve interchanges were identified as being most suitable for field study. Every effort was made to identify two interchanges for each of the six types identified in a previous section; however, this goal could not be achieved in some instances. Table B-2 describes the distribution of the twelve study sites, as categorized by interchange type and study location.

Table B-2. Interchanges studied by type and location

Interchange Type	Study Location					Total
	Nebraska	Arizona	Texas	Kansas	California	
Compressed Diamond	1			2		3
Tight Urban Diamond (no frontage)		1				1
Tight Urban Diamond (with frontage)			2		1	3
Single Point Urban Diamond		2				2
Parclo B (2-quad)					2	2
Parclo AB (2-quad)	1					1
Total:	2	3	2	2	3	12

The traffic and geometric characteristics of each site are listed in Table B-3. In general, the study sites satisfied almost all of the geometric and traffic demand criteria previously described. In a few instances, the distance to the adjacent intersection exceeded the desired 275-meter maximum distance; however, the traffic demands at these sites were sufficiently high as to precipitate the extensive queueing considered desirable for study purposes.

A capacity study was conducted at each of the twelve study sites. In addition, weaving studies were conducted at six of the sites. All total, eighteen studies were conducted in eight cities and five states.

The traffic signal characteristics of each study site are listed in Table B-4. It is interesting to note that a few of the interchanges are not coordinated with the adjacent intersection. The reason for this lack of coordination is different in each case. The Peoria Road site is not coordinated because it is currently standard practice in Arizona for the state DOT to operate the interchanges and the city to operate the adjacent intersections. The Towneast Boulevard site is not coordinated due to a lack of funds for coordination hardware. The Stevenson Boulevard site is not coordinated because the existing coordination hardware failed in service and resources were not available to replace it.

Table B-3. Traffic and geometric characteristics of the study sites

Interchange Type	Arterial	City, State	Arterial AADT	Arterial Thru Lanes	Ramp to Ramp Distance (meters) ¹	Ramp to Intersection Distance (meters) ¹	Speed Limit (km/h)
Compressed Diamond	Metcalf Ave 110th to I-435	Overland Park, Kansas	58,600	6	200	204	72
	75th Street I-35 to Frontage	Overland Park, Kansas	32,000	4	174	155	56
	Maple Street 102nd to I-680	Omaha, Nebraska	34,200	4	268	198	72
Tight Urban Diamond	Peoria Road 25th Ave. to I-17	Phoenix, Arizona	34,400	6	107	276	64
	Mathilda Ave SR-237 to Ross	Sunnyvale, California	34,540	6	87	110	72
Texas Diamond	Arapaho Road US75 to Greenville	Richardson, Texas	39,000	6	99	265	64
	Towneast Blvd Emporium to I-635	Mesquite, Texas	35,000	6	137	223	56
Parclo AB (2 quad)	60th Street I-80 to Grover	Omaha, Nebraska	31,800	4	259	216	64
Parclo B (2 quad)	Somersville Rd Delta Fair to SR-4	Antioch, California	39,700	4	265	119	56
	Stevenson Blvd Balentine to I-880	Newark, California	55,600	4	264	157	56
Single Point Urban Diamond (SPUI)	7th Street I-10 to McDowell	Phoenix, Arizona	42,000	6	78	331	56
	Indian School Rd 16th St. to SR-51	Phoenix, Arizona	54,500	6	91	316	56

Notes:

- 1 - Distance measured from stop line to stop line in the same direction, except at SPUI's. At SPUI's, the "same direction" concept is also applied but the opposing direction through stop line is used as the reference point at the second ramp terminal (since the through stop line at the second ramp terminal does not exist at the SPUI).

Table B-4. Signal characteristics of the study sites

Interchange Type	Arterial	Interchange Signal Control				Adjacent Intersection Signal Control		
		Arterial Left-Turn Protection	No. of controllers	Control Type ¹	Cycle Length ² (sec)	Control Type ¹	Coordination ³	Cycle Length ² (sec)
Compressed Diamond	Metcalfe Ave	Protected	2	SA	126, 105	SA	C	126, 105
	75th Street	Protected	1	SA	82, 90	SA	C	82, 90
	Maple Street	Prot./Perm.	1	SA	90	SA	C	90
Tight Urban Diamond	Peoria Road	Prot./Perm.	1	FA	<u>166</u>	SA	N	90
	Mathilda Ave	Protected	2	SA	100	SA	C	100
Texas Diamond	Arapaho Road	Protected	1	SA	90, 120, 128	SA	C	90, 120, 128
	Towneast Blvd	Prot./Perm.	1	FA	<u>118</u>	FA	N	<u>113</u>
Parclo AB (2 quad)	60th Street	Prot./Perm.	2	SA	90	SA	C	90
Parclo B (2 quad)	Somersville Rd	Protected	2	SA	110	SA	C	110
	Stevenson Blvd	Protected	2	SA	110, 100	FA	N	<u>124</u>
Single Point Urban Diamond	7th Street	Protected	1	SA	90	SA	C	90
	Indian School Rd	Protected	1	SA	78, 90, 102	SA	C	78, 90, 102

Notes:

1 - FA = fully actuated, SA = semi-actuated.

2 - Values listed are those observed during the study periods. Underline values are averages.

3 - C = coordinated with interchange signal(s), N = not coordinated with interchange signal(s)

B.3 DATA COLLECTION**B.3.1 Approach**

The data collection equipment used to collect the field data included video cameras and computer-monitored tape switch sensors placed in the traffic lanes. As described in a preceding section, the equipment deployment followed one of two study types (i.e., a capacity or weaving study). Data collected during these studies is described in this section.

All data were collected during weekday, daytime periods between the hours of 7:00 am and 7:00 pm. The study period generally included the hours of peak traffic demand at the study site. Data were not collected during inclement weather nor during unusual traffic conditions (e.g., a traffic accident).

B.3.2 Capacity and Lane Utilization Data

During each of the capacity studies, the equipment was deployed in a manner consistent with that shown in Figure B-1. At some sites, the location of power lines, fences, or private property required slight deviations from the desired camera position. Typical camera positions are shown in Figure B-3 for one study site. The corresponding fields of view obtained with the video cameras are shown in Figure B-4.

Data collected during the capacity studies were used to calibrate the capacity and lane utilization models. Data for the capacity model were collected using both the computer-monitored tape switches and the video recorders at each of the twelve interchange study sites. In general, three traffic movements were monitored during each study. These movements included an interchange off-ramp or arterial left-turn, an interchange arterial through, and either the second interchange through movement (for the downstream case) or the upstream intersection through movement (for the upstream case).

The tape switches were used to record traffic flow behavior in two traffic lanes for each movement monitored. A speed trap consisting of two parallel tape switches was typically located in the inside-most lane. This trap provided information on vehicle headway, speed, acceleration, and wheelbase. A single tape switch was located in the adjacent lane. This single tape switch provided additional headway information. The computer monitoring the tape switches was also connected to the signal controller (using photo-cell sensors) and used to monitor the signal indication status. The video recorders were positioned to provide a visual record of traffic crossing the tape switches as well as a view of the downstream street segment.

Data for the lane utilization model were collected using two video cameras and recorders at each of the twelve interchange study sites. All traffic movements entering the study boundary were tracked as they traveled through the study area. The data collected included the approach lane used at each intersection or terminal and the travel time through the system.

75th Street & I-35, Overland Park, Kansas

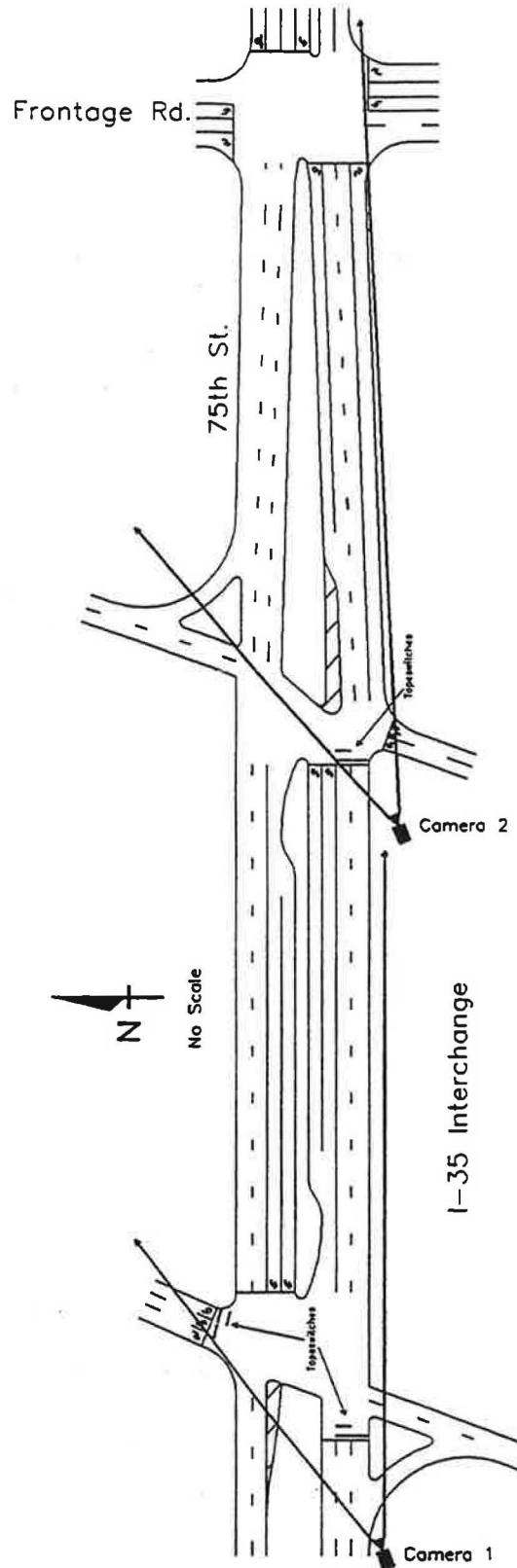


Figure B-3. Typical data collection equipment locations for a capacity study.

conducted in the 1970's, it is likely that the shorter storage length found in this research is due to reductions in average vehicle length during the last 20 years.

The average reaction time for the first-in-queue drivers was found to vary from 1.17 to 1.73 seconds among the study sites. The overall average reaction time was found to be 1.52 seconds. In contrast, the average reaction time for the subsequent queued drivers was found to be 1.06 seconds, with a range of 0.96 to 1.19 seconds among sites. This trend was expected because the first driver has more of a "surprise" situation (i.e., the signal indication changing from red to green) than the subsequent queued drivers who can look ahead, see that the indication is green, and anticipate their time of departure. As a result, the first drivers require slightly more reaction time than subsequent queued drivers.

The average reaction times found in this research are generally consistent with those reported in the literature. Specifically, the study by Messer and Fambro (2) found that the first queued driver required about 2.0 seconds of reaction time and that subsequent queued drivers required about 1.0 seconds. An earlier study by George and Heroy (4) at five intersections found that the first queue driver required about 1.8 seconds and that subsequent queued drivers required about 1.3 seconds.

B.5.5 Weaving Model Database Summary

The analysis of the weaving model database focussed on the volume and speed of the weaving and non-weaving traffic streams. These data were collected because it was hypothesized that the volume of the two conflicting streams would affect their individual running speeds through the weaving section. It was theorized that these speeds would decrease with increasing volume. The weaving movement that is the subject of this analysis is the off-ramp right-turn movement that weaves across the arterial to make a left-turn at the downstream signalized intersection. The average volumes and speeds through the weaving section for the six study sites are listed in Table B-12.

Table B-12. Weaving model database - summary statistics

Variable	Statistic			
	No.	Average	Minimum	Maximum
Volume				
Total arterial volume entering weaving section, vph	6	1,409	954	1,813
Arterial lane volume entering weaving section, vphpl	6	575	465	856
Weaving volume (off-ramp right to downstream left), vph	6	151	100	230
Speed				
Arterial veh. spot speed at entry to weaving section, m/s	324	14.1	10.3	19.3
Arterial veh. running speed through weaving section, m/s	324	10.6	8.3	12.1
Arterial veh. speed reduction due to weaving activity, m/s	324	3.4	1.4	7.2
Weaving veh. running speed through weaving section, m/s	421	8.0	6.6	10.3

As the volumes in Table B-12 indicate, the six study sites have relatively high weaving volumes. On average, the weaving vehicles accounted for about one-half of the off-ramp right-turn volume at any one site. The arterial lane volumes were also relatively high such that weaving opportunities were limited during a significant portion of the signal cycle. It should be noted that the off-ramp right-turn movement at three of the sites was signalized (with right-turn on red allowed); the other three were yield-controlled.

Two types of speed statistic were reported for the arterial vehicles. One statistic is the spot speed of the arterial vehicles at a point just upstream of the off-ramp. The second statistic is the running speed of the same arterial vehicles. This speed related the distance traveled through the weaving section to the corresponding travel time. The distance and time were measured from the point of entry to the weaving section to the downstream intersection stop line or to the first point of joining the stopped queue associated with the downstream signal, whichever was reached first. The running speed measured in this manner was reasoned to be the better measure of weaving vehicle impact because it excluded the effect of downstream signal delay on the speed estimate.

The difference between the arterial spot speed and the running speed is an indicator of a speed reduction in the weaving area due to weaving activity. The average speed reduction at the study sites was 3.4 m/s. This statistic is more useful than the spot or running speeds alone because it eliminates the effect of differing speed limits among the sites. A preliminary examination of this speed difference indicates a strong correlation between it and the total arterial and weaving volumes. Increases in either volume level tended to increase the speed reduction.

The average weaving vehicle speed is also shown in Table B-12. The weaving vehicle speed tends to be lower than that of the arterial vehicles because the weaving vehicle enters the weaving section at a relatively slow speed due to the ramp control (i.e., signal or yield sign). Some preliminary analysis of this speed indicates that it decreases with increasing arterial lane volume.

B.6 APPENDIX B REFERENCES

1. *TRB Special Report 209: Highway Capacity Manual*, 3rd ed. TRB, National Research Council, Washington, D.C. (1994).
2. Messer, C.J., and Fambro, D.B. "Effects of Signal Phasing and Length of Left-Turn Bay on Capacity." In *Transportation Research Record 644*, TRB, National Research Council, Washington, D.C. (1977) pp. 95-101.
3. Herman, R., Lam, T., and Rothery, R.W. "The Starting Characteristics of Automobile Platoons." *Proc., 5th International Symposium on the Theory of Traffic Flow and Transportation*, American Elsevier Publishing Co., New York (1971) pp. 1-17.
4. George, E.T., and Heroy, F.M. "Starting Response of Traffic at Signalized Intersections." *Traffic Engineering*, Institute of Transportation Engineers, Washington, D.C. (July 1966) pp. 39-43.

APPENDIX C

CAPACITY CHARACTERISTICS FOR INTERCHANGES AND CLOSELY-SPACED INTERSECTIONS

This appendix describes the development, calibration, and application of models that collectively can be used to predict the capacity of traffic movements at signalized interchange ramp terminals and other closely-spaced intersections. Specifically, these models predict three important capacity characteristics: saturation flow rate, start-up lost time, and end lost time. The model for each characteristic is developed using theoretic constructs that incorporate the factors that have an influence on the characteristic's magnitude or duration. These models were calibrated with data collected at twelve interchanges (the field studies are described in Appendix B). It should be noted that the traffic characteristics described in this appendix reflect passenger car performance as all heavy vehicles were excluded from the database.

This appendix includes five main sections. The first four sections describe the development and calibration of models of saturation flow rate, start-up lost time, clearance lost time, and lane utilization. The last section describes the proposed form of these models and their application to capacity analysis.

C.1 SATURATION FLOW RATE

This section describes the development and calibration of a saturation flow rate model applicable to the signalized movements at interchanges and closely-spaced intersections. Separate models are developed for the left-turn and through movements at these junctions because of their unique operational character. In each case, the saturation flow rate model is derived from a model of the discharge headway process for queued vehicles. The saturation flow rate is defined as the minimum discharge headway reached and sustained by the discharging queue. The models are sensitive to factors that affect the discharge process at interchanges and closely-spaced intersections such as distance to the downstream queue and signal timing.

Three topics are discussed in the remainder of this section. First, the basic issues related to the measurement of the minimum discharge headway and start-up lost time are described. Then, the minimum discharge headway and resulting saturation flow rate models are developed for through movements. Finally, the minimum discharge headway and saturation flow rate models are developed for the left-turn movements. The development of start-up lost time models for each of these movements is described in the next section.

C.1.1 Minimum Discharge Headway and Start-up Lost Time

Methods of Computation. The discharge of a traffic queue, upon presentation of the green indication, is characterized by the reaction time of the queued drivers to the indication followed by

their steady acceleration to a desired discharge speed. As a result, the first few vehicles have relatively long headways; however, the headways of subsequent vehicles gradually decrease as they approach the desired discharge speed. Ultimately, the discharging stream converges to a relatively constant headway in the range of 1.8 to 2.0 seconds per vehicle. This constant value represents the minimum discharge headway H of the queue; its inverse represents the saturation flow rate s .

Typical discharge headways observed for each of the first ten queue positions are shown in Figure C-1. As this figure suggests, the minimum discharge headway is effectively reached by the sixth queue position; however, it should be noted that the headways for each subsequent queue position continue to decrease slightly. The first four or five queue positions, having headways larger than the minimum, each incur an increment of "lost-time" due to the reaction-time process and the subsequent acceleration to the desired discharge speed (i.e., the speed at saturation flow). The sum of these lost-time increments for each vehicle represents the total start-up lost time l_s of the queue.

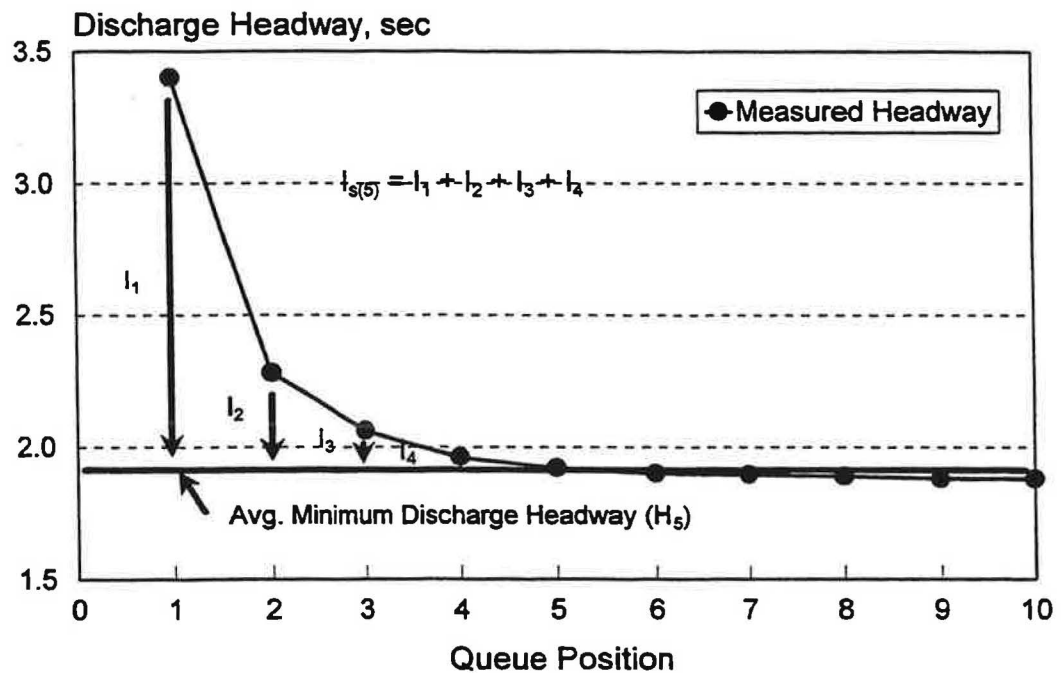


Figure C-1. Discharge headway by queue position.

In recognition of the relatively constant headway achieved by the higher queue positions, the 1994 *Highway Capacity Manual* (HCM) (1) recommends that the minimum discharge headway be estimated as the total headway of the combined higher queue positions divided by the number of vehicles observed in these queue positions. In this context, the total headway is computed as the difference between the discharge time of the last vehicle to discharge $T_{d(j)}$ and that of the last vehicle to incur some start-up lost time $T_{d(j-l)}$. Thus, the minimum discharge headway can be computed as:

$$\begin{aligned}
 H_j &= \frac{T_{d(j)} - T_{d(j-1)}}{J - (j - 1)} \\
 &= \frac{1}{J - (j - 1)} \sum_{i=j}^J h_i
 \end{aligned}
 \tag{C-1}$$

where:

H_j = minimum discharge headway based on specification of the j th queue position as the first to achieve the minimum discharge headway, sec/veh;

$T_{d(i)}$ = discharge time of the i th queued vehicle ($i = j-1$ to J), sec;

h_i = headway of the vehicle in the i th queue position, sec;

j = "specified" first queue position to discharge at the minimum discharge headway; and

J = last queue position to discharge.

The HCM (1) indicates that the saturation flow rate s can be computed from the minimum discharge headway using the following equation:

$$s_j = \frac{3,600}{H_j} \tag{C-2}$$

where:

s_j = saturation flow rate for the subject lane based on specification of the j th queue position as the first to achieve the minimum discharge headway, vphgpl.

By definition, the start-up lost time can be estimated from the discharge time of the m th vehicle and the average minimum discharge headway as follows:

$$\begin{aligned}
 l_{s(j)} &= T_{d(j)} - j H_j \\
 &= \sum_{i=1}^j h_i - (j H_j)
 \end{aligned}
 \tag{C-3}$$

where:

$l_{s(j)}$ = start-up lost time based on H_j , sec.

For most practical applications, the HCM (1) recommends that the fifth and higher queue positions can be used to estimate the minimum discharge headway (i.e., $j = 5$). In general, these averages are sufficiently accurate as to permit the estimation of the capacity of an intersection traffic movement; however, they are likely to be higher than the true minimum discharge headway H . As

Potential Bias in Headway and Lost Time Estimation. A closer examination of the data was undertaken to explore the causes of the variability in the tabulated headways shown in Table C-1. This examination focussed on the possibility that the method used to calculate the minimum discharge headway introduces some bias. Specifically, the specification of queue position five as the first to achieve the minimum discharge headway is likely to introduce some bias into the resulting average (as obtained from Equation C-1) if the headways recorded for position five at a given site are not, on average, equal to the average value of the higher queue positions.

Further examination of the data at the study sites indicated that this bias does exist and that the amount of bias is dependent on two factors. The first factor relates to the inclusion of headways from lower queue positions, that are not discharging at the minimum headway, in the estimate of the average minimum discharge headway. Specifically, the average headway for a lower queue position can be larger than the minimum discharge headway and, if included in its estimate, will bias the estimate to values larger than the true minimum discharge headway. This effect can be seen in Figure C-1 where the line representing average minimum discharge headway is "pulled" upward slightly by headway observations in queue positions five and six.

The second factor affecting the amount of bias due to the use of headways not at the minimum value relates to the frequency of headway observation at each queue position. In general, the number of observations is highest for the lowest queue position and decreases with each increasing queue position. This variation in observations can amplify the effect noted above (i.e., differences in average headway between queue positions) by giving greater "weight" in the overall average to the more frequently observed lower queue positions.

The following example is posed to illustrate the effect of the bias due to the use of lower numbered queue positions in the average minimum discharge headway computation. Consider a traffic lane at each of two different intersections. Each lane has a demand volume of 200 vphpl and an identical queue discharge character that yields the headways shown in Figure C-2. These headways represent the "true" headways that are known in advance for the purpose of this example. Inspection of this figure indicates that the true minimum discharge headway H of each lane is 1.88 sec/veh (1,915 vphgpl).

A field study is conducted to estimate the true minimum discharge headway H for each intersection through the use of Equation C-1 (i.e., H_5). Headways are measured for three hours at each intersection. The cycle length is 120 seconds at one intersection and 60 seconds at the other. These cycle lengths are found to yield average queues of 6.7 and 3.3 vpcpl, respectively, and the frequencies of observation by queue position shown in Figure C-2. The average minimum discharge headway is computed for each lane and shown in Table C-2. As this table indicates, the two estimates are different solely because of the bias effects described previously. Moreover, neither estimate yields the true minimum discharge headway of 1.88 sec/veh (although, the lane with the longer average queue length yields a closer estimate).

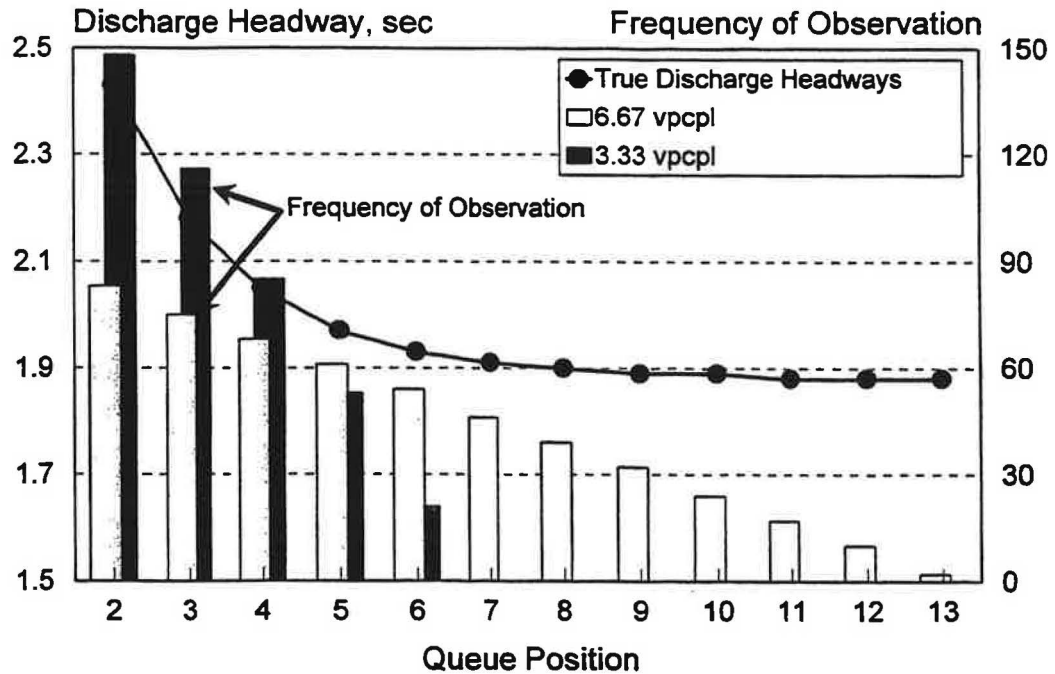


Figure C-2. Data used to demonstrate potential bias in the estimate of minimum discharge headway.

Bonneson (2) has investigated the amount of bias observed in average minimum discharge headway estimates for several left-turn and through traffic movements at several sites. From this investigation, he found that minimum discharge headway estimates obtained from the "average of the fifth through last position" H_s were typically 0.13 and 0.06 sec/veh higher for the left and through movements, respectively, than unbiased estimates of H obtained from other methods. This increase translates into the saturation flow rate being underestimated by about 110 and 50 vphgpl for the respective movements. From a practical standpoint, these differences may be considered small. However, from the standpoint of statistically quantifying the effect of various factors (e.g., lane width) on minimum discharge headway or saturation flow rate, these differences can be very important.

There are several implications that stem from the use of a biased estimate of the true minimum discharge headway. First, the estimate always exceeds the true minimum discharge headway; however, as mentioned in the previous paragraph, this effect may be small for practical purposes. Second, the estimate of the start-up lost time will be biased toward a value that is smaller than the true value because it is computed using the biased estimate of minimum discharge headway (as per Equation C-3). Third, and most important, the bias will likely cloud any statistical analysis of cause and effect by introducing added variability in the data set. As a result, the true effect of a treatment or factor (e.g., lane width, grade, interchange configuration, etc.) may be obscured by data from sites having different degrees of bias. In fact, it is this type of bias that causes most of the differences between the estimates of minimum discharge headway for the interchange configurations shown in Table C-1.

Table C-2. Effect of headway observation frequency on the estimate of minimum discharge headway

Number of Vehicles Observed:		600			
Cycle Length, sec:		120		60	
Average Vehicles / Cycle, vpcpl:		6.7		3.3	
Queue Position	True Headway ($h_{T,i}$) sec/veh	Observations ($n_{o,i}$)	$h_{T,i} * n_{o,i}$ sec	Observations (n)	$h_{T,i} * n_{o,i}$ sec
1	3.63	90	—	180	—
2	2.43	83	—	148	—
3	2.18	75	—	116	—
4	2.05	68	—	85	—
5	1.97	61	120.36	53	104.58
6	1.93	54	104.36	21	40.59
7	1.91	46	87.88	0	0
8	1.90	39	74.03	0	0
9	1.89	32	60.52	0	0
10	1.89	24	45.30	0	0
11	1.89	17	32.06	0	0
12	1.88	10	18.85	0	0
13	1.88	2	3.77	0	0
14	1.88	0	0	0	0
15	H = 1.88	0	0	0	0
Sum:		285	547.13	74	145.17
Avg. Headway ($H_s = \Sigma(h_{T,i} * n_{o,i}) / \Sigma n$):		1.92 sec (1,875 vphgpl)		1.96 sec (1,835 vphgpl)	

C.1.2 Saturation Flow Rate Model for Through Movements

Factors Affecting Discharge Headway. In addition to the biases described in the preceding section, the differences in headway and lost-time among the interchange configurations shown in Table C-1 can be partly explained by differences among the individual study sites. A review of the literature on the topic of through movement headways suggests that several site-specific factors exist that can have an effect on the discharge process. For example, Bonneson (2), in a previous study of headways at intersections and single-point urban interchanges for NCHRP Project 3-40 (3), found that the number of vehicles served per cycle had an effect on the minimum discharge headway. Specifically, he found that the headways observed for each queue position were lower when there were more vehicles queued behind that position. He called this effect that of “traffic pressure.” In this context, traffic pressure is believed to result from the presence of aggressive drivers (e.g., commuters) that are anxious to minimize their travel time in otherwise high-volume conditions. As these drivers are typically traveling during the morning and evening peak traffic periods, they are

typically found to be concentrated in the large queues associated with these periods. It should be noted that Stokes, Messer, and Stover (4) found a similar effect of traffic queues on headways; they termed this effect "headway compression."

Bonneson (2) recommended the following equation for predicting the minimum discharge headway of a single-point urban interchange through movement as a function of traffic pressure:

$$H_{th} = 1.57 + \frac{7.70}{u_s} - 0.0086 v_l' \quad (C-4)$$

where:

H_{th} = through movement minimum discharge headway, sec/veh;

u_s = speed at saturation flow, m/s; and

v_l' = demand flow rate per lane (i.e., traffic pressure), vpcpl.

The speed in Equation C-4 represents the maximum speed drivers tend to reach as they discharge from a traffic queue. In theory, it represents the speed associated with a traffic stream flowing at its saturation flow rate. This speed was found to vary between 12 and 15 m/s in the sites studied by Bonneson (2) (it was denoted by the variable " V_{max} " in Bonneson's work). One reason offered for this variation was the proximity of some sites to adjacent intersections. Specifically, Bonneson noted that lower speeds were associated with those sites where the distance to the downstream intersection (and its associated queue) was relatively short. This suggests that discharge headways may be lower because of lower discharge speeds that result from the impending downstream stop faced by the discharging drivers.

The HCM (1) describes many additional factors that can affect discharge headway. These factors include: lane width, vehicle classification, local bus frequency, parking activity, approach grade, and area type. To avoid confounding the effect of these factors with those specifically being considered in this study (e.g., distance to back of queue), several steps were taken to avoid or remove the aforementioned factors from the data collected for this project. Specifically, the study sites all had lane widths of about 3.6 meters, approach grades of less than ± 2.5 percent, no local busses, and no parking activity. In addition, all heavy vehicles (i.e., vehicles with more than two axles) and all queued vehicles that followed heavy vehicles were removed from the data base.

Analysis Approach. The analysis of the through movement headways focussed on an examination of the headways for passenger cars in the fifth and higher queue positions at twelve study sites. Several precautions were taken to eliminate the effects of bias and to account for factors that might confound the analysis. Specifically, analysis of variance (ANOVA) techniques were used to control for differences in sample size (i.e., unbalanced data) and to account for extraneous differences among otherwise similar sites (e.g., bias by queue position). The ANOVA was implemented with the Statistical Analysis System's (SAS) (5) general linear model (GLM) because of its ability to provide statistics corrected for unbalanced data. The distribution of the residual errors was checked graphically to verify: (1) that they did not deviate significantly from a normal

distribution, and (2) that their variance was essentially constant over the range of the independent variables. All significance tests were conducted at a 95 percent confidence level (i.e., $\alpha = 0.05$).

The effects of bias in the examination of factors affecting headway were eliminated by including queue position as a "blocking factor" in the ANOVA model and by "nesting" it within each potentially influential factor (e.g., traffic pressure or phase duration) being considered. By blocking and nesting on queue position, all of the ANOVA comparisons and parameter estimates are made on a queue-position-by-queue-position basis, thereby eliminating any bias by differing sample sizes or queue positions. In modeling terms, blocking gives each queue position its own intercept whereas nesting allows the predicted headway to vary independently with each influential factor. If the factor is continuous, this latter effect is equivalent to allowing each factor to have a unique slope for each queue position. Using these techniques, the effects of potentially influential factors were examined for each queue position on an individual basis with the effect of bias eliminated.

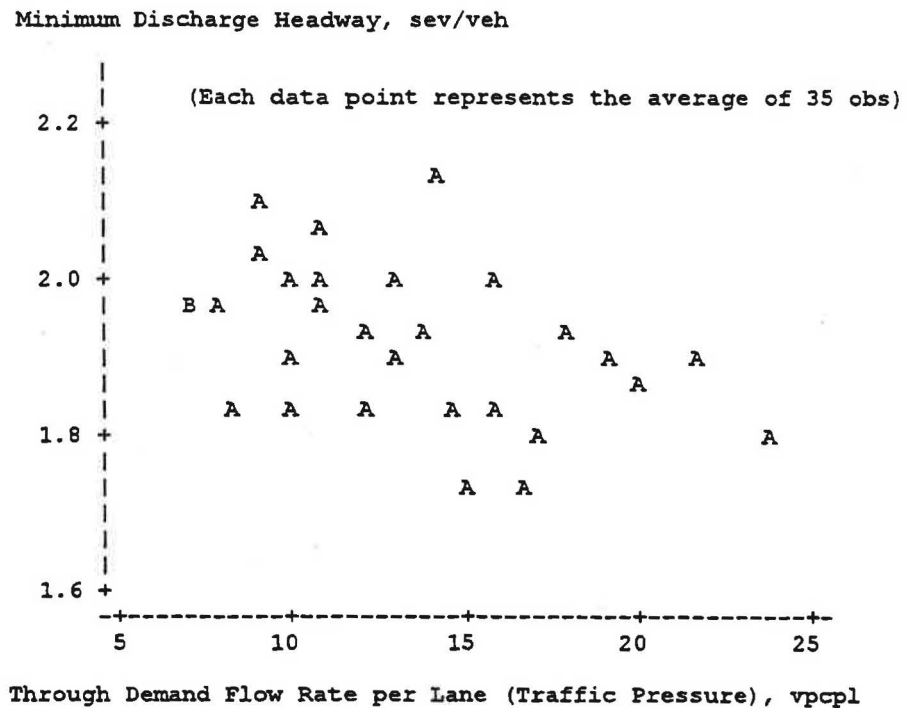
If the effect of a factor was found to be the same for each queue position (i.e., it had the same slope), then the nesting technique was eliminated and a common slope fitted to all positions. Similarly, if any queue positions were found to have the same coefficient (i.e., the same intercept), then they were combined, or pooled, to determine a single representative intercept. These types of combinations are desirable because they yield better estimates of the model parameter coefficients.

Once the influential factors were identified from the ANOVA, regression techniques were used to fit the data (via these factors) to the proposed model. Linear or nonlinear regression techniques were used to quantify the calibration parameters, depending on the model formulation. The linear regression was implemented with the SAS (5) regression model (REG); the nonlinear regression was implemented with the nonlinear model (NLIN).

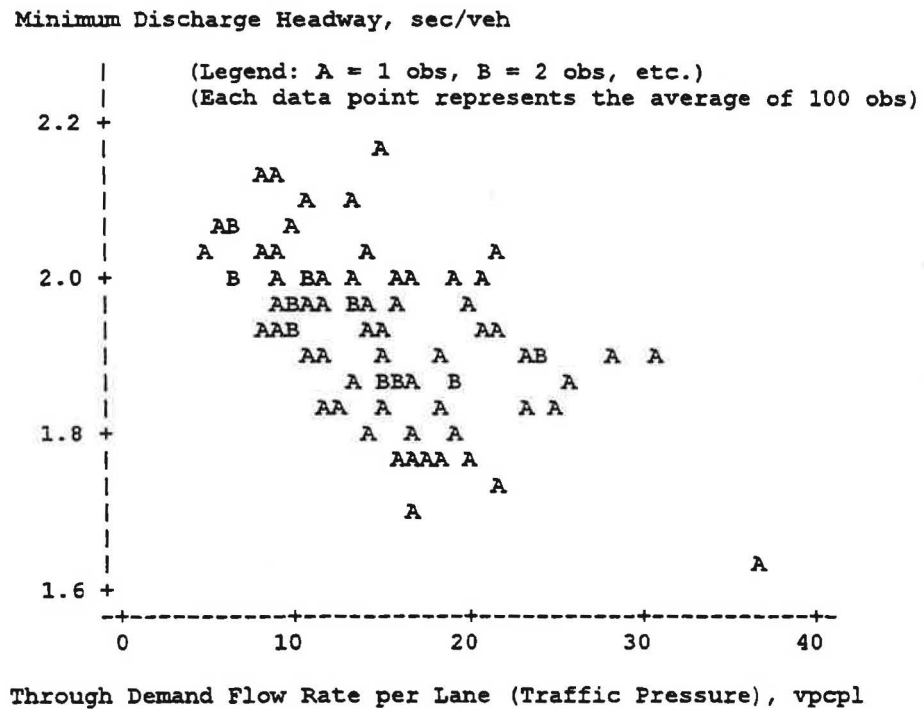
Effect of Traffic Pressure. The ANOVA analysis indicated that two factors had a significant influence on the headways at each queue position. These factors include: traffic pressure and distance to the back of the downstream queue at the start of the upstream/subject phase. The first factor is described in this section, the second factor is described in the next section.

The average headway recorded for each queue position at each site was found to decrease slightly as the number of vehicles served per cycle increased. This trend is consistent with the findings of Bonneson (2) and Stokes *et al* (4). The effect of traffic pressure on discharge headway is shown in Figure C-3.

Figure C-3 shows the effect of traffic pressure as it exists in one queue position at all sites and at all sites and queue positions combined. The analysis indicated that there were no statistically significant differences in the effect of traffic pressure among sites and queue positions. Comparison of Figures C-3a and C-3b support this finding.



a) Traffic pressure effect for queue position seven at all study sites.



b) Traffic pressure effect for all queue positions at all study sites.

Figure C-3. Effect of traffic pressure on through movement minimum discharge headway.

The data shown in Figure C-3 represent the average of several observations. The number of individual observations was so large that, when they were plotted, they tended to become a black mass of ink which obscured the examination of trend-wise effects. To overcome this problem, the data were first sorted by lane volume, segregated into contiguous groups of 35 (or 100) observations, and then used to compute average headways and lane volumes for each group.

Effect of Distance to the Back of Downstream Queue. As mentioned previously, the ANOVA indicated that the distance to the back of downstream queue had a significant effect on the queue discharge headways. This distance is measured from the subject movement stop line to the "effective" back of queue at the start of the subject (or upstream) phase. The effective back of queue represents the location of the back of queue if all vehicles on the downstream street segment (moving or stopped) at the start of the subject phase were joined into a stopped queue. If there are no moving vehicles at the start of the phase, then the effective and actual distance to queue are the same. If there are no vehicles on the downstream segment at the start of the phase, then the effective distance to queue would equal the distance to the through movement stop line at the downstream intersection.

Figure C-4 shows the relationship between the discharge headway of the fifth and higher queue positions and the corresponding distance to the back of downstream queue. As the headways for these queue positions are considered to be effectively at their minimum value, the data shown represents individual estimates of the minimum discharge headway H . The trend shown in this figure indicates that the minimum discharge headway decreases with increasing distance to queue.

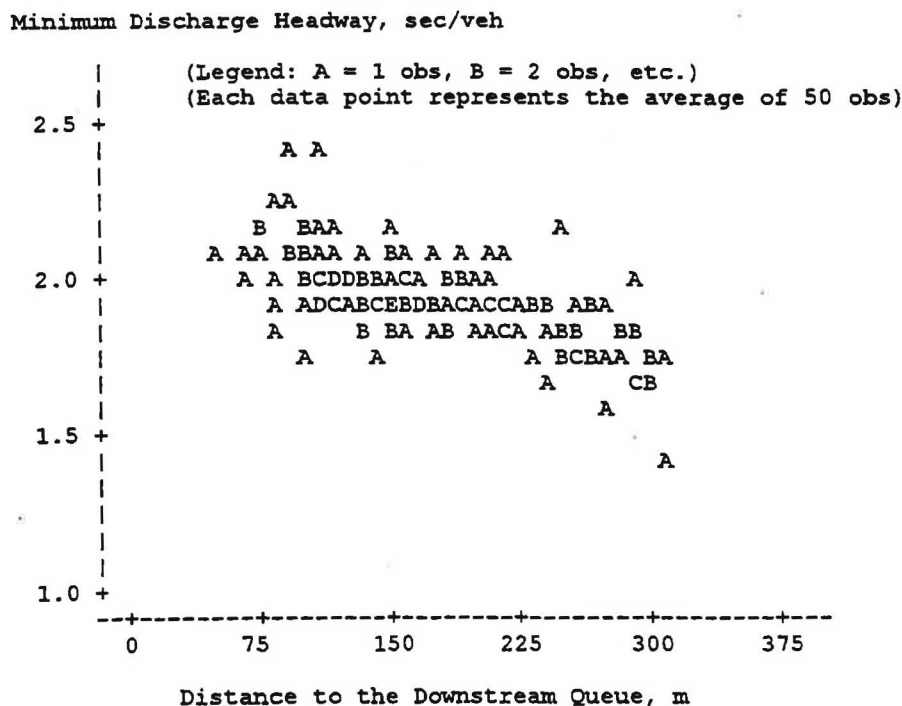


Figure C-4. Effect of distance-to-queue on through movement minimum discharge headway.

As there are more than 7,700 headways in the data base, plots of individual data points were problematic. Therefore, the data were sorted by distance-to-queue, segregated into contiguous groups of 50 observations, and then average headways were computed for each distance group. The data shown in Figure C-4 represent these average headways.

The effect of distance-to-queue was found to vary depending on whether the downstream traffic queue spilled back into the upstream intersection. Spillback conditions are characterized by the backward propagation of a downstream queue into the upstream intersection such that subject (or upstream) intersection movement is effectively blocked from discharging during some or all of the signal phase. In general, the headways of vehicles able to discharge prior to spillback were larger than the headways of similar vehicles measured during phases in which spillback did not occur.

Model of Minimum Discharge Headway. Several alternative model forms were considered in an attempt to find the best fit to the data. A simple linear model was considered, however, this form did not provide logical boundary values. For example, it is logical that the headway would tend to very large values at short distances-to-queue and would converge to a nominal minimum discharge headway for large distances-to-queue. Based on this rationalization, the following empirical model form was developed and fit to the data:

$$H_{th} = b_0 \left(1 + \frac{b_1 (1 - I_s) + b_2 I_s}{D} \right) (1 - b_3 v_l') \quad (C-5)$$

where:

H_{th} = through movement minimum discharge headway, sec/veh;

D = effective distance to the back of downstream queue (or stop line if no queue) at the start of the subject (or upstream) phase, m;

I_s = indicator variable (1.0 if spillback occurs during phase, 0.0 otherwise);

v_l' = demand flow rate per lane (i.e., traffic pressure), vpcpl; and

b_0, b_1, b_2, b_3 = calibration coefficients.

Model Calibration. The field data were used to calibrate the minimum discharge headway model, as shown in Equation C-5. The details of the model calibration analysis and statistics describing the model's predictive performance are provided in Table C-3.

As the statistics in Table C-3 indicate, the calibrated model explains only about four percent of the variability in the headway data. The remaining variability is primarily due to the random (or unexplainable) variability inherent in headway data. Some of the variability is also due to differences among the traffic lanes and sites studied. Nevertheless, the statistics in Table C-3 indicate that there is a statistically significant relationship between minimum discharge headway, traffic pressure, and distance to the back of queue. The root mean square error and number of observations can be used to estimate the minimum standard deviation (or precision) of the predicted average minimum discharge headway as ± 0.006 sec/veh.

Table C-3. Calibrated through movement minimum discharge headway model

Model Statistics		Value		
R^2 :		0.04		
Root Mean Square Error:		0.56 sec/veh		
Observations:		7,704		
Range of Model Variables				
Variable	Variable Name	Units	Minimum	Maximum
H_{th}	Through movement min. discharge headway	sec/veh	0.61	6.8
v_l'	Demand flow rate per lane (traffic pressure)	vpcpl	5	37
D	Distance to back of downstream queue	meters	35	315
Calibrated Coefficient Values				
Coeff.	Coefficient Definition	Value	Std. Dev.	t-statistic
b_0	Min. discharge headway for ideal D and v_l	1.94	0.028	69.3
b_1	Effect of distance to queue (no spillback)	8.13	1.45	5.6
b_2	Effect of distance to queue (with spillback)	21.8	1.56	14.0
b_3	Effect of traffic pressure	0.00453	0.0005	9.1

The data used for model calibration were screened to identify the effects of queue spillback. This screening was accomplished by using the videotape record of traffic events obtained during the field studies. During the screening process, each headway observation included in the database was visually verified as to whether it was affected by spillback. This approach resulted in there being two categories of observations in the database. One category included those headways that occurred during phases not having spillback. The other category included those headways that occurred during phases with spillback but prior to the first occurrence of spillback. Headways occurring during phases with spillback but subsequent to the first occurrence of spillback were excluded from the database.

As the coefficient values in Table C-3 indicate, the magnitude of the effect of distance-to-queue is dependent on whether queue spillback occurred during the phase. Phases without spillback had a smaller regression coefficient indicating less sensitivity to distance. In general, the coefficients predict a larger minimum headway for those queues discharging prior to the occurrence of spillback than for those that discharge without spillback ever occurring.

In addition to its intended purpose, the distance-to-queue variable D addresses the effect of intersection spacing (including differences among interchange types). In general, it suggests that discharge efficiency will likely be lower for the inbound movements at closely-spaced intersections (e.g., interchanges with relatively short distances between ramp terminals) than it would be at otherwise similar intersections but with larger spacings.

The quality of fit of the calibrated model to the headway data is shown in Figure C-5. The plotted data points represent averages of 50 observations each. As discussed previously, this

averaging was performed to overcome the presentation problems incurred when plotting hundreds of data points. It should be noted that the individual headway observations were used for the model calibration; averaged data were used for plotting purposes only.

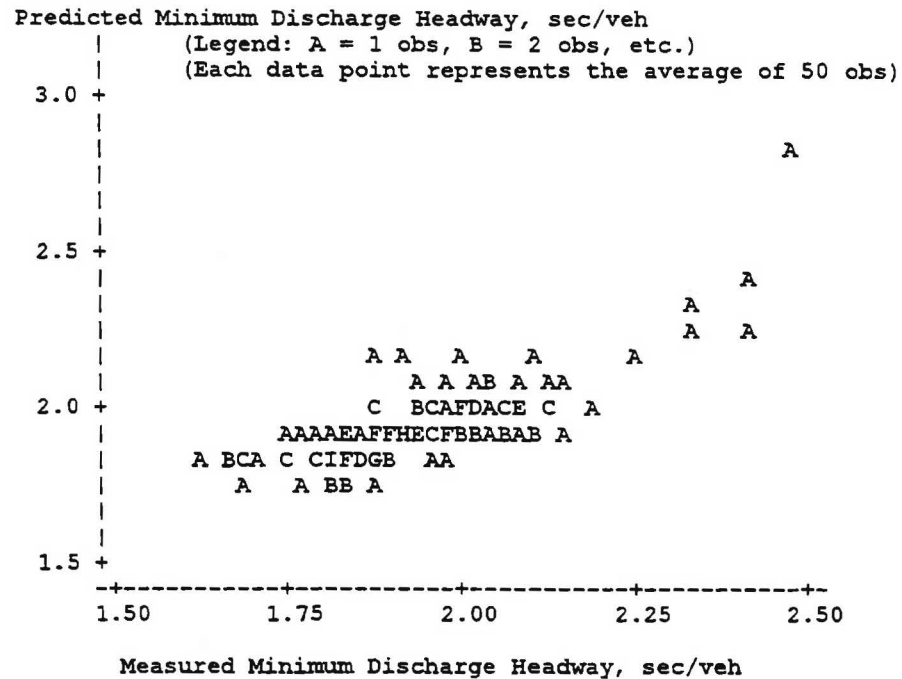


Figure C-5. Comparison of predicted and measured through movement minimum discharge headways.

Several additional effects were also evaluated during the model calibration process. Specifically, the effect of junction type, phase duration, and downstream signal indication were also evaluated. This latter factor was considered because it was reasoned that drivers might discharge at a more efficient rate if the downstream signal indication was green (as opposed to red), particularly if there was no downstream queue. Based on this additional analysis, it was concluded that these factors did not significantly affect discharge headway after the effects of distance to queue and traffic pressure were removed.

Interpretation of Model Statistics. Three statistics are provided in Table C-3 to indicate the quality of fit of the calibrated model. First, the "t-statistic" is provided for each independent variable to test the hypothesis that its regression coefficient equals to zero. When the t-statistic exceeds 1.96, the hypothesis is rejected and it can be concluded that the corresponding variable has a significant effect on the dependent variable. In this situation, there is a 5 percent (or less) chance of this conclusion being in error.

In some situations, the effect of a variable can be misrepresented when there is a large amount of data. This problem was avoided in this research by using an analysis of variance

(ANOVA) approach. An ANOVA is much more robust than a regression analysis at detecting the true effect of an independent variable. In addition, a graphical examination was conducted for each independent variable. This examination provides a practical means of confirming any relationship that may exist between the dependent and independent variables; it can also indicate the trend between these variables (e.g., linear, log, etc.). Figures C-3 and C-4 are examples of this type graphical examination. The influence of the independent variables is clearly indicated in these figures; the t-statistic provided for each variable in Table C-3 is further confirmation of the significance of the observed trend.

The second measure of quality of fit is the root mean square error. This statistic represents the standard deviation of the dependent variable. Presumably, the error represented by this statistic is from random sources; however, there could also be some variation due to systematic effects. Knowledge of typical values of the root mean square error for the dependent variable can be a useful gage to assess whether additional systematic error exists in the data. For example, the standard deviation of vehicle headways is rarely reported in the literature to be less than 0.45 sec. Therefore, as the root mean square error of 0.56 reported in Table C-3 exceeds 0.45, it is possible that there is some additional systematic error in the data that additional model variables could explain.

The third measure of quality of fit is the R^2 statistic. This statistic represents the portion of the variability explained by the model relative to the total variability in the data. As such, it can range in value from 0.0 to 1.0. In general, larger values of R^2 indicate a good fit; however, it must be remembered that the value (or range of values) used to denote a "good" fit is dependent on the amount of random variability in the data. For example, the only way that an R^2 of 1.0 can be achieved is when all of the variability in the data is due systematic sources (i.e., there is no random error) and the model properly includes an independent variable for each systematic effect. To illustrate this relationship, the following relationship between the actual and "best-possible" R^2 was derived using mathematical relationships between R^2 and the partitioned variance structure:

$$R_{best}^2 = 1 - (1 - R_{act}^2) \frac{\sigma_r^2}{\sigma_{act}^2} \quad (C-6)$$

where:

R_{best}^2 = largest possible R^2 value obtainable (i.e., all systematic error explained);

R_{act}^2 = R^2 value obtained from a regression analysis;

σ_r^2 = variance in data due to random sources; and

σ_{act}^2 = variance obtained from a regression analysis (i.e., = root mean square error²).

To illustrate the implications of this equation, consider the "actual" R^2 and σ^2 values provided in Table C-3. If it can be assumed that the variance due to random sources is about 0.20 (= 0.45²), then the largest R^2 value possible is 0.38 (= 1 - (1 - 0.04) 0.45² / 0.56²); it is *not* the 1.0 that is traditionally assumed. Data sets with larger amounts of random variability will yield even lower maximum values for R^2 .

Other factors that have a significant effect on the value of R^2 are the effect of the independent variable (i.e. the absolute value of the corresponding variable coefficient) and its variability. In general, the R^2 will be limited to relatively low values when the independent variable's coefficient or its variance are small (and there is some variability due to random sources). The following equation can be derived for computing R^2 using mathematical relationships between R^2 and the partitioned variance structure:

$$R^2_{act} = \frac{b_1^2 \sigma_x^2}{\sigma_r^2 + b_1^2 \sigma_x^2} \quad (C-7)$$

where:

σ_x^2 = variance in the independent variable; and

b_1 = regression coefficient for the independent variable.

To illustrate the implications of this equation, consider the equation $H = b_0 + b_1/D$ (this equation is similar in form to Equation C-5). If the variance of the independent variable ($1/D$) is 0.000036, the value of b_1 is 15, and the variance due to random sources is 0.20, then the R^2 will be about 0.04. These variable values were chosen to be consistent with the data used in calibrating Equation C-5 and serve to illustrate the effect of a low variance in the independent variable on the R^2 . Based on this computation, the R^2 value of 0.04 cited in Table C-3 is probably about as large a value as could be expected given the relatively small variance of the independent variable.

It should be noted that Equation C-7 is generally insensitive to the *form* of the model. To illustrate this attribute, consider the equation $H = b_0 + b_1 D$. Using the same data as for the preceding example, values of -0.002 and 2.025 are obtained for b_1 and σ_x^2 , respectively. As before, the resulting R^2 is about 0.04. Hence, the R^2 obtained from either model is essentially the same. Moreover, this example illustrates the effect of a relatively "small" (but significant) regression coefficient on R^2 ; that is, small regression coefficients inherently yield small R^2 values. Thus, in a comparison of two regression models with identical root mean square errors and t-statistics, Equation C-7 guarantees that the model with the larger coefficient (i.e., slope) will have the larger R^2 . Yet, the similarity in root mean square errors and t-statistics suggest that each model is equally good.

In general, the models developed in this and subsequent sections deal with traffic data representing the reactions of individual drivers to the roadway geometry, the traffic signal, and other vehicles. As a result, the data have a relatively large random error component making it impossible to obtain an R^2 near 1.0, regardless of the model form. More importantly, the magnitude of the systematic effect (as represented by the independent variable coefficient) is generally small for this type of traffic data which limits the R^2 to low values. In conclusion, the R^2 statistic is a useful measure of model quality of fit *when used for relative comparisons among similar models and data sets*. It is impossible to universally define one single R^2 value as the "minimum acceptable value."

Sensitivity Analysis. The calibrated discharge headway model was used to examine the effect of distance-to-queue, spillback, and traffic pressure on the queue discharge rate. These effects are shown in Figures C-6 and C-7 for the discharge headway and saturation flow rate, respectively.

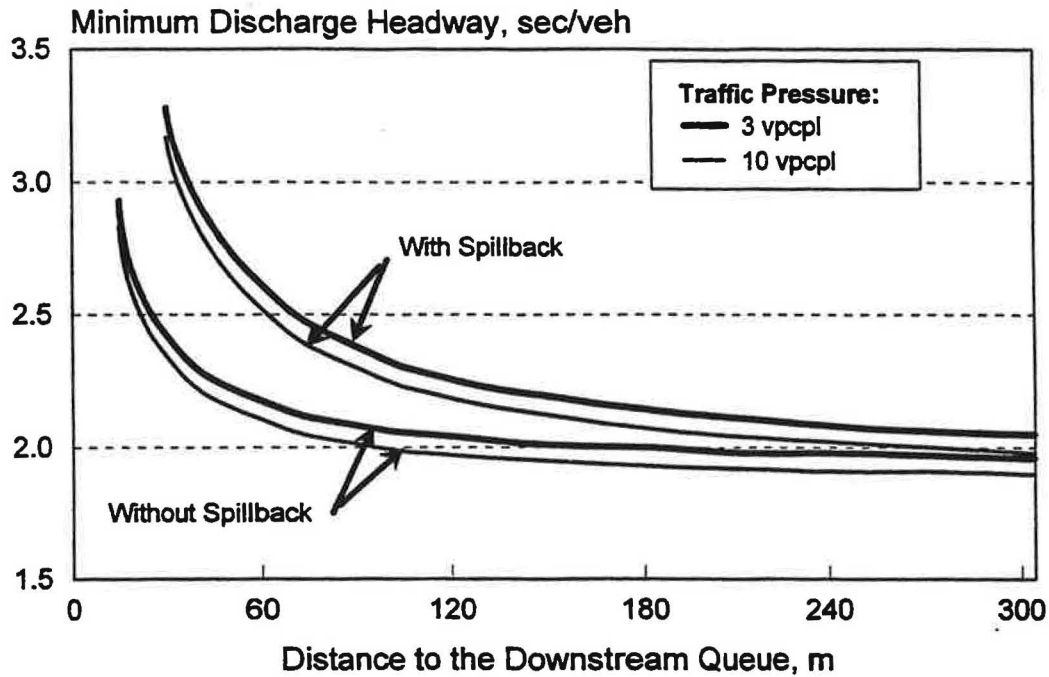


Figure C-6. Effect of distance-to-queue, spillback occurrence, and traffic pressure on through movement minimum discharge headway.

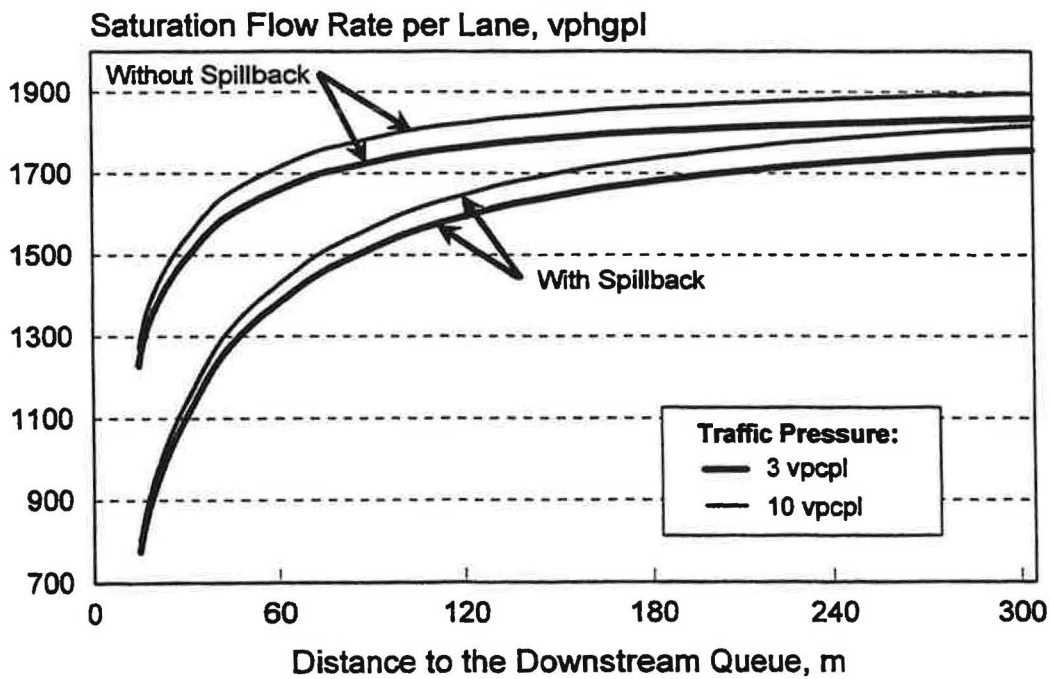


Figure C-7. Effect of distance-to-queue, spillback occurrence, and traffic pressure on through movement saturation flow rate.

Observation of Figure C-6 indicates that the calibrated model has a trend of decreasing headway with increasing distance-to-queue which is consistent with the trend shown in Figure C-4. Phases that incurred spillback tended to be associated with larger headways and lower saturation flow rates. Specifically, discharge headways are higher for those vehicles that are able to discharge prior to the first occurrence of spillback. Traffic pressure decreases the minimum discharge headway; although, the effect of traffic pressure is small relative to that of distance-to-queue or spillback.

The calibrated model predicts a minimum headway of 1.94 sec/veh (= 1,856 pcphgpl) when the distance-to-queue is infinitely long and there is no traffic pressure effect. Inclusion of a nominal traffic pressure of 5.0 vpcpl yields a minimum headway of 1.90 sec/veh (= 1,900 pcphgpl); this value is consistent with the ideal saturation flow rate recommended by the HCM (1) for signalized intersections. The calibrated model also logically predicts infinitely long headways when the distance-to-queue is zero.

Saturation Flow Rate Model. The calibrated through movement minimum discharge headway model was converted into an equivalent saturation flow rate model. The form of this model was patterned after that used in Chapter 9 of the HCM (1). Specifically, the saturation flow rate for a particular location is estimated as the product of the ideal saturation flow rate and the various site-specific adjustment factors. In this context, the adjustment factors found in this research relate to the effect of distance-to-queue, spillback occurrence, and traffic pressure. The basic form of the model is:

$$s_i = s_0 \times f_D \times f_v \quad (C-8)$$

where:

- s_i = saturation flow rate per lane under prevailing conditions, vphgpl;
- s_0 = saturation flow rate per lane under ideal conditions, pcphgpl;
- f_D = adjustment factor for distance to downstream queue at green onset; and
- f_v = adjustment factor for volume level (i.e., traffic pressure).

The ideal saturation flow rate represents the saturation flow rate when not affected by any external environmental factors (i.e., grade), atypical vehicles (i.e., trucks), and constrained geometrics (e.g., less than 3.6-meter lane widths, curved travel path). In this regard, the saturation flow rate would be equal to the ideal rate when all factor effects are optimum for efficient traffic flow and the corresponding adjustment factors are equal to 1.0. Based on this definition, it was determined that an infinite distance-to-queue under non-spillback conditions and a traffic pressure of 15.0 vpcpl were representative of ideal conditions. Using this definition of ideal conditions and associated parametric values, the resulting ideal saturation flow rate and adjustment factors were algebraically derived from Equation C-5 as:

$$\begin{aligned} s_0 &= \frac{3,600}{b_0 (1 - b_3 15.0)} \\ &= 1,990 \text{ pcphgpl} \end{aligned} \quad (C-9)$$

$$f_D = \begin{cases} \frac{1}{1 + \frac{8.13}{D}} & : \text{ no spillback} \\ \frac{1}{1 + \frac{21.8}{D}} & : \text{ with spillback} \end{cases} \quad (\text{C-10})$$

$$\begin{aligned} f_v &= \frac{1}{(1 - b_3 v_l') / (1 - b_3 15.0)} \\ &= \frac{1}{1.07 - 0.00486 v_l'} \end{aligned} \quad (\text{C-11})$$

where:

D = effective distance to the back of downstream queue (or stop line if no queue) at the start of the subject (or upstream) phase, m;

v_l' = demand flow rate per lane (i.e., traffic pressure), vpcpl; and

b_0, b_3 = calibration coefficients from Table C-3.

The precision of the ideal saturation flow rate predicted by Equation C-9 is estimated as ± 12 pcphgpl, based on the root mean square error and number of observations shown in Table C-3. As the resulting range of possible true ideal values includes 2,000 pcphgpl, this value is recommended as the ideal saturation flow rate for through movements.

The relationship between distance-to-queue and the corresponding saturation flow adjustment factor f_D is shown in Figure C-8. The trends shown are similar to those noted for the saturation flow rates shown in Figure C-7. Specifically, the adjustment factor (and saturation flow rate) increase as the distance to the back of queue becomes longer. Phases that incur spillback have a smaller factor value, for the same distance to queue, than phases that do not incur spillback.

C.1.3 Saturation Flow Rate Model for Left-Turn Movements

With one exception, the left-turn movements included in this study represent left-turns at interchange ramp terminals. The one exception was a left-turn movement at an adjacent signalized intersection. Of the two types of left-turn movements at ramp terminals (i.e., off-ramp and arterial), the majority of the data were collected for the off-ramp left-turn movement. Nevertheless, it is believed that the factors identified in this section are sufficiently general that they are applicable to off-ramp and arterial left-turn movements at interchanges and to left-turn movements at adjacent intersections.

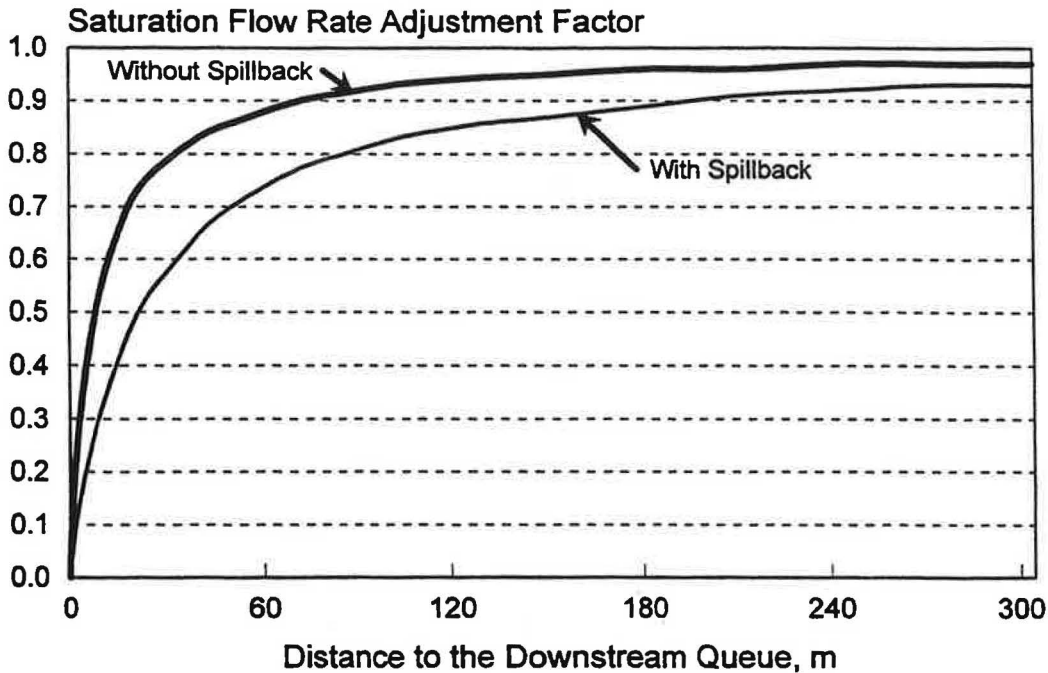


Figure C-8. Effect of distance-to-queue and spillback occurrence on the through movement saturation flow rate adjustment factor.

The left-turn movements studied rarely, if ever, experienced queue spillback during the study periods due to the nature of the signal phase coordination between the two interchange ramp terminals. Hence, in contrast to the through movements studied, the variability in left-turn headways among sites cannot be explained by differences in the distance to the downstream queue. This restriction is a characteristic of the twelve sites studied; certainly, left-turn movements can be affected by downstream queueing conditions at other sites. In fact, it is likely that the effect will be very similar to that found for the through movements.

Factors Affecting Discharge Headway. A review of the literature on the topic of left-turn headways suggests that several site-specific factors exist that can have an effect on the left-turn discharge process. For example, Kimber *et al* (6) measured saturation flows on curves with radii ranging from 6 to 35 meters and developed an equation for predicting saturation flow rate as a function of turn radius. An equivalent relationship, as it relates to minimum discharge headway, is:

$$H_u = 1.73 + \frac{2.60}{R} \quad (C-12)$$

where:

H_u = left-turn movement minimum discharge headway, sec/veh; and

R = radius of curvature of the left-turn travel path (at center of path), m.

In a previous study of single-point urban interchanges conducted by Messer *et al* (3), Bonneson (2) found an effect of radius on headway consistent with that found by Kimber (6). The

range of radii included in this study was 18 to 84 meters. Bonneson also found that the number of vehicles served per cycle had an effect on left-turn headway. This latter effect was referred to as that of "traffic pressure" (as discussed in a preceding section). Bonneson recommended the following equation for predicting the minimum discharge headway of a left-turn movement as a function of radius and traffic pressure:

$$H_{\mu} = 1.58 + \frac{0.830}{R^{0.245}} - 0.0121 v_l' \quad (C-13)$$

where:

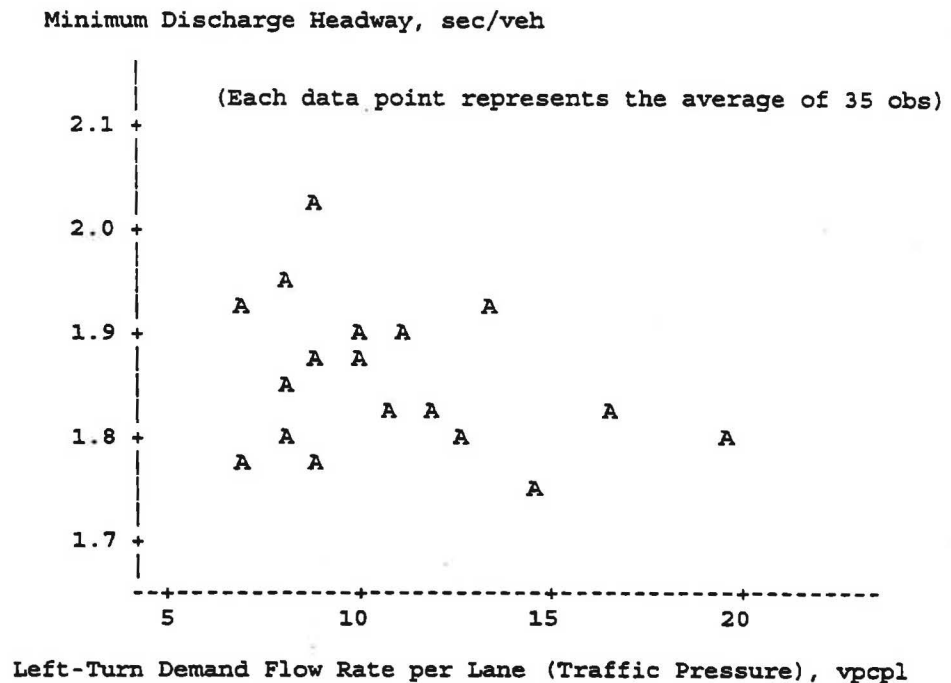
v_l' = demand flow rate per lane (i.e., traffic pressure), vpcpl.

The HCM (1) describes many additional factors that can affect discharge headway. These factors include: lane width, vehicle classification, local bus frequency, parking activity, and approach grade. To avoid confounding the effect of these factors with those specifically being considered in this study (e.g., turn radius), several steps were taken to avoid or remove the aforementioned factors from the data collected for this project. Specifically, the study sites all had lane widths of about 3.6 meters, approach grades of less than ± 2.5 percent, no local busses, and no parking activity. In addition, all heavy vehicles (i.e., vehicles with more than two axles) and all queued vehicles that followed heavy vehicles were removed from the data base.

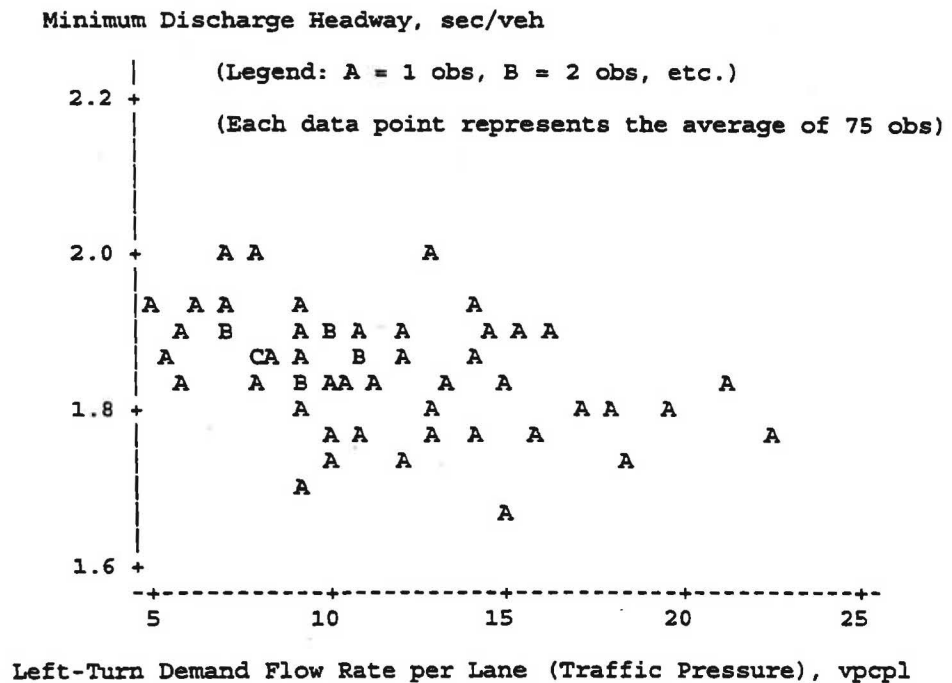
Analysis Approach. The analysis of the left-turn headways focused on an examination of the headways for passenger cars in the fifth and higher queue positions at eleven study sites. Analysis of Variance (ANOVA) techniques were used to control for differences in sample size and to account for extraneous differences among otherwise similar sites. All significance tests were conducted at a 95 percent confidence level (i.e., $\alpha = 0.05$). Once the influential factors were identified from the ANOVA, both linear and non-linear regression techniques were used to calibrate the data (via these factors) to the proposed model. More details of the modeling approach are provided in a preceding section describing the through movement saturation flow model.

Traffic Pressure Effect. The ANOVA analysis indicated that three factors had a significant influence on the headways at each queue position. These factors include: traffic pressure, signal timing, and turn radius. The first factor is described in this section, the latter two are described in the next two sections. The combined effect of distance to queue and spillback were not considered in the left-turn movement analysis. This omission was a consequence of the signal coordination used at the study sites. This coordination was generally good for the left-turn movements such that they did not experience spillback nor were they faced with lengthy downstream queues.

The average headway recorded for each queue position at each site was found to decrease slightly as the number of vehicles served per cycle increased. This trend is consistent with the findings of Bonneson (2) and Stokes *et al* (4). The effect of traffic pressure on discharge headway is shown in Figure C-9.



a) Traffic pressure effect for queue position seven at all study sites.



b) Traffic pressure effect for all queue positions at all study sites.

Figure C-9. Effect of traffic pressure on left-turn movement minimum discharge headway.

Figure C-9 shows the effect of traffic pressure as it exists in one queue position at all sites and at all sites and queue positions combined. The analysis indicated that there were no statistically significant differences in the effect of traffic pressure among sites and queue positions. Comparison of Figures C-9a and C-9b support this finding.

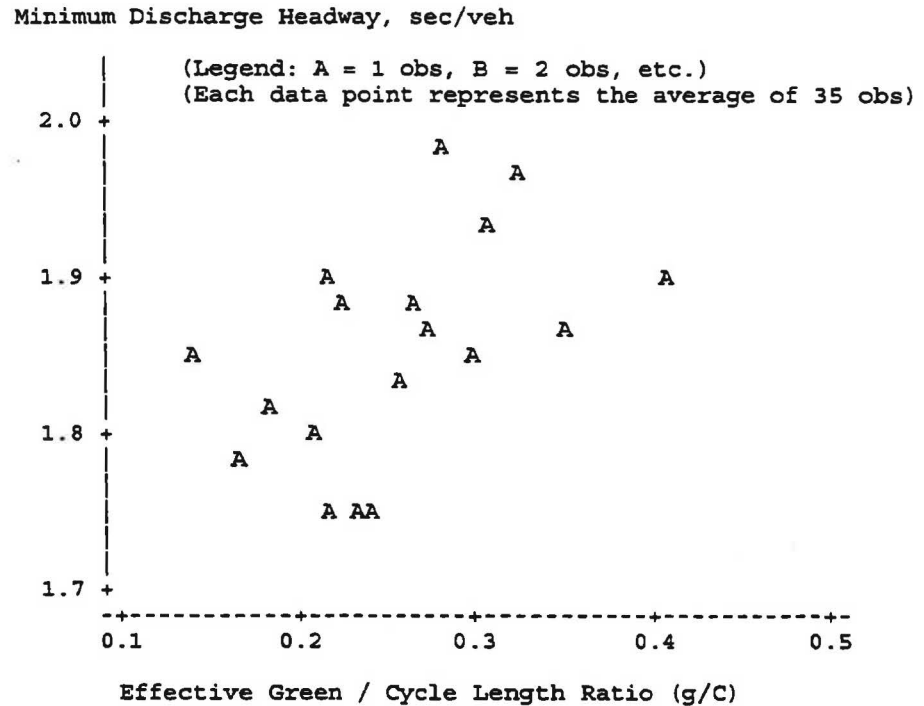
The data shown in Figure C-9 represent the average of several observations. The number of individual observations was so large that, when they were plotted, they tended to become a black mass of ink which obscured the examination of trend-wise effects. To overcome this problem, the data were first sorted by lane volume, segregated into contiguous groups of 35 (or 75) observations, and then used to compute average headways and lane volumes for each group.

Signal Timing Effect. The ANOVA tests also indicated that the left-turn movements were affected by the duration of the phase and the cycle length. The effects of these two factors were dissimilar in the sense that drivers adopted shorter discharge headways during *shorter* phase durations or *longer* cycle lengths or both. A similar effect of phase duration on headway has been noted by Stokes *et al* (4). The implication of these findings is that drivers will adopt shorter headways to avoid the additional delay associated with having to wait for the next phase, particularly when the expected delay is large (due to a large cycle length). As delay is theoretically related to the effective-green-to-cycle-length ratio g/C , the two factors were combined into this ratio for further examination. The results of this examination are shown in Figure C-10.

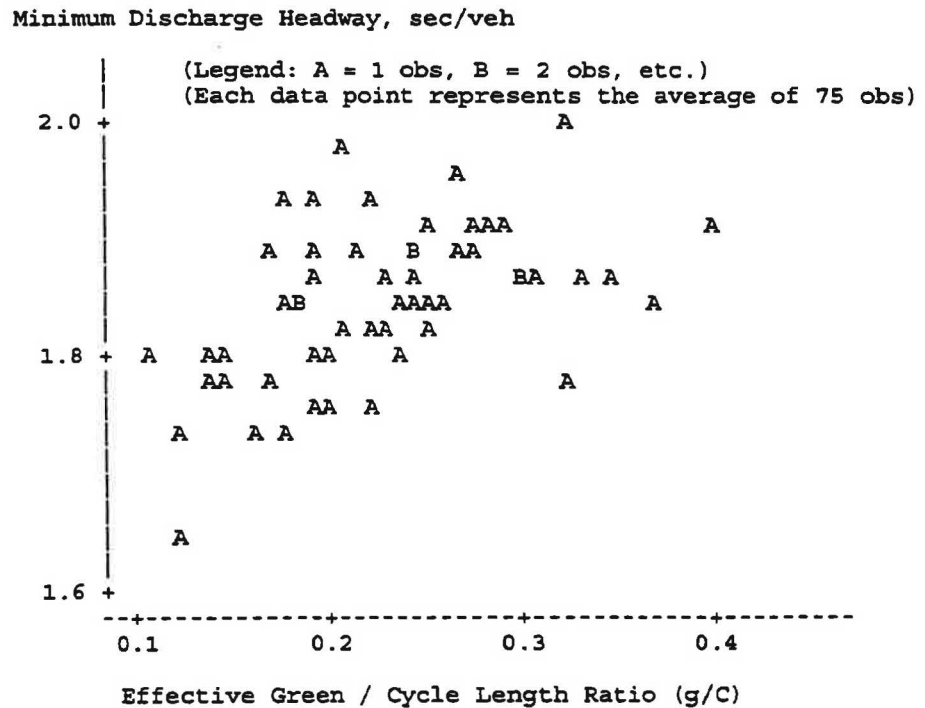
Figure C-10 shows the effect of signal timing as it exists in one queue position at all sites and at all sites and queue positions combined. This analysis indicated that there was no statistically significant difference in the effect of signal timing among sites and queue positions. Comparison of Figures C-10a and C-10b support this finding. Closer examination of Figure C-10b indicates that the effect of g/C ratio is not constant over the range of g/C ratios observed. Specifically, it appears that the effect of g/C ratio is limited to values less than about 0.27. Beyond this value, the data suggest that the effect of g/C ratio is negligible.

As with Figure C-9, the data shown in Figure C-10 represent the average of several observations. This averaging was performed to facilitate the examination of trends in the data rather than its variability.

Unlike the left-turn movements, the g/C ratio was not found to have significant effect on the through movement discharge rate. This finding is likely due to the large g/C ratios experienced by the through movements, relative to the left-turn movements. The examination of the left-turn movements found that the effect of g/C ratio was limited to situations where the ratio was relatively small (i.e., less than 0.27). As the average g/C ratio for the through movements was about twice that of the left-turn movements, there were rarely any through movement phases that had a g/C ratio low enough to warrant driver anxiety about being delayed for an additional signal cycle. Therefore, the existence of an effect of g/C ratio on through headways may exist but it cannot be determined (as it was for left-turn headways) because there is insufficient data from sites with through movements operating at low g/C ratios.



a) Signal timing effect for queue position seven at all study sites.



b) Signal timing effect for all queue positions at all study sites.

Figure C-10. Effect of signal timing on left-turn movement minimum discharge headway.

Travel Path Radius Effect. The radius of the travel path was examined for its effect on the minimum discharge headway. To facilitate this examination, the minimum discharge headways were tabulated for each site and traffic lane studied. In addition, the minimum discharge headways of several sites studied during a previous NCHRP project (3) were included in the tabulation to increase its range of coverage. The results of this effort are summarized in Table C-4.

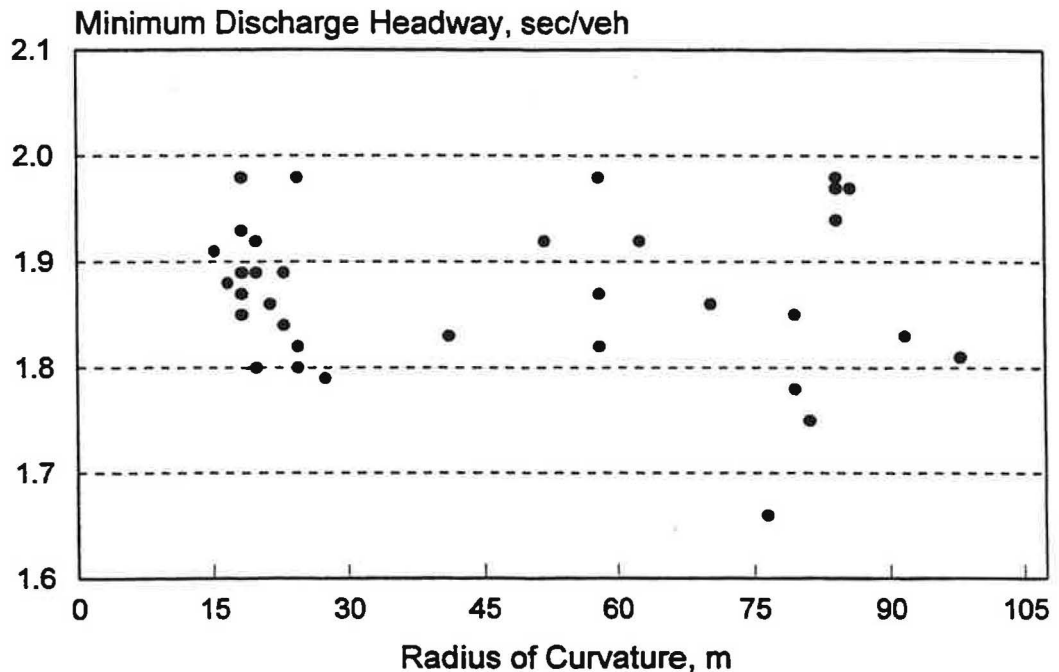
Table C-4. Left-turn movement minimum discharge headway database summary

Left-Turn Lane Type	Junction Type ¹	Location	City, State	Left-Turn Radius ⁴ (m)	Single Lane or Inside Lane of Dual Lanes ²			Outside Lane of Dual Lanes ²			g/C Ratio ³
					Obs.	H _b (s/v)	v _i ['] (vp/cpl)	Obs.	H _b (s/v)	v _i ['] (vp/cpl)	
Single Lane	AGI	SR 60 & Belcher Road	Clearwater, FL	18	166	1.89	4.7	na	na	na	0.10
		SR 60 & Belcher Road	Clearwater, FL	18	229	1.93	5.2	na	na	na	0.10
		Wellborn Rd. & Old Main Dr	College Station, TX	18	268	1.87	3.4	na	na	na	0.18
		Wellborn Rd. & Old Main Dr	College Station, TX	18	74	1.98	2.9	na	na	na	0.18
		Mathilda Ave & Moffitt Rd.	Sunnyvale, CA	41	42	1.83	6.7	na	na	na	0.20
	Parclo B	I-880 & Stevenson Blvd	Newark, CA	23	458	1.89	5.3	na	na	na	0.22
		SR 4 & Somersville Rd.	Antioch, CA	24	488	1.98	5.8	na	na	na	0.22
	SPUI	US 19 & SR 694	St. Petersburg, FL	52	410	1.92	4.6	na	na	na	0.18
		US 19 & SR 694	St. Petersburg, FL	70	1,899	1.86	9.1	na	na	na	0.20
		US 19 & SR 694	St. Petersburg, FL	85	1,086	1.97	6.7	na	na	na	0.20
		US 19 & SR 60	Clearwater, FL	58	1,451	1.87	8.4	na	na	na	0.22
		US 19 & SR 60	Clearwater, FL	58	1,316	1.82	8.3	na	na	na	0.22
	Average:				40	657	1.90	5.9	na	na	na
Dual-Lane	CDI	Metcalf Ave. & I-435	Overland Park, KS	18	76	1.85	5.8	49	1.84	3.6	0.19
		75th Street & I-35	Overland Park, KS	15	73	1.91	3.0	282	1.80	4.1	0.17
		Maple Street & I-680	Omaha, NE	20	499	1.89	5.8	192	1.82	3.6	0.22
	SPUI	US 19 & SR 60	Clearwater, FL	58	584	1.98	6.1	329	1.92	4.8	0.18
		Indian School Rd. & SR-51	Phoenix, AZ	76	367	1.66	5.6	275	1.75	5.0	0.15
		US 19 & SR 686	Largo, FL	79	2,318	1.78	13.3	452	1.97	8.5	0.23
		US 19 & SR 686	Largo, FL	79	392	1.85	5.2	284	1.94	4.9	0.13
		7th Street & I-10	Phoenix, AZ	91	51	1.83	2.7	154	1.81	3.4	0.20
		Peoria Ave. & I-17	Phoenix, AZ	23	919	1.84	13.1	538	1.79	8.9	0.32
	TUDIw/f	Arapaho Rd. & US-75	Richardson, TX	17	408	1.88	5.4	200	1.86	4.1	0.20
		Towneast Blvd. & I-635	Mesquite, TX	20	192	1.92	4.1	64	1.80	2.9	0.21
	Average:				45	534	1.85	6.4	256	1.85	4.9
Overall Average:				43	599	1.88	6.1				0.19

Notes:

- 1 - AGI - at-grade intersection; SPUI - single-point urban interchange; CDI - compressed diamond interchange; TUDI - tight urban diamond interchange (no frontage road or U-turn lanes); TUDI w/f - tight urban diamond interchange (with frontage road and U-turn lanes); Parclo B - partial cloverleaf with off-ramps beyond overpass.
- 2 - H_L = left-turn movement minimum discharge headway.
 v_L' = demand flow rate per lane (i.e., traffic pressure).
- 3 - g/C = effective green duration g of the subject left-turn phase divided by the cycle length C .
- 4 - Turn radius of the left-turn lane (inside lane of dual-lane bays), measured to the center of the travel path. Radii represent the visual "best fit" approximation of the travel path of the turn movement along a single radius curve.

A preliminary examination of the relationship between minimum discharge headway and radius was conducted using the average headways shown in Table C-4. This relationship is shown in Figure C-11. As suggested by the scatter in the data shown in this figure, the results of this analysis were somewhat inconclusive; however, there appeared to be a slight trend toward shorter headways on travel paths of larger radius.



Model of Minimum Discharge Headway. Based on a review of model forms used by other researchers for predicting the minimum discharge headway for left-turn movements and the trend analysis described in the preceding sections, the following empirical model form was developed and fit to the headway data:

$$H_{lt} = b_0 \left(1 + \frac{b_1}{R} \right) (1 - b_2 v_l') (1 + b_3 t_g) \quad (C-14)$$

with,

$$t_g = \left(\frac{g}{C} \right) I_g + g_x (1 - I_g) \quad (C-15)$$

where:

- H_{lt} = left-turn movement minimum discharge headway, sec/veh;
- R = radius of curvature of the left-turn travel path (at center of path), m;
- v_l' = demand flow rate per lane (i.e., traffic pressure), vpcpl;
- g = effective green time where platoon motion (flow) can occur, sec;
- C = cycle length, sec;
- t_g = signalization variable ($0.0 < t_g < g_x$);
- I_g = indicator variable (1.0 if $g/C < g_x$, 0.0 otherwise);
- g_x = maximum g/C ratio (larger g/C ratios have no additional effect on headway); and
- b_0, b_1, b_2, b_3 = calibration coefficients.

Model Calibration. The field data were used to calibrate the left-turn movement minimum discharge headway model, as shown in Equation C-14. The details of the model calibration analysis and statistics describing the model's predictive performance are provided in Table C-5.

As the statistics in Table C-5 indicate, the calibrated model explains only about five percent of the variability in the headway data. The remaining 95 percent of the variability is primarily due to the inherent randomness in headway data (as described in a preceding section dealing with the saturation flow rate model for through movements). Nevertheless, the statistics in Table C-5 indicate a statistically significant relationship between minimum discharge headway, radius, traffic pressure, and g/C ratio. The minimum precision of the average headway estimate is about ± 0.007 sec/veh.

The quality of fit of the calibrated model to the headway data is shown in Figure C-12. As discussed previously, the plotted data points represent averages of 75 observations each. This averaging was performed to facilitate the examination of trends rather than variability. The data were first sorted by predicted headway, segregated into contiguous groups of 75 observations, and then average headways were computed for each group. It should be noted that the individual headway observations were used for the model calibration; averaged data were used for plotting purposes only.

Table C-5. Calibrated left-turn movement minimum discharge headway model

Model Statistics		Value		
R^2 :		0.05		
Root Mean Square Error:		0.44 sec/veh		
Observations:		4,153		
Range of Model Variables				
Variable	Variable Name	Units	Minimum	Maximum
H_h	Left-turn movement min. discharge headway	sec/veh	0.83	3.5
R	Radius of curvature of travel path	meters	15	98
v_l'	Demand flow rate per lane (traffic pressure)	vpcpl	5	26
g/C	Effective green to cycle length ratio	na	0.06	0.55
Specified Parameter Values				
Variable	Variable Name	Units	Value	
g_x	Maximum g/C ratio	na	0.27	
Calibrated Coefficient Values				
Coeff.	Coefficient Definition	Value	Std. Dev.	t-statistic
b_0	Intercept	1.55	0.034	45.6
b_1	Effect of radius of travel path	1.71	0.29	5.9
b_2	Effect of traffic pressure	0.00630	0.00090	7.0
b_3	Effect of g/C ratio	0.868	0.128	6.8

Predicted Minimum Discharge Headway, sec/veh

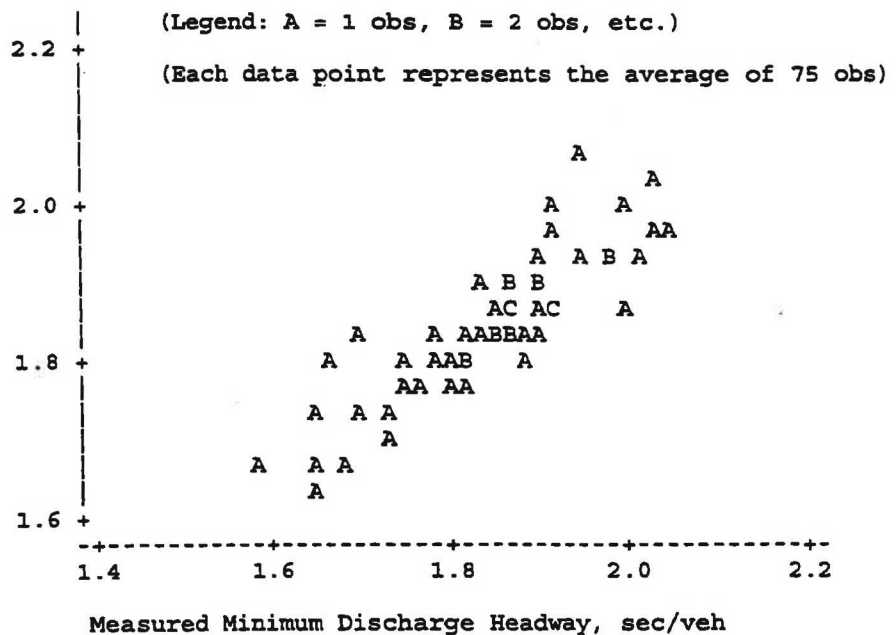


Figure C-12. Comparison of predicted and measured left-turn minimum discharge headways.

The predictive ability of the calibrated model can also be assessed by comparing it with the models developed by Bonneson (2) and by Kimber *et al* (6). The three models are compared in Figure C-13 over a range of left-turn radii. As this figure indicates, the model developed for this research is in good agreement with both of these previously calibrated models.

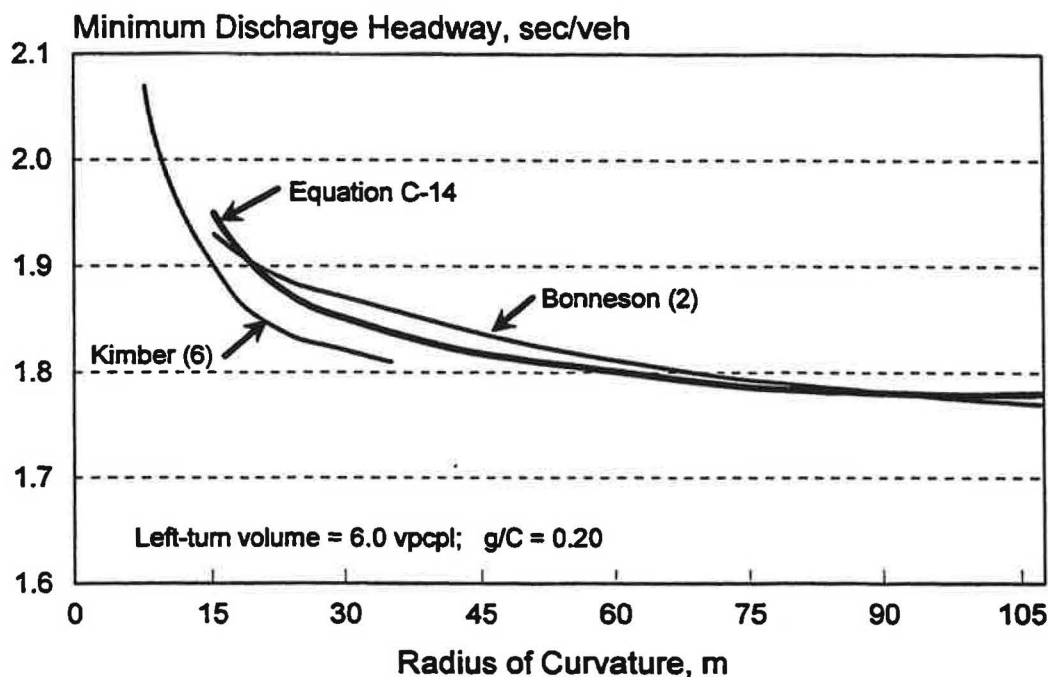


Figure C-13. Comparison of alternative models of left-turn minimum discharge headway.

Sensitivity Analysis. The calibrated model was used to examine the effect of radius, traffic pressure, and g/C ratio on the minimum discharge headway and saturation flow rate of left-turn movements. This effect is shown in Figures C-14 and C-15 for the respective characteristics. The trend lines shown in these figures reflect left-turn volumes of 3 and 10 vpcpl and g/C ratios of 0.16 and 0.27. These ranges were selected to be inclusive of about 90 percent of the observations in the database. In general, the model has a trend of decreasing headway with radius. The range in traffic pressure considered makes a difference of about 0.08 sec/veh in minimum discharge headway and about 100 vphgpl in saturation flow rate. In contrast, the g/C ratio has almost twice the effect as traffic pressure (i.e., a change of about 0.15 sec/veh and 160 vphgpl). Of course, g/C ratio has no effect when the ratio increases beyond 0.27.

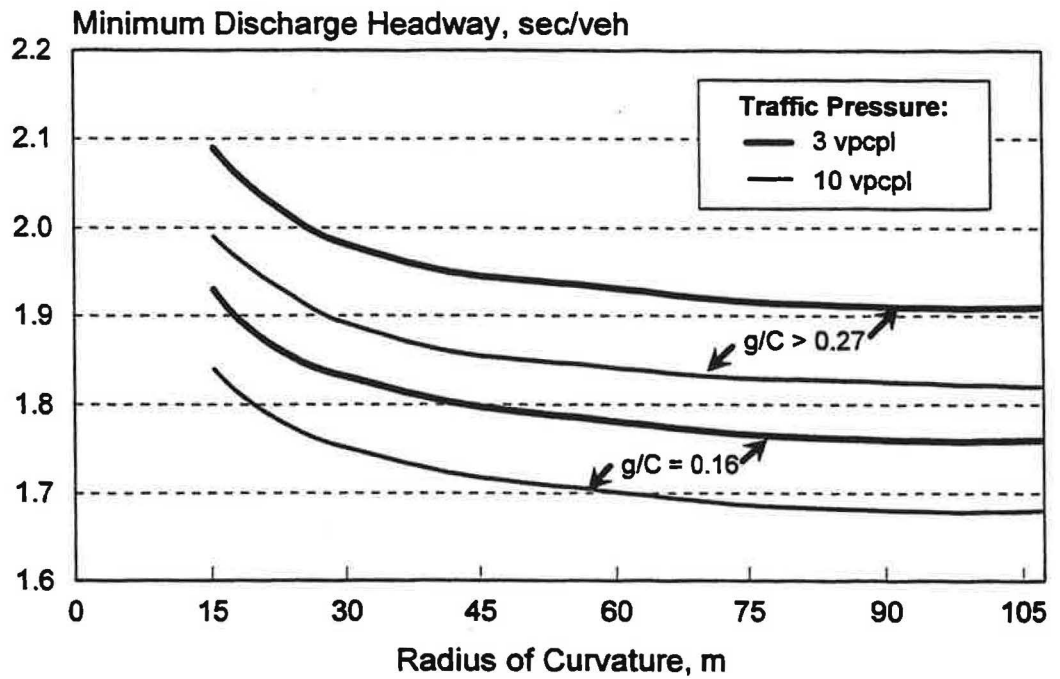


Figure C-14. Effect of traffic pressure, signal timing, and radius on left-turn movement minimum discharge headway.

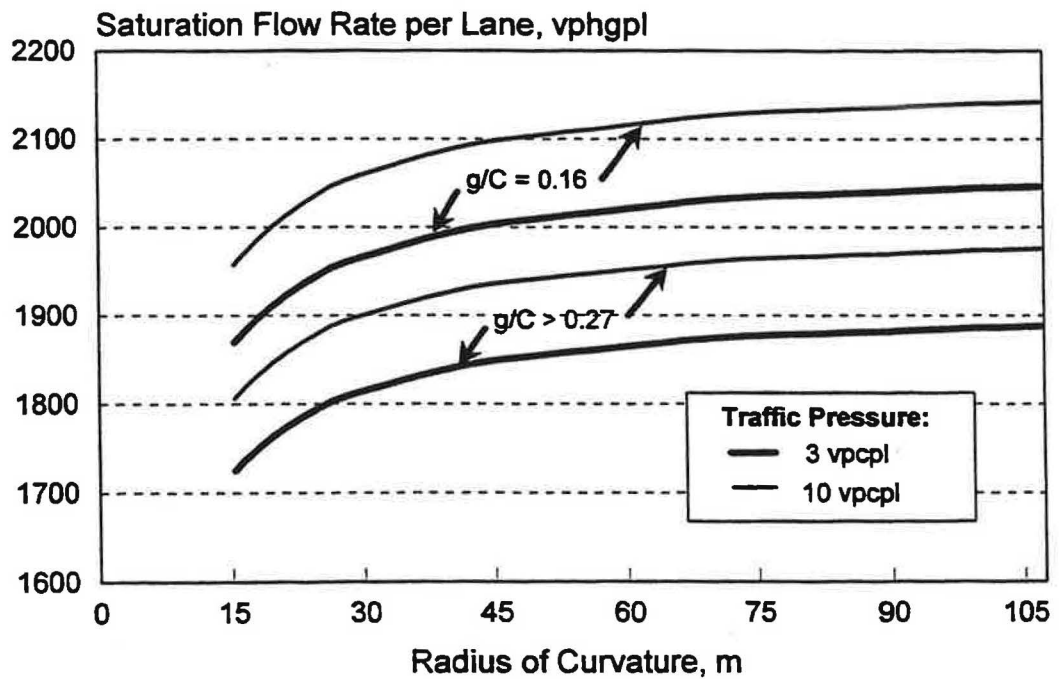


Figure C-15. Effect of traffic pressure, signal timing, and radius on left-turn movement saturation flow rate.

Saturation Flow Rate Model. The calibrated left-turn movement minimum discharge headway model was converted into an equivalent saturation flow rate model. The form of this model was patterned after that used in Chapter 9 of the HCM (1). Specifically, the saturation flow rate for a particular location is estimated as the product of the ideal saturation flow rate and the various site-specific adjustment factors. In this context, the adjustment factors found in this research relate to the effect of traffic pressure, signal timing, and turn radius. The basic form of the model is:

$$s_l = s_0 \times f_R \times f_v \times f_{g/C} \quad (\text{C-16})$$

where:

- s_l = saturation flow rate per lane under prevailing conditions, vphgpl;
- s_0 = saturation flow rate per lane under ideal conditions, pcphgpl;
- f_R = adjustment factor for the radius of the travel path;
- f_v = adjustment factor for volume level (i.e., traffic pressure); and
- $f_{g/C}$ = adjustment factor for signal timing.

The ideal saturation flow rate represents the saturation flow rate when not affected by any external environmental factors (i.e., grade), atypical vehicles (i.e., trucks), and constrained geometrics (e.g., less than 3.6-meter lane widths, curved travel path). In this regard, the saturation flow rate would be equal to the ideal rate when all factor effects are optimum for efficient traffic flow and the corresponding adjustment factors are equal to 1.0. Based on this definition, it was determined that an infinite radius, a traffic pressure of 10.0 vpcpl, and a g/C ratio greater than 0.27 were representative of ideal conditions. Using this definition of ideal conditions and associated parametric values, the resulting ideal flow rate and adjustment factors were algebraically derived from Equation C-14 as:

$$\begin{aligned} s_0 &= \frac{3,600}{b_0 (1 - 10.0 b_2) (1 + 0.27 b_3)} \\ &= 2,010 \text{ pcphgpl} \end{aligned} \quad (\text{C-17})$$

$$f_R = \frac{1}{1 + \frac{1.71}{R}} \quad (\text{C-18})$$

$$\begin{aligned} f_v &= \frac{1}{(1 - b_2 v_l') / (1 - b_2 10.0)} \\ &= \frac{1}{1.07 - 0.00672 v_l'} \end{aligned} \quad (\text{C-19})$$

$$f_{g/C} = \frac{1}{(1 + b_3 t_g) / (1 + b_3 0.27)} \quad (C-20)$$

$$= \frac{1}{0.810 + 0.703 t_g}$$

$$t_g = \left(\frac{g}{C} \right) I_g + 0.27 (1 - I_g) \quad (C-21)$$

where:

R = radius of curvature of the left-turn travel path (at center of path), m;

v_l' = demand flow rate per lane (i.e., traffic pressure), vpcpl;

g = effective green time where platoon motion (flow) can occur, sec;

C = cycle length, sec;

t_g = signalization variable ($0.0 < t_g < 0.27$);

I_g = indicator variable (1.0 if $g/C < 0.27$, 0.0 otherwise); and

b_0, b_2, b_3 = calibration coefficients from Table C-5.

The precision of the ideal saturation flow rate predicted by Equation C-17 is estimated as ± 12 pcphgpl, based on the root mean square error and the number of observations shown in Table C-5. As the resulting range of possible true ideal values includes 2,000 pcphgpl, that this latter value is recommended as the ideal saturation flow rate for left-turn movements.

The relationship between turn radius and the corresponding saturation flow adjustment factor f_R is shown in Figure C-16. The trends shown are similar to those noted for the saturation flow rate shown in Figure C-15. This figure extends Figure C-15 by including the adjustment factor developed by Messer *et al* (3, p. 46). Inspection of the trend line corresponding to Messer's model indicates that it yields slightly higher values than those of Equation C-18. This difference is due primarily to the fact that Messer's model includes the effect of traffic pressure.

C.2 START-UP LOST TIME

The effective time "lost" at the start of a phase (i.e., the start-up lost time) stems from the fact that the headways of the vehicles in the first few queue positions are larger than those of vehicles in the higher queue positions. The headways of these first few queued vehicles are large because of the acceleration they are undergoing as they cross the stop line. Once the vehicles in these positions near the "desired" discharge speed, their headways converge to the minimum discharge headway. Thus, factors that influence this speed (e.g., distance to queue, radius, etc.) also affect minimum discharge headway and saturation flow rate. As a result, there is an inherent relationship between saturation flow rate and start-up lost time.

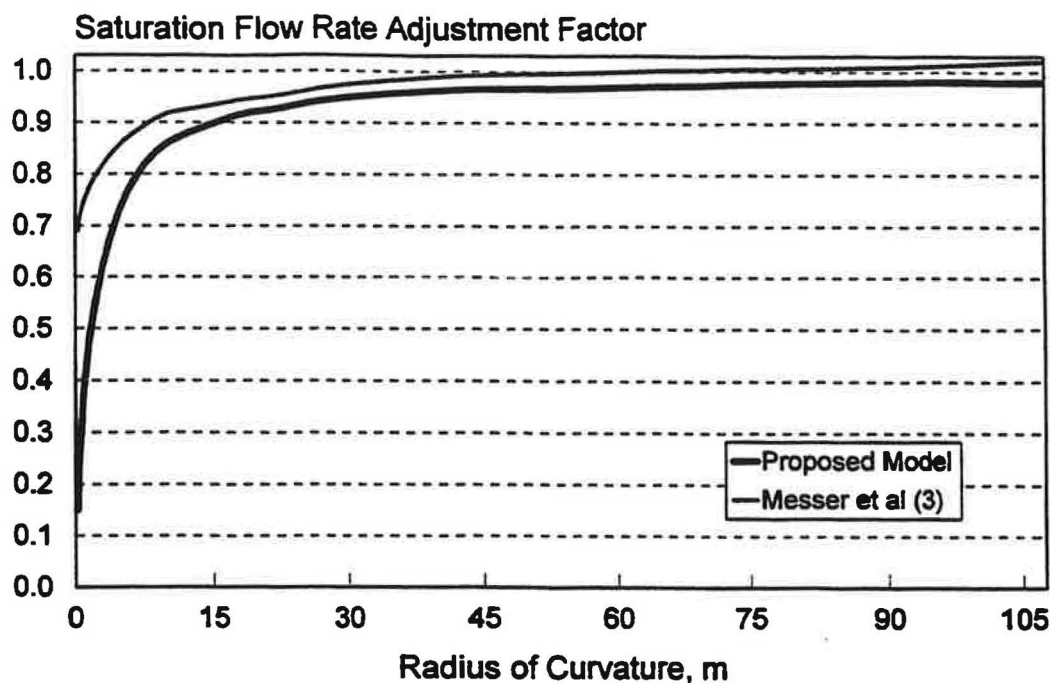


Figure C-16. Effect of turn radius on the saturation flow rate adjustment factor.

This section describes the development and calibration of the start-up lost time model. As with the saturation flow rate model, a start-up lost time model is developed for both the left-turn and the through movement.

C.2.1 Start-Up Lost Time Model for Through Movements

Effect of Saturation Flow Rate on Start-Up Lost Time. Equations C-1, C-2, and C-3, were used to compute the minimum discharge headway, saturation flow rate, and start-up lost time for the through movements studied. The relationship found between start-up lost time and saturation flow rate is shown in Figure C-17. The data shown represent the average of 50 observations. These observations were generated by first sorting the database by saturation flow rate, combining the sorted values into groups of 50, and then computing the average saturation flow rate and start-up lost time for each group. This averaging was performed to better illustrate the trend-wise relationship between the two variables in light of the presentation problems associated with plotting hundreds of individual observations.

As Figure C-17 indicates, there is a strong correlation between start-up lost time and saturation flow rate. Specifically, start-up lost time increases in a linear manner with increasing saturation flow rate. Start-up lost times for the database range from about 0.5 to 4.0 seconds for saturation flow rates ranging from 1,400 to 2,300 vphgpl.

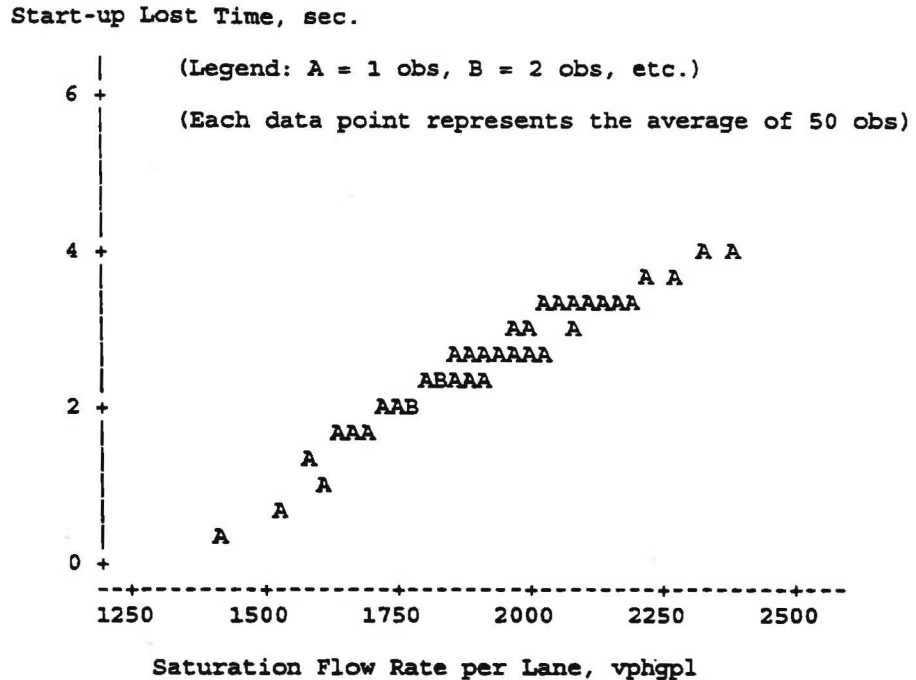


Figure C-17. Effect of through movement saturation flow rate on start-up lost time.

The trend in Figure C-17 supports the contention that the effective amount of time lost at the start of a phase is dependent on the saturation flow rate. In particular, it suggests that more time is lost when the saturation flow rate (and corresponding desired speed) is higher. In a summary of cause and effect, start-up lost time increases with saturation flow rate because it takes more time for the discharging queue to attain the higher discharge speed associated with a higher saturation flow rate. Of course, any resulting loss in capacity due to the added lost time is offset by an increased saturation flow rate.

Model of Start-up Lost Time. Based on the trend in the data shown in Figure C-17, the following model form was developed for the start-up lost time model:

$$l_s = b_0 + b_1 s_l \quad (C-22)$$

where:

l_s = start-up lost time, sec; and

s_l = saturation flow rate per lane under prevailing conditions; vphgpl.

Model Calibration. The details of the model calibration analysis and statistics describing the predictive performance of the model are presented in Table C-6. As the statistics in this table indicate, the calibrated model explains about 41 percent of the variability in the data which is indicative of a strong correlation between start-up lost time and saturation flow rate. Based on the root mean square error and number of observations, the minimum precision of the average start-up lost time estimate is about ± 0.02 sec.

Table C-6. Calibrated through movement start-up lost time model

Model Statistics		Value		
R^2 :		0.41		
Root Mean Square Error:		1.07 sec		
Observations:		1,927		
Range of Model Variables				
Variable	Variable Name	Units	Minimum	Maximum
l_s	Start-up lost time	sec	-3.1	8.0
s_l	Saturation flow rate per lane	vphgpl	1,257	3,326
Calibrated Coefficient Values				
Coeff.	Coefficient Definition	Value	Std. Dev.	t-statistic
b_0	Intercept	-4.64	0.199	-23.3
b_1	Effect of saturation flow rate	0.00373	0.00010	37.3

The quality of fit of the calibrated model to the data is shown in Figure C-18. For reasons discussed previously, the plotted data points represent averages of 50 observations each. The data shown in Figure C-18 confirm that the model is able to accurately predict the start-up lost time over the range of saturation flow rates included in the database.

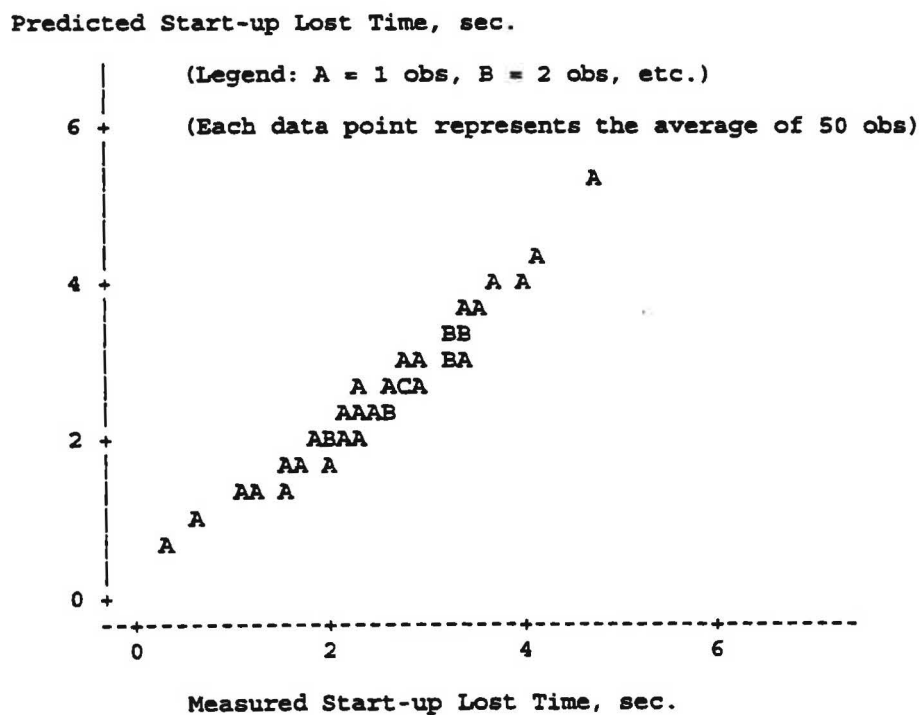


Figure C-18. Comparison of predicted and measured through movement start-up lost times.

Sensitivity Analysis. The calibrated model of start-up lost time was used to examine the sensitivity of through movement start-up lost time to saturation flow rate. This relationship is shown in Figure C-19. As this figure indicates, the start-up lost times for saturation flow rates of 1,800 and 1,900 vphgpl are about 2.0 and 2.5 seconds, respectively. These values are slightly larger than the 1.0 to 2.0 seconds recommended in Chapter 2 of the HCM (1).

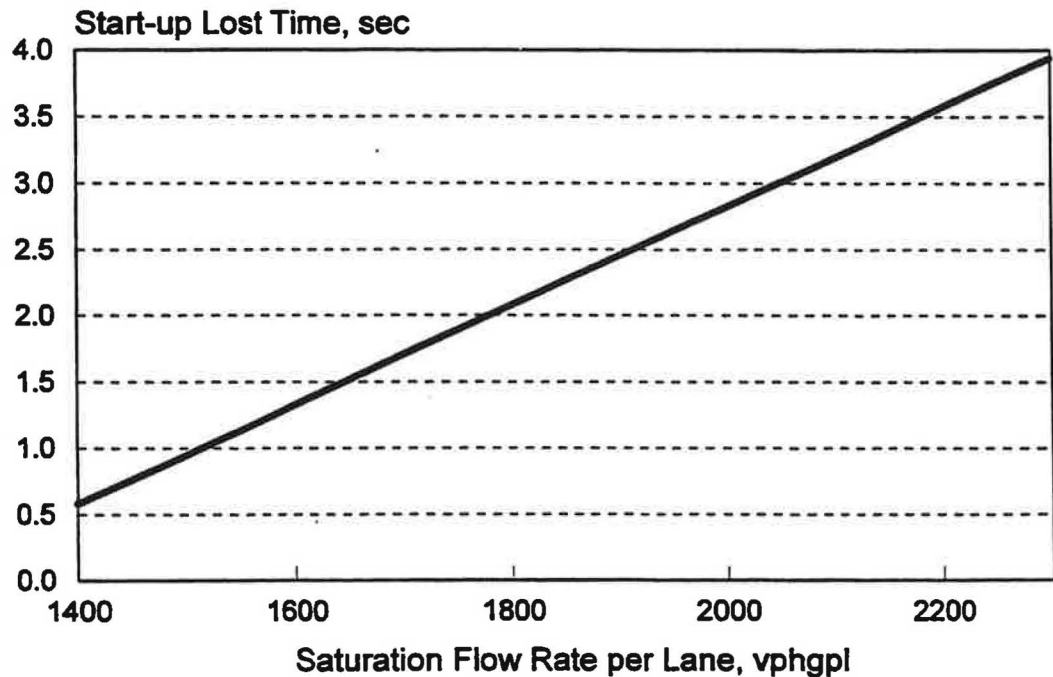


Figure C-19. Expected through movement start-up lost time as a function of saturation flow rate.

C.2.2 Start-up Lost Time Model for Left-Turn Movements

Effect of Saturation Flow Rate on Start-up Lost Time. Equations C-1, C-2, and C-3, were used to compute the minimum discharge headway, saturation flow rate, and start-up lost time for the left-turn movements studied. These movements primarily include the off-ramp and arterial left-turns at interchange ramp terminals; however, the relationships found are also applicable to the left-turn movements at adjacent signalized intersections. The relationship found between start-up lost time and saturation flow rate is shown in Figure C-20.

The data shown in Figure C-20 represent the average of 50 observations. This averaging was performed to better illustrate the relationship between the saturation flow rate and start-up lost time in light of the presentation problems associated with plotting hundreds of individual observations. These observations were generated by first sorting the database by saturation flow rate, combining the sorted values into groups of 50, and then computing the average saturation flow rate and start-up lost time for each group. These averages were computed for graphical presentation purposes only, the individual observations were analyzed during model calibration.

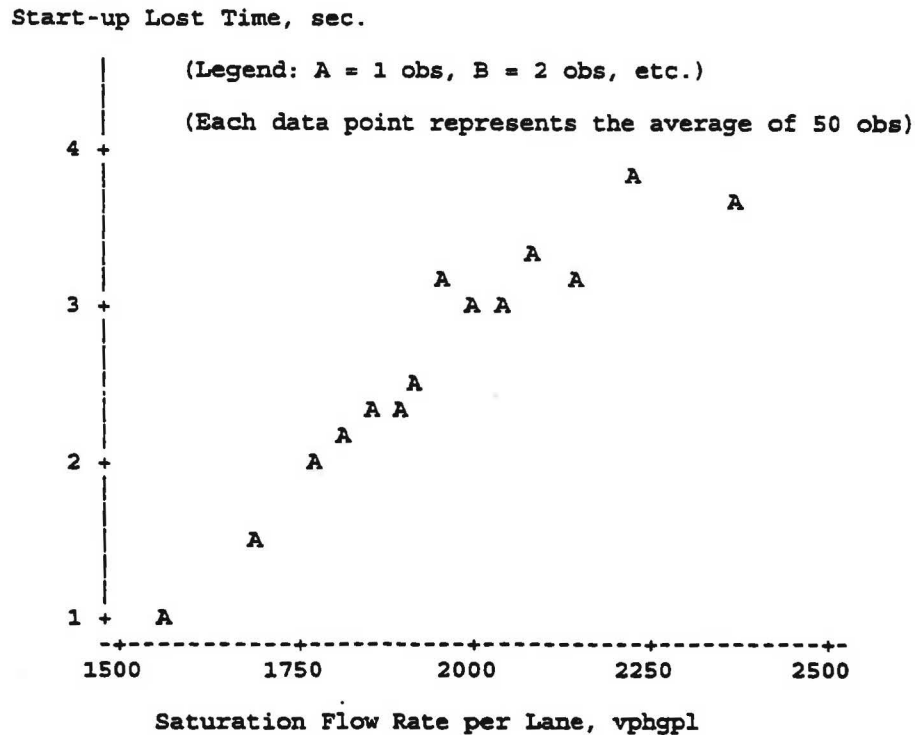


Figure C-20. Effect of left-turn movement saturation flow rate on start-up lost time.

As this figure indicates, there is a strong correlation between start-up lost time and saturation flow rate. Specifically, start-up lost time increases in a linear manner with increasing saturation flow rate. Start-up lost times for the database range from about 1.0 to 4.0 seconds for saturation flow rates ranging from 1,550 to 2,300 vphgpl.

Model of Start-up Lost Time. Based on the trend in the data shown in Figure C-20, the following model form was developed for the left-turn movement start-up lost time model:

$$l_s = b_0 + b_1 s_l \quad (C-23)$$

where:

l_s = start-up lost time, sec; and

s_l = saturation flow rate per lane under prevailing conditions, vphgpl.

Model Calibration. The details of the model calibration analysis and statistics describing the predictive performance of the model are presented in Table C-7. As the statistics in this table indicate, the calibrated model explains about 34 percent of the variability in the headway data which is indicative of a good correlation between start-up lost time and saturation flow rate. The minimum precision of the average start-up lost time estimate is about ± 0.04 sec.

Table C-7. Calibrated left-turn movement start-up lost time model

Model Statistics		Value		
R^2 :		0.34		
Root Mean Square Error:		1.14 sec		
Observations:		714		
Range of Model Variables				
Variable	Variable Name	Units	Minimum	Maximum
l_s	Start-up lost time	sec	-1.5	7.8
s_l	Saturation flow rate per lane	pcphgpl	1,339	2,770
Calibrated Coefficient Values				
Coeff.	Coefficient Definition	Value	Std. Dev.	t-statistic
b_0	Intercept	-4.43	0.373	-11.9
b_1	Effect of saturation flow rate	0.00362	0.00019	19.1

The quality of fit of the calibrated model to the data is shown in Figure C-21. For reasons discussed previously, the plotted data points represent averages of 50 observations each. The data shown in Figure C-21 confirm that the model is able to accurately predict the start-up lost time over the range of saturation flow rates included in the database.

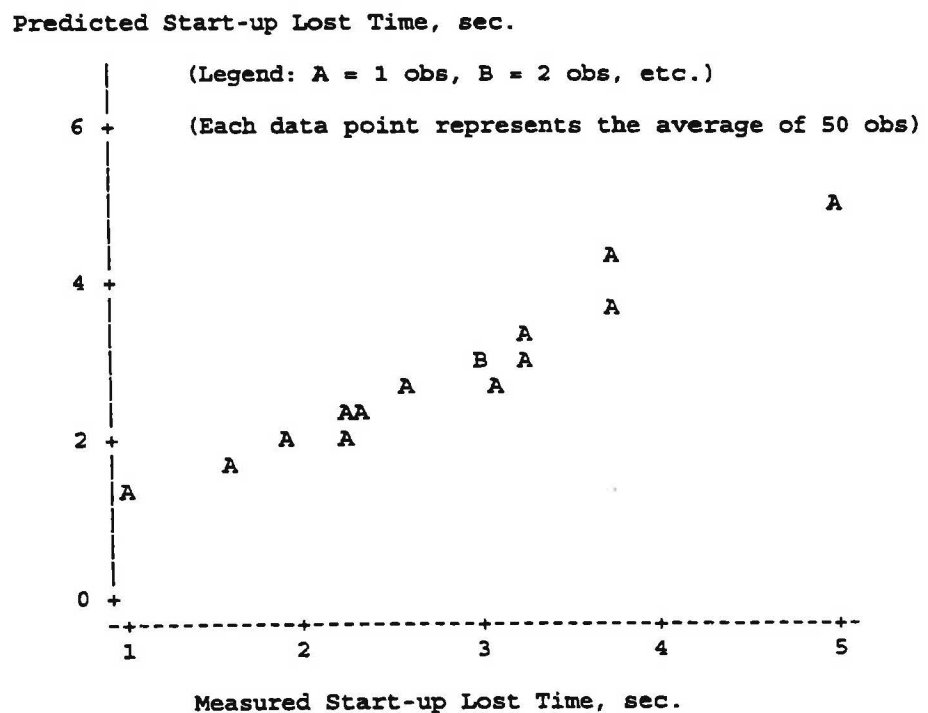


Figure C-21. Comparison of predicted and measured left-turn movement start-up lost times.

Sensitivity Analysis. The calibrated model of start-up lost time was used to examine the sensitivity of left-turn movement start-up lost time to saturation flow rate. This relationship is shown in Figure C-22. As this figure indicates, the start-up lost times for saturation flow rates of 1,800 and 1,900 pcphgpl are about 2.0 and 2.5 seconds, respectively. These values are slightly larger than the 1.0 to 2.0 seconds recommended in Chapter 2 of the HCM (1).

The trend line for through movements found in the preceding section is also included in Figure C-22. Comparison of this line with that for the left-turn movement indicates very little difference between the two movements. Therefore, it appears reasonable to conclude that the effect of saturation flow rate on start-up lost time is independent of movement type.

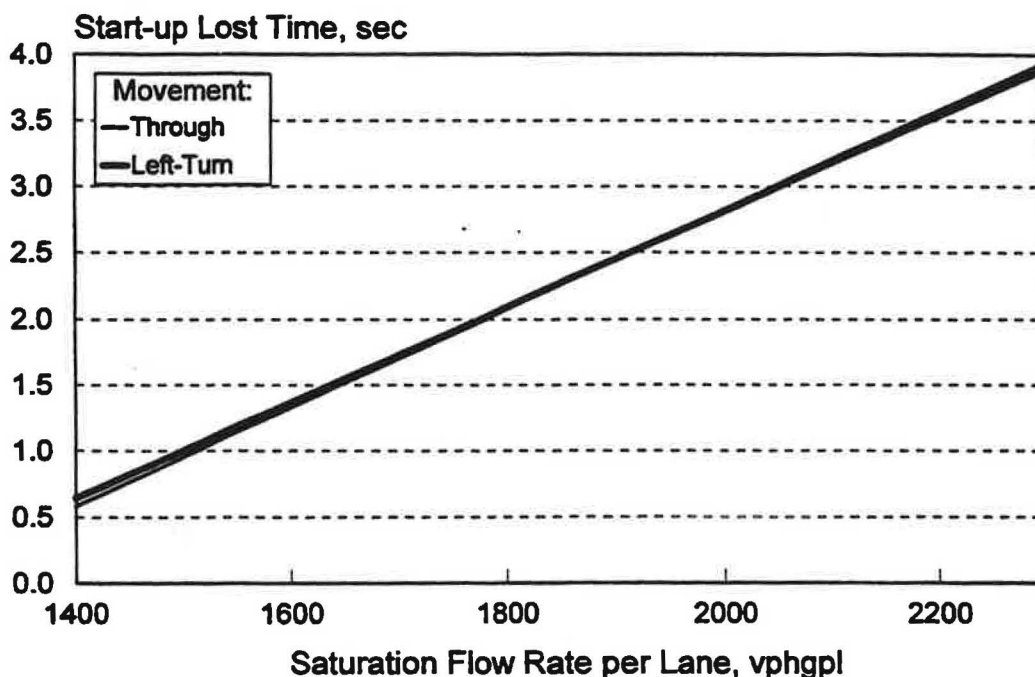


Figure C-22. Expected left-turn movement start-up lost time as a function of saturation flow rate.

C.3 CLEARANCE LOST TIME

When the yellow indication is presented at the end of a phase, drivers close to the intersection generally continue on through the intersection because stopping would be impossible or, at least, very uncomfortable. Thus, these "clearing" drivers tend to use the first few seconds of the yellow interval. The remaining portion of the yellow interval that is not used by the average clearing driver plus the red clearance interval is defined as clearance lost time. Based on this definition, the following equation can be used to compute clearance lost time:

$$l_c = Y + R_c - g_Y \quad (C-24)$$

Table C-7. Calibrated left-turn movement start-up lost time model

Model Statistics		Value		
R^2 :		0.34		
Root Mean Square Error:		1.14 sec		
Observations:		714		
Range of Model Variables				
Variable	Variable Name	Units	Minimum	Maximum
l_s	Start-up lost time	sec	-1.5	7.8
s_l	Saturation flow rate per lane	pcphgpl	1,339	2,770
Calibrated Coefficient Values				
Coeff.	Coefficient Definition	Value	Std. Dev.	t-statistic
b_0	Intercept	-4.43	0.373	-11.9
b_1	Effect of saturation flow rate	0.00362	0.00019	19.1

The quality of fit of the calibrated model to the data is shown in Figure C-21. For reasons discussed previously, the plotted data points represent averages of 50 observations each. The data shown in Figure C-21 confirm that the model is able to accurately predict the start-up lost time over the range of saturation flow rates included in the database.

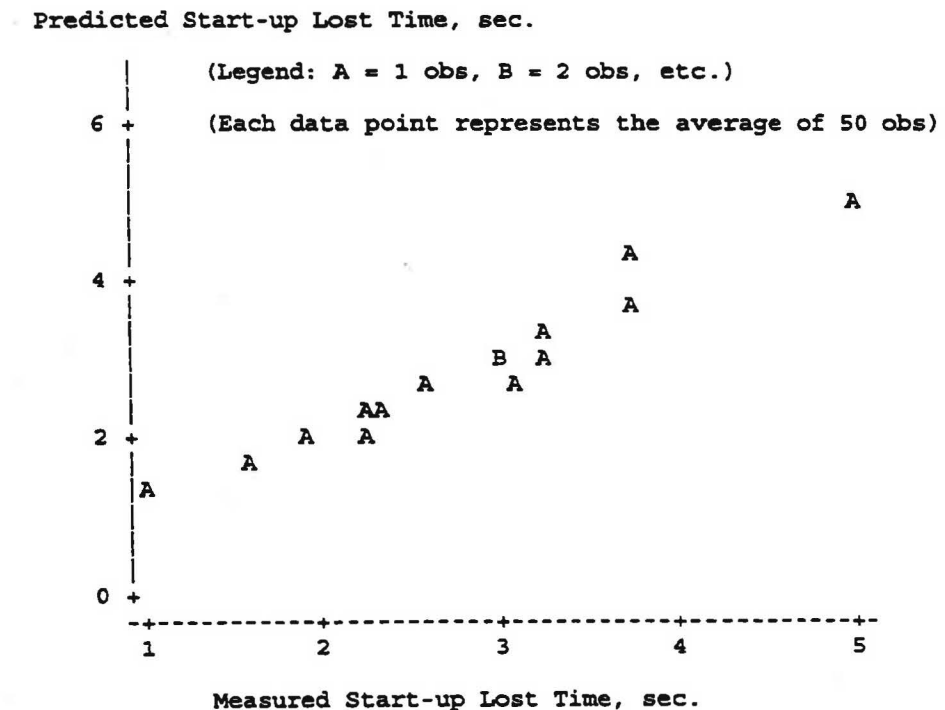


Figure C-21. Comparison of predicted and measured left-turn movement start-up lost times.

Sensitivity Analysis. The calibrated model of start-up lost time was used to examine the sensitivity of left-turn movement start-up lost time to saturation flow rate. This relationship is shown in Figure C-22. As this figure indicates, the start-up lost times for saturation flow rates of 1,800 and 1,900 pcphgpl are about 2.0 and 2.5 seconds, respectively. These values are slightly larger than the 1.0 to 2.0 seconds recommended in Chapter 2 of the HCM (1).

The trend line for through movements found in the preceding section is also included in Figure C-22. Comparison of this line with that for the left-turn movement indicates very little difference between the two movements. Therefore, it appears reasonable to conclude that the effect of saturation flow rate on start-up lost time is independent of movement type.

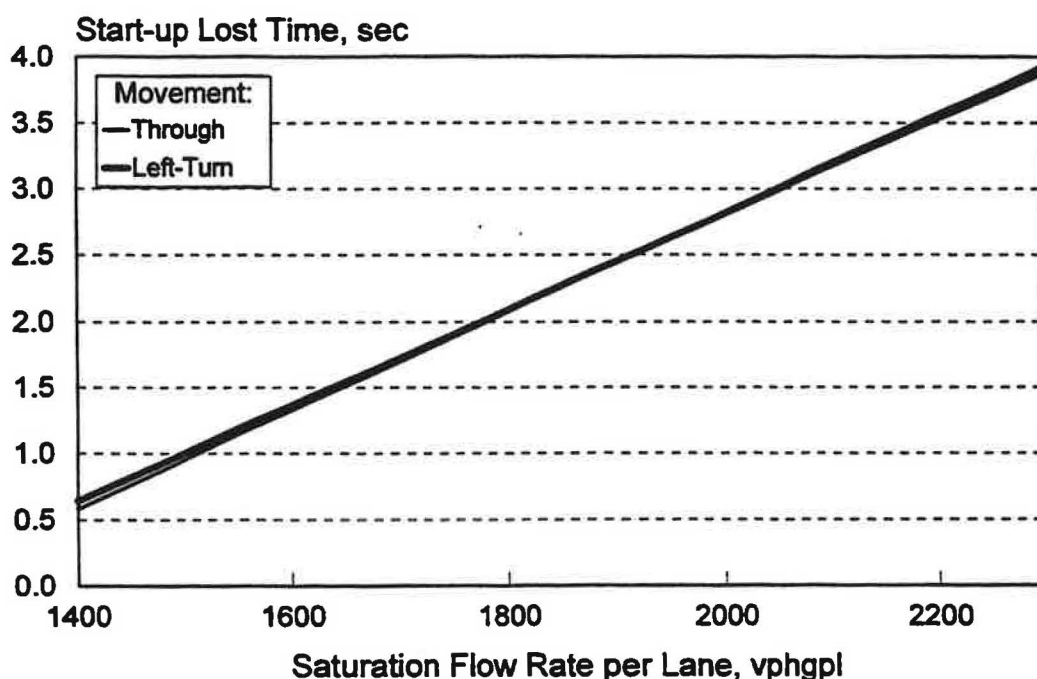


Figure C-22. Expected left-turn movement start-up lost time as a function of saturation flow rate.

C.3 CLEARANCE LOST TIME

When the yellow indication is presented at the end of a phase, drivers close to the intersection generally continue on through the intersection because stopping would be impossible or, at least, very uncomfortable. Thus, these "clearing" drivers tend to use the first few seconds of the yellow interval. The remaining portion of the yellow interval that is not used by the average clearing driver plus the red clearance interval is defined as clearance lost time. Based on this definition, the following equation can be used to compute clearance lost time:

$$l_c = Y + R_c - g_T \quad (C-24)$$

where:

- l_c = clearance lost time, sec;
- Y = yellow interval, sec;
- R_c = red clearance interval, sec; and
- g_Y = effective green extension into the yellow interval, sec.

It was hypothesized that extent of driver encroachment into the yellow interval could be affected by the clearing vehicle's speed, the width of the intersection, and the "cost" of not clearing (i.e., delay). Thus, the calibration activity for this model focused on defining a relationship between green extension, speed, intersection width, and signal timing (as a surrogate for delay). Green extension was quantified as the average amount of the yellow interval that was used during phases where it was observed to be used to some degree. Phases where the yellow interval was not used were excluded from the analysis because the reason for this lack of use was not determinable from the available data. In general, the yellow interval was used in about 27 percent of the phases; although, this frequency varied widely among the twelve study sites.

C.3.1 Model Development

Based on the preceding discussion, it was reasoned that the prediction of clearance lost time would require a model of green extension. Thus, the objective of the model development process was to determine the nature and magnitude of the effect of various factors on the amount of the yellow interval used by drivers. Several potentially influential factors were identified from past research on lost time and from insights obtained during the field studies, these factors include:

1. Number of vehicles per cycle.
2. Yellow interval duration.
3. Red clearance interval duration.
4. Saturation flow rate for the respective movement.
5. Green interval duration.
6. Intersection width (clearance distance).
7. Speed of the clearing vehicle.

After an exploratory analysis, several of the factors listed above were found to have an insignificant correlation with green extension and were eliminated from further consideration. Specifically, the yellow, red clearance, and green interval durations as well as the intersection width were not found to be correlated with green extension. In contrast, the number of vehicles per cycle, minimum discharge headway, and clearance speed were found to be correlated with green extension. Further examination of the first two of these effects suggested that the number of vehicles per cycle relative to the maximum number that can be served per phase (i.e., volume-to-capacity ratio) would be an appropriate means of examining their combined effect on green extension. Specifically, these variables were combined as follows:

$$X_i = \frac{v_i'}{c_i'} \quad (C-25)$$

where:

- X_i = volume-to-capacity ratio in lane i , $i = 1, 2, \dots, N$;
- v_i = demand flow rate in lane i , vpcpl;
- c_i = capacity of lane i (estimated as $c_i = G / H_i$), vpcpl;
- G = green signal interval, sec; and
- H_i = minimum discharge headway in lane i , sec/veh.

The volume-to-capacity ratio has an intuitive correlation with the amount of the yellow a driver would be willing to use. In particular, as the phase becomes busier (i.e., as X approaches 1.0), drivers will increase the amount of the yellow interval they are willing to use to avoid lengthy delays in spite of reduced safety and an increased chance of receiving a citation.

The relationship between the X -ratio and green extension is shown in Figure C-23. As this figure indicates, the effect increases exponentially with increasing values of X . However, a closer examination indicated that a correlation between green extension and X -ratio only existed (in a statistically significant way) for X values above about 0.90. The data shown in this figure represent averages of 50 observations, sorted by X -ratio. This averaging was performed to facilitate the examination of trends by eliminating some of the variability in the individual observations.

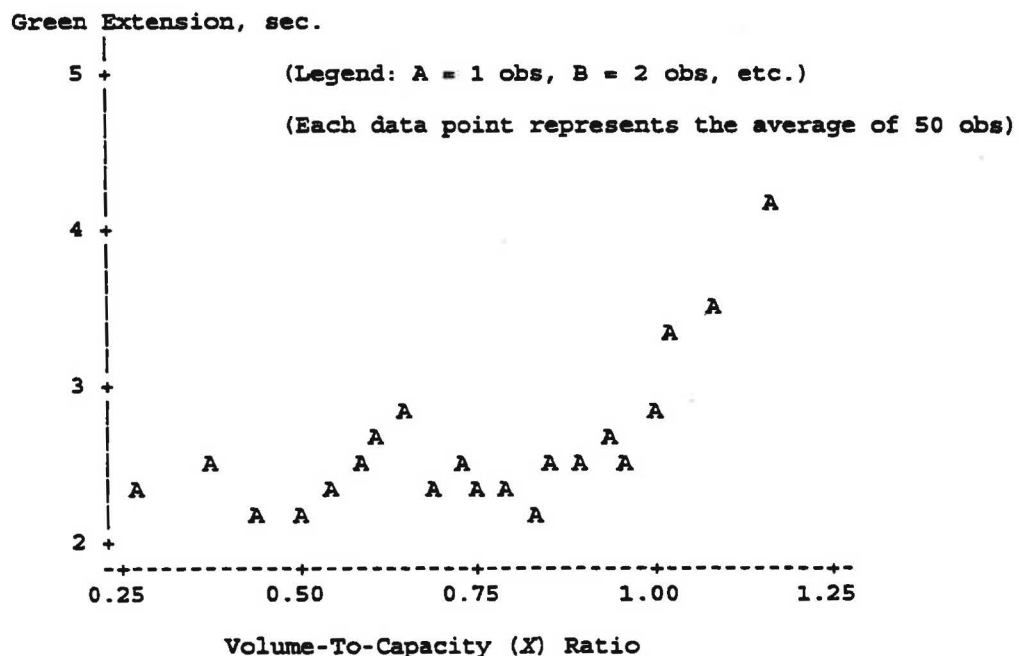


Figure C-23. Effect of x -ratio on green extension.

In addition to X -ratio, clearance speed was also found to be correlated with green extension. Specifically, drivers used more of the yellow interval when they were traveling at higher speeds. This trend suggests that drivers may be more challenged to estimate the adequacy (and safety) of their stopping distance at higher speeds and, as a consequence, err on the side of maintaining speed

and proceeding through the intersection on yellow. Although clearance speed is likely to be more strongly correlated with green extension than any other speed measure, it represents an impractical variable because it is generally an unknown quantity at most intersections. As a more useful alternative, the approach speed limit was substituted in the green extension model as a surrogate measure for clearance speed.

Based on the preceding discussion of variable effects on green extension, the following form of the green extension model was chosen:

$$g_T = b_0 + b_1 SL + b_2 (X_i - b_3) I_X \quad (C-26)$$

where:

- g_T = effective green extension into the yellow interval, sec;
- SL = approach speed limit, km/h;
- X_i = volume-to-capacity ratio in lane i , $i = 1, 2, \dots, N$;
- I_X = indicator variable (1.0 if $X_i > b_3$, 0.0 otherwise); and
- b_0, b_1, b_2, b_3 = calibration coefficients.

C.3.2 Model Calibration

The analysis considered 1,044 signal phases with observed driver use of the yellow (and in some cases, red clearance) interval. These phases were observed at twelve interchange ramp terminals and at twelve intersection approaches. The green extension data used in this analysis represent observations made for both left-turn and through movements. The left-turns at the interchanges were made from either the off-ramp or the arterial. Details of the model calibration analysis and the performance of the calibrated model are presented in Table C-8.

The R^2 statistic in Table C-8 indicates that the calibrated model accounts for eleven percent of the variability in the green extension data. The remaining variability is likely due to random sources; although, some of it may be due to differences among the study sites (that was not explained by speed limit and volume-to-capacity ratio). Nevertheless, it is believed that the calibrated model provides a relatively good fit to the data and that it can be used to predict the average green extension with reasonable precision (i.e., a minimum of ± 0.04 sec.). The quality of fit to the data is also shown in Figure C-24.

Table C-8. Calibrated green extension model

Model Statistics		Value		
R^2 :		0.11		
Root Mean Square Error:		1.33 seconds		
Observations:		1,044		
Range of Model Variables				
Variable	Variable Name	Units	Minimum	Maximum
g_Y	Effective green extension into the yellow interval	sec	0.02	7.3
SL	Approach speed limit	km/h	56	72
G	Green interval duration	sec	10	99
H	Minimum discharge headway	sec	1.7	2.2
v_i'	demand flow rate in lane i	vpcpl	2	37
X_i	Volume-to-capacity ratio in lane i	na	0.08	1.3
Calibrated Parameter Values				
Variable	Definition	Value	Std. Dev.	t-statistic
b_0	Intercept	1.48	0.42	3.5
b_1	Effect of speed limit	0.014	0.0068	2.1
b_2	Effect of X ratio	6.40	0.87	7.4
b_3	Threshold X ratio	0.88	0.022	40.0

Predicted Green Extension, sec

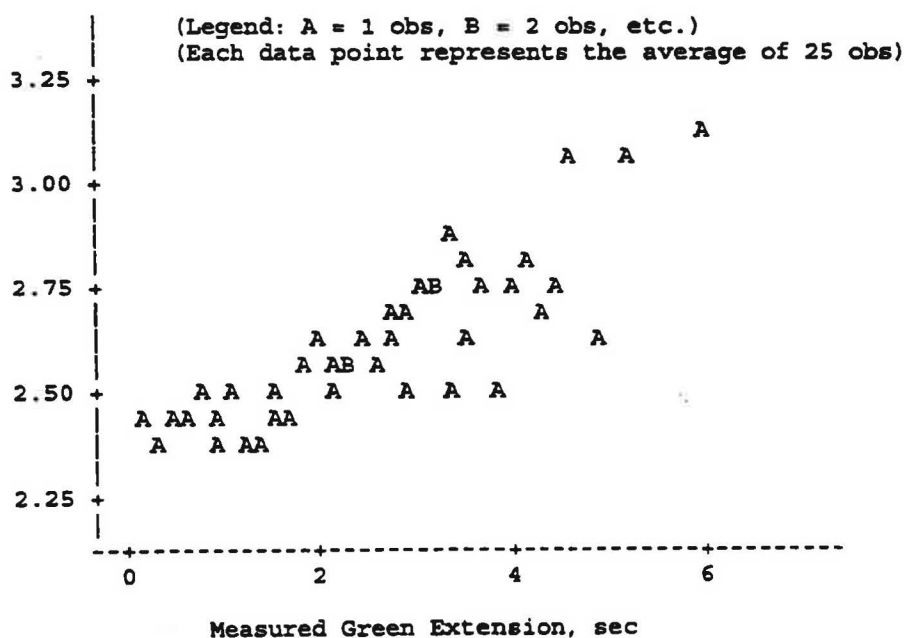


Figure C-24. Comparison of predicted and measured green extension.

C.3.3 Sensitivity Analysis

The calibrated model was used to examine the effect of speed limit and volume-to-capacity ratio on clearance lost time. Prior to conducting this examination, it was necessary to define the duration of the yellow and red clearance intervals. Recognizing that the yellow interval duration is often dependent on the approach speed and that there are a wide range of methods being used to determine yellow interval duration, it was decided to set the yellow interval equal to 0.062 times the approach speed (i.e., $Y = 0.062 * \text{speed limit (km/h)}$). This approach yields values that are generally consistent with those obtained from other methods or policies. The red clearance interval was established as 1.0 second. Using these values, the clearance lost time was computed for a range of speed limits and volume-to-capacity ratios. The results of this examination are shown in Figure C-25.

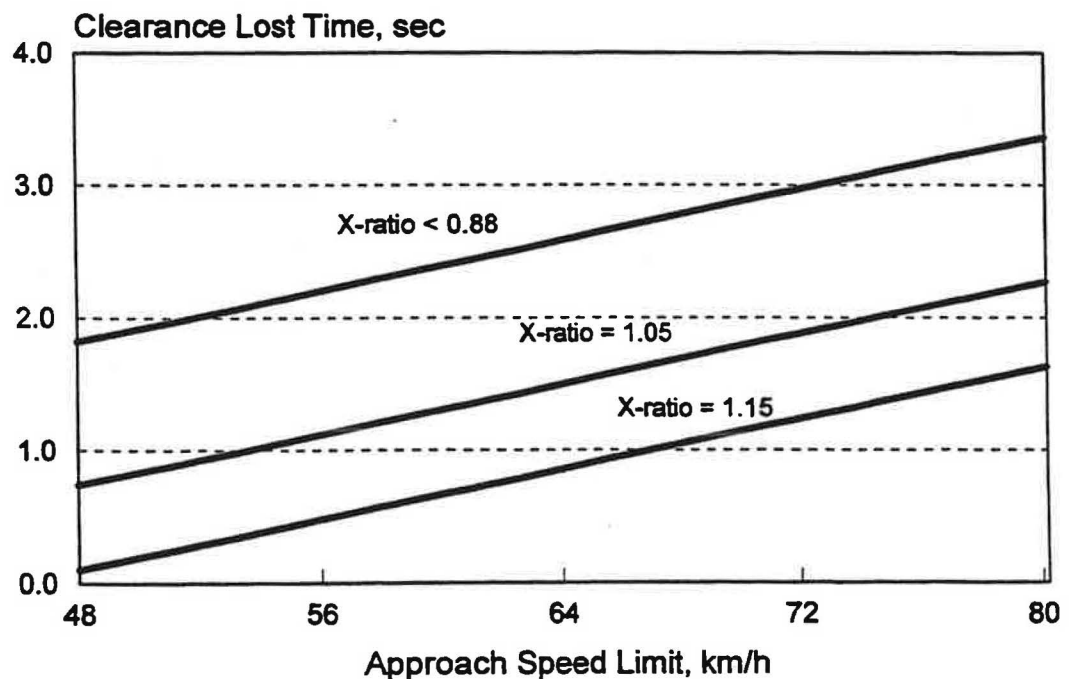


Figure C-25. Effect of approach speed limit on clearance lost time.

As Figure C-25 illustrates, clearance lost time increases with approach speed limit and decreases with increasing volume-to-capacity ratio. In general, it ranges from 1.0 to 3.0 seconds for typical speed limits and uncongested conditions. This range compares with the 1.2 to 2.8-second range for clearance lost time suggested in Chapter 2 of the HCM (1). Clearance lost time increases with speed because of a corresponding increase in the yellow interval; however, it should be noted that this effect is offset to some degree by the increase in green extension associated with higher speeds.

C.4 LANE UTILIZATION

The quality of service provided by a signalized intersection is highly dependent on the volume-to-capacity ratio of the intersection and its associated signal phases. One consideration in determining demand volume for the phase is the distribution of traffic among the lanes it serves. More specifically, these would be the lanes available to a "lane group," as defined in Chapter 9 of the HCM (1). Obviously, if the traffic for a given lane group is concentrated in only one of the several available lanes, then the phase duration would need to be long enough to serve traffic in this one lane. Alternatively, if the phase duration is not increased, then the capacity of the lane group is effectively reduced by the degree of underutilization of its lower volume lanes.

Unfortunately, an even distribution of traffic among the lanes in a lane group is rarely achieved at most intersections for a variety of reasons. One reason is the inherent probability that the number of vehicles arriving per cycle will not be evenly divisible by the number of lanes serving the movement. When this event occurs, the number of vehicles in each through lane will differ by at least one vehicle; perhaps more, if drivers do not distribute themselves evenly among the available lanes. To illustrate this effect, consider six vehicles arriving to an intersection while the signal indication is red. If the subject lane group is provided two lanes, then presumably three vehicles would queue in each lane. However, if seven vehicles arrive, then one lane will have four vehicles and one lane will have three vehicles (as a minimum).

It is important to note that the effect of unbalanced lane volumes is defined herein in the context of "arrivals per cycle." This definition is consistent with the analysis approach described in the HCM (1) and the methods contained therein that account for unequal lane use.

Another reason for an uneven distribution of traffic among the available lanes is driver desire to "preposition" for a turn maneuver at a downstream intersection. This activity commonly occurs on the arterial cross street at interchange ramp terminals and any associated closely-spaced intersections. Specifically, drivers in these areas typically concentrate in one through lane at the upstream ramp terminal or intersection in anticipation of a turn onto the freeway at the downstream ramp terminal.

The uneven allocation of traffic to the available traffic lanes is typically quantified in terms of the lane group Lane Utilization Factor U . This utilization factor can be computed using the following equation, where the volumes used represent averages per signal cycle:

$$U = \frac{v'_{\max} N}{\sum v'_i} \quad (\text{C-27})$$

where:

- U = lane utilization factor for the lane group;
- v'_{\max} = maximum demand flow rate in any of N lanes, vpcpl;
- v'_i = demand flow rate in lane i , $i = 1, 2, \dots, N$, vpcpl; and
- N = number of lanes in the lane group.

The overall effect of number-of-lanes and movement-type on lane use were previously shown in Table B-10. Two trends were suggested by the data in this table. First, the magnitude of the utilization factor tends to increase with number of lanes. This increase is indirectly due to the probability of an uneven distribution of vehicles among the available lanes each cycle (i.e., the arrival volume is not evenly divisible by the number of lanes). The potential for an uneven distribution increases with increasing number of lanes; hence, the utilization factors should tend to increase with an increasing number of lanes.

The second trend suggested by the data in Table B-10 is that the utilization factor for left-turn movements is higher than for through movements, given the same number of lanes. This trend is likely due to the lower volume of the left-turn movements relative to the through movements studied. The effect of volume level on lane utilization can be demonstrated by the following example. Consider a lane group with two traffic lanes where the drivers always distribute themselves as evenly as possible. When there are three arrivals per cycle, Equation C-27 predicts a lane utilization factor of 1.33 ($= 2 \text{ veh} * 2 \text{ lanes} / 3 \text{ arrivals}$); however, when there are 13 arrivals per cycle, Equation C-27 predicts a factor of 1.08 ($= 7 * 2 / 13$). Thus, a low-volume movement will generally have a larger lane utilization factor than a high-volume movement.

C.4.1 Model Development

The lane utilization model developed in this research is based on a quantitative description of the two problems previously described: (1) drivers not distributing themselves as evenly as possible, and (2) drivers prepositioning for a downstream turn. The first problem is more fundamental in nature and deals specifically with the random nature of vehicle arrivals per cycle and the extent that drivers collectively can and will distribute themselves among available traffic lanes. Unbalanced lane use stemming from this problem would be found in any multi-lane lane group on an intersection approach. The second problem is of a site-specific nature as it relates to the effects of downstream turn movements on a driver's lane choice at the upstream intersection. This problem would not necessarily be found in all multi-lane lane groups.

Random Behavior Model. The first problem described above can be modeled by identifying the boundary values of the "most" and the "least" even distribution possible and then determining where drivers naturally fall between these boundaries. The most even distribution possible does not always equate to an equal number of vehicles in each lane due to the integer nature of vehicles and available lanes. When the "most-evenly possible" distribution is achieved, the flow rate in the higher-volume lane can be computed using the probability that the arriving vehicles are (or are not) evenly divisible by the number of available lanes. This probability can then be used to compute an expected maximum lane flow rate $v'_{\max(m)}$.

Consider a two-lane lane group with v arrivals per cycle. One-half of the time, the arrivals would be evenly distributed among available lanes and one-half of the time there would be one more vehicle in the largest lane (assuming a "most-even" posture). Thus, the average (or expected) maximum lane flow rate can be computed as:

$$\begin{aligned}
 E[v'_{\max(m)}] &= \frac{1}{2} \left(\frac{v'}{2} \right) + \frac{1}{2} \left(\frac{v'-1}{2} + 1 \right) \\
 &= \frac{v'}{2} + \frac{1}{4}
 \end{aligned}
 \tag{C-28}$$

where:

$E[v'_{\max(m)}]$ = expected maximum lane flow rate based on the "most-even" distribution possible, vpcpl;
 v' = demand flow rate for the lane group, vpc.

A key assumption made in the development of Equation C-28, that the "even" and "uneven" arrival cases will occur with 50 percent probability (i.e., the "1/2" factor shown), is based on the reasonable presumption that total number of arrivals per cycle is a random variable with a mean arrival rate of v' . An empiric examination of the strength of this assumption was conducted using several types of arrival distributions. Based on this examination, it was concluded that the assumption is valid, provided that the distribution of arrivals has a nonzero variance and is single-moded (i.e., does not have a multi-modal distribution). This examination also indicated that the probability of an "even" distribution for any number of lanes N can be assumed to be equal to $1/N$, provided that the number of arrivals per cycle equals or exceeds the number of lanes.

Based on the preceding discussion, the largest lane flow rate can be computed for the general case of N lanes as:

$$\begin{aligned}
 E[v'_{\max(m)}] &= \frac{1}{N} \left[\sum_{i=0}^{N-1} \left(\frac{v'-i}{N} + 1 \right) - 1 \right] \\
 &= \frac{v'}{N} \left(1 + \frac{N-1}{2} \right)
 \end{aligned}
 \tag{C-29}$$

This equation predicts the flow rate in the highest-volume lane under the assumption that drivers will distribute themselves as evenly as possible upon their arrival. It forms a lower bound on the maximum lane volume.

The counterpart to the "most-even" distribution assumption would be the case where drivers remain in the lane in which they arrived and, thereby, combine in a purely random manner to yield relatively large flow rates in the highest-volume lane. This "least-evenly distributed" scenario forms a reasonable upper bound on the maximum lane volume (precluding the possibility of prepositioning).

When drivers make no attempt to redistribute themselves as they arrive, then the flow rate in the highest-volume lane is a random variable that follows the distribution of maximum values. This behavior represents an extreme case where the resulting lane distribution is the most uneven possible (neglecting prepositioning). Using the Poisson distribution to model the arrival process, the maximum value distribution can be written as:

$$f(v'_{max}) = N \left[\sum_{i=0}^{v'_{max}} \left(\frac{e^{-v'_i} (v'_i)^i}{i!} \right) \right]^{N-1} \frac{e^{-v'_i} (v'_i)^{v'_{max}}}{(v'_{max})!} \quad (C-30)$$

where:

$f(v'_{max})$ = distribution of the maximum demand flow rate in any of N lanes;

N = number of lanes in the lane group; and

v'_i = demand flow rate per lane, vpcpl.

Using this distribution, the expected flow rate in the highest-volume lane $v'_{max(l)}$ can be computed using the following equation:

$$E[v'_{max(l)}] = \sum_{v'_{max}=0}^{\infty} v'_{max} f(v'_{max}) \quad (C-31)$$

where:

$E[v'_{max(l)}]$ = expected maximum lane flow rate based on the "least-even" distribution possible, vpcpl.

Unfortunately, the combination of Equations C-30 and C-31 does not yield a tractable, closed-form solution. Therefore, Equations C-30 and C-31 were used to develop the following empirical model for predicting $v'_{max(l)}$:

$$E[v'_{max(l)}] = \frac{v'}{N} \left(1 + \frac{3N}{4} \sqrt{\frac{N-1}{2v'}} \right) \quad (C-32)$$

This equation was developed from an examination of Equations C-30 and C-31 and the use of heuristic techniques. Other forms are possible, however, Equation C-32 was found to predict the theoretic maximum lane flow rates obtained from Equation C-31 with a relatively high degree of accuracy (i.e., $R^2 = 0.85$) for lane group flow rates in the range of 0.0 to 40 vpc.

To demonstrate the predictive ability of Equation C-32, Equation C-31 was used to compute the theoretic lane utilization factors for a range of flow rates and traffic lanes. These factors are plotted as data points in Figure C-26 along with the trend lines representing Equation C-32. As shown in this figure, the empirical model predicts the lane utilization factors quite well for the two and three-lane cases while it overestimates the factors for the four-lane case by about 4.0 percent.

As Equations C-29 and C-32 represent boundary values for the maximum lane volume, the actual maximum lane volume would fall somewhere between these two extremes. Hence, the expected maximum lane volume can be computed as the proportional combination of both equations:

$$E[v'_{max}] = \frac{v'}{N} \left(1 + \frac{N-1}{2v'} \right) (1 - f_u) + \frac{v'}{N} \left(1 + \frac{3N}{4} \sqrt{\frac{N-1}{2v'}} \right) f_u \quad (C-33)$$

where:

$E[v'_{max}]$ = expected maximum demand flow rate in any of N lanes, vpcpl; and

f_u = proportion of drivers that do not attempt to evenly distribute themselves.

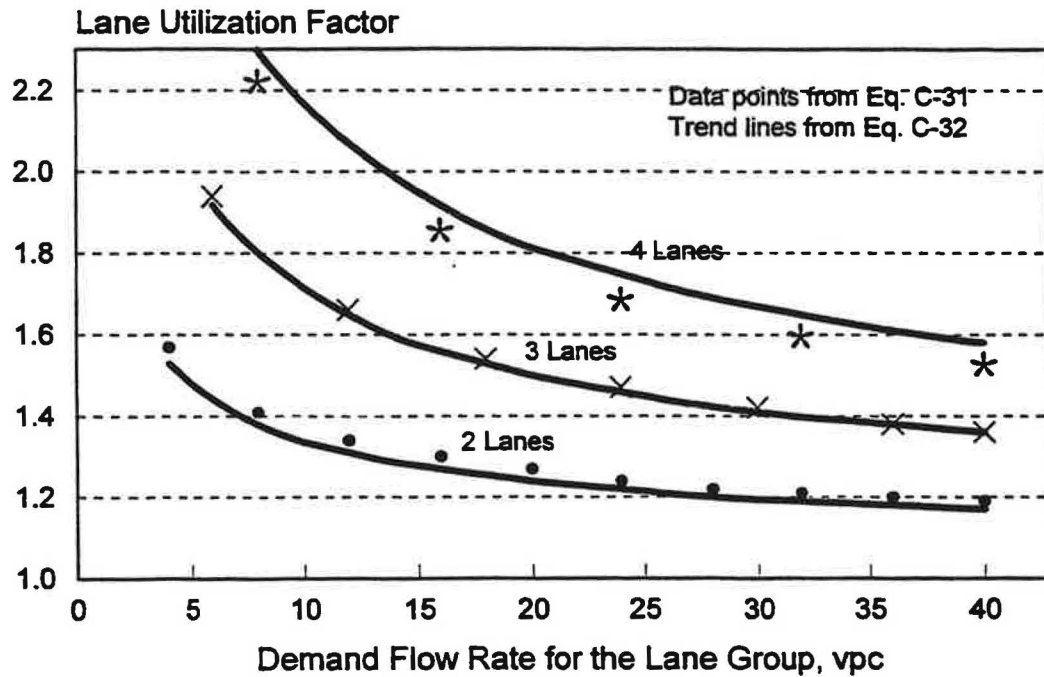


Figure C-26. Comparison of lane utilization factors obtained from Equations C-31 and C-32.

Equations C-27 and C-33 can be combined to yield the expected lane utilization factor for isolated intersections (i.e., no prepositioning) U_r :

$$U_r = 1 + \left(\frac{N-1}{2v'} \right) (1 - f_u) + \frac{3N}{4} \sqrt{\frac{N-1}{2v'}} f_u \quad (C-34)$$

The effect of number-of-lanes, flow rate, and driver propensity to distribute evenly on the lane utilization factor is shown in Figure C-27. As the trends shown in this figure indicate, the lane utilization factor decreases with increasing flow rate. It also decreases with a decreasing number of lanes. This latter trend is consistent with the data shown in Table B-10.

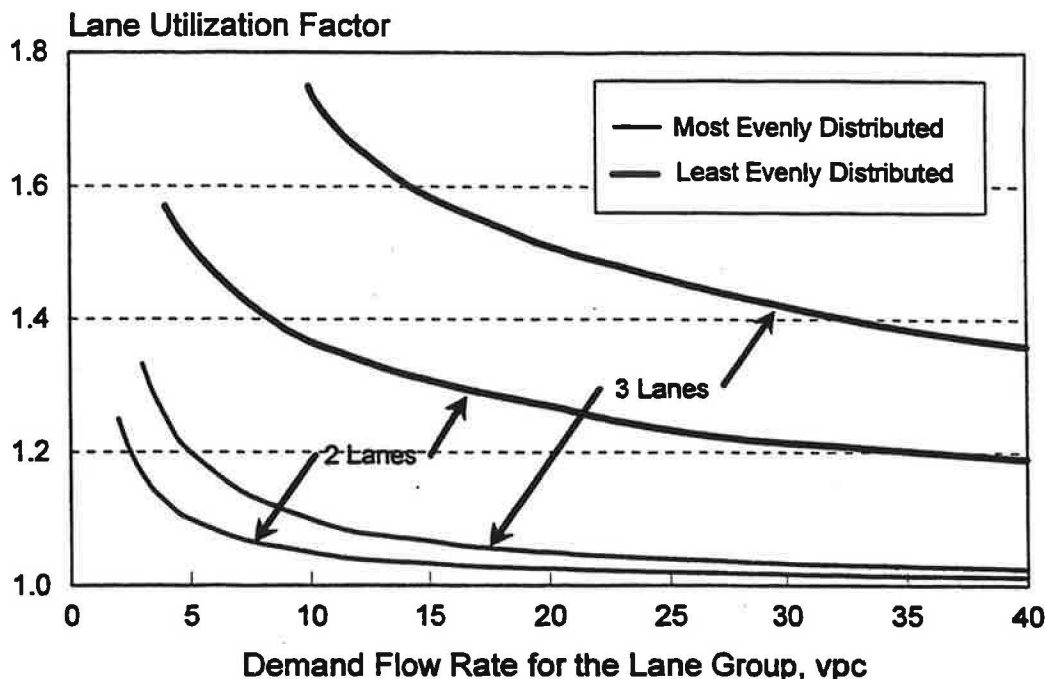


Figure C-27. Range of possible lane utilization factors as affected by volume and lanes.

Also shown in Figure C-27 is the range of possible values of the lane utilization factor, based on the drivers' tendency to distribute themselves on the intersection approach. In general, the range width increases with number of lanes and decreases with lane group flow rate.

The utilization factors recommended by the HCM (1) (i.e., 1.05 at two through lanes, 1.10 at three through lanes) fall within the range of expected lane utilization factors shown in Figure C-27 when the volume exceeds about 10 vpc. The HCM factors do not vary with traffic demand even though the preceding theoretic development indicates that they should have some sensitivity to it. As a result, the HCM factors would appear to be too low when traffic demands are less than 10 vpc.

Prepositioning Model. Driver prepositioning for a turn maneuver at a downstream intersection can affect lane utilization at an upstream intersection if the demand for this turn maneuver is high relative to the volume *not* turning at the next intersection. If the number of drivers not turning are sufficiently high as to maintain an even distribution of traffic in the upstream lanes, then the lane utilization would follow that described for the "Random Behavior Model." On the other hand, if drivers not turning are relatively low such that they cannot maintain an even distribution of traffic, then the prepositioned drivers will dominate one lane and yield a lane utilization factor significantly in excess of 1.0.

The model of this process begins with a test to determine which of the two cases described in the preceding paragraph will occur. This test determines if the number of drivers that will not be turning are sufficient in number to maintain an even lane distribution. The test compares the proportion of prepositioned traffic with the proportion of traffic in any lane when this traffic is

evenly distributed (i.e., $1/N$). When the proportion of prepositioned traffic exceeds $1/N$, then prepositioning will dictate the maximum lane volume. The test equation is:

$$\begin{aligned} \frac{\text{Max}(v'_{dl}, v'_{dr})}{v'} &> \frac{1}{N} : \text{prepositioning} \\ \frac{\text{Max}(v'_{dl}, v'_{dr})}{v'} &\leq \frac{1}{N} : \text{no prepositioning} \end{aligned} \quad (\text{C-35})$$

where:

v'_{dl} = flow rate in the lane group that will be turning left at the downstream intersection, vpc;
 v'_{dr} = flow rate in the lane group that will be turning right at the downstream intersection, vpc;
 N = number of lanes in the lane group; and
 $\text{Max}(v'_{dl}, v'_{dr})$ = larger of v'_{dl} and v'_{dr}

It should be noted that the test equation does not include the distance to the downstream intersection. Logically, the motivation to preposition is dependent on this distance, as drivers are not likely to preposition if the next intersection is quite distant. However, the test equation is believed to be applicable to interchange areas because the inter-signal distances at these locations are sufficiently short that prepositioning is necessary (say, less than 300 meters).

If prepositioning dictates the lane volume distribution, then the largest lane flow rate will be:

$$E[v'_{\max}] = \text{Max}(v'_{dl}, v'_{dr}) \quad (\text{C-36})$$

Similarly, the lane utilization factor for the prepositioning case can be computed as:

$$U_p = \frac{\text{Max}(v'_{dl}, v'_{dr}) N}{v'} \quad (\text{C-37})$$

Combined Models. Equations C-34 and C-37 were combined to yield the generalized lane utilization model. This model is applicable to interchanges, adjacent intersections, and other intersections where prepositioning may occur. The form of this model, including two empirical calibration constants, is:

$$U = \left[1 + \left(\frac{N-1}{2 v'} \right) (1 - b_0) + \frac{3 N}{4} \sqrt{\frac{N-1}{2 v'}} b_0 \right] (1 - I_p) + \left[\frac{\text{Max}(v'_{dl}, v'_{dr}) N}{v'} b_1 \right] I_p \quad (\text{C-38})$$

where:

I_p = indicator variable (1.0 if $\text{Max}(v'_{dl}, v'_{dr})/v' > 1/N$, 0.0 otherwise); and
 b_0, b_1 = calibration coefficients.

C.4.2 Model Calibration

The lane utilization model was calibrated using data for 97 five-minute intervals. The data in each interval represent the maximum lane flow rate and the lane group flow rate that was observed each cycle during the five-minute interval, these cyclic values were then averaged to obtain the five-minute interval data. The 97 intervals represent the observation of 10,585 through vehicles in two, three, or four-lane lane groups.

The results of the model calibration are shown in Table C-9. As the statistics provided in this table indicate, the calibrated model provides a reasonably good fit to the data. The minimum precision of the average lane utilization factor estimate is about ± 0.01 , based on the root mean square error and the number of observations.

The R^2 of 0.18 is lower than values traditionally expected; however, it must be remembered that there is considerable random variability in the lane utilization factor. This variability stems from the fact that two random variables (i.e., v'_{max} and v') are being used in the computation of the lane utilization factor. Thus, the variability in this factor represents the combined variability of the two underlying random variables. It should be noted that a regression analysis using Equations C-33 and C-36 to predict the maximum observed lane volume v'_{max} yielded an R^2 of 0.95.

Table C-9. Calibrated lane utilization model

Model Statistics		Value		
R^2 :		0.18		
Root Mean Square Error:		0.11		
Observations:		97		
Range of Model Variables				
Variable	Variable Definition	Units	Minimum	Maximum
U	Lane utilization factor	na	1.0	1.56
v'_{max}	Maximum lane flow rate in any lane	vpcpl	5.8	40.3
N	Number of lanes in the lane group	na	2	4
v'_{dl}	No. of vehicles turning left downstream	vpc	0	23.3
v'_{dr}	No. of vehicles turning right downstream	vpc	0	16.1
v'	Demand flow rate for the lane group	vpc	12.4	69.7
Calibrated Coefficient Values				
Coeff.	Coefficient Definition	Value	Std. Dev.	t-statistic
b_0	Portion of drivers not evenly distributed	0.577	0.044	13.1
b_1	Effect of additional non-turning drivers	1.05	0.030	35.0

Two observations can be made based on the parameter coefficient values. First, the b_0 coefficient in the calibrated model indicates that about 58 percent of drivers at the sites studied do

not attempt to evenly distribute themselves among the available lanes. Second, the fact that the b_1 coefficient exceeds 1.0 suggests that additional "non-turning" drivers are choosing the same lane as the prepositioning drivers. Specifically, lane groups that experience prepositioning typically experience a five percent increase in the number of vehicles in the "maximum lane" due to non-turning drivers.

The calibrated model's quality of fit to the data is also shown in Figure C-28. The data points shown in this figure represent the average of three observations each. This averaging was done to facilitate an examination of model fit over the range of the measured data. This type of assessment can be made difficult when individual observations overlap and appear as only one data point. This problem is particularly evident when comparing predicted and measured values where the tendency is for the many data points in agreement to overlap and appear as only one observation.

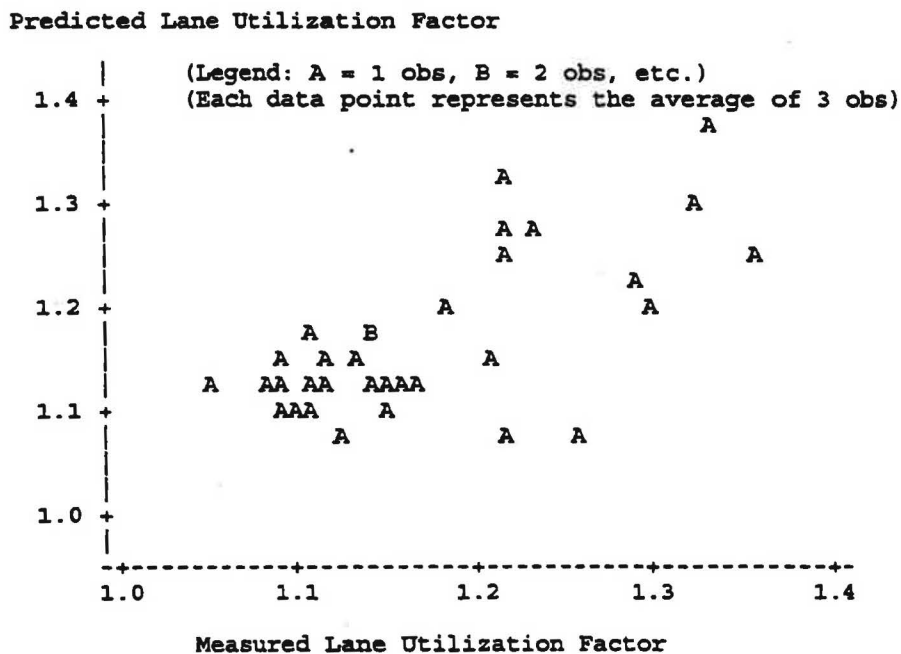


Figure C-28. Comparison of predicted and measured lane utilization factors.

C.4.3 Sensitivity Analysis

The calibrated lane utilization model can be used to examine the relationship between lane utilization, lane group flow rate, and number-of-lanes. Figure C-29 illustrates these relationships for the Random Behavior component of the model (i.e., no prepositioning). As expected, the lane utilization factors increase with number of lanes and decrease with increasing volume. It should be noted that the predicted lane utilization factors exceed the values recommended by the HCM (1) (i.e., 1.05 at two through lanes, 1.10 at three through lanes).

C.4.2 Model Calibration

The lane utilization model was calibrated using data for 97 five-minute intervals. The data in each interval represent the maximum lane flow rate and the lane group flow rate that was observed each cycle during the five-minute interval, these cyclic values were then averaged to obtain the five-minute interval data. The 97 intervals represent the observation of 10,585 through vehicles in two, three, or four-lane lane groups.

The results of the model calibration are shown in Table C-9. As the statistics provided in this table indicate, the calibrated model provides a reasonably good fit to the data. The minimum precision of the average lane utilization factor estimate is about ± 0.01 , based on the root mean square error and the number of observations.

The R^2 of 0.18 is lower than values traditionally expected; however, it must be remembered that there is considerable random variability in the lane utilization factor. This variability stems from the fact that two random variables (i.e., v'_{max} and v') are being used in the computation of the lane utilization factor. Thus, the variability in this factor represents the combined variability of the two underlying random variables. It should be noted that a regression analysis using Equations C-33 and C-36 to predict the maximum observed lane volume v'_{max} yielded an R^2 of 0.95.

Table C-9. Calibrated lane utilization model

Model Statistics		Value		
R^2 :		0.18		
Root Mean Square Error:		0.11		
Observations:		97		
Range of Model Variables				
Variable	Variable Definition	Units	Minimum	Maximum
U	Lane utilization factor	na	1.0	1.56
v'_{max}	Maximum lane flow rate in any lane	vpcpl	5.8	40.3
N	Number of lanes in the lane group	na	2	4
v'_a	No. of vehicles turning left downstream	vpc	0	23.3
v'_d	No. of vehicles turning right downstream	vpc	0	16.1
v'	Demand flow rate for the lane group	vpc	12.4	69.7
Calibrated Coefficient Values				
Coeff.	Coefficient Definition	Value	Std. Dev.	t-statistic
b_o	Portion of drivers not evenly distributed	0.577	0.044	13.1
b_1	Effect of additional non-turning drivers	1.05	0.030	35.0

Two observations can be made based on the parameter coefficient values. First, the b_0 coefficient in the calibrated model indicates that about 58 percent of drivers at the sites studied do

not attempt to evenly distribute themselves among the available lanes. Second, the fact that the b_1 coefficient exceeds 1.0 suggests that additional "non-turning" drivers are choosing the same lane as the prepositioning drivers. Specifically, lane groups that experience prepositioning typically experience a five percent increase in the number of vehicles in the "maximum lane" due to non-turning drivers.

The calibrated model's quality of fit to the data is also shown in Figure C-28. The data points shown in this figure represent the average of three observations each. This averaging was done to facilitate an examination of model fit over the range of the measured data. This type of assessment can be made difficult when individual observations overlap and appear as only one data point. This problem is particularly evident when comparing predicted and measured values where the tendency is for the many data points in agreement to overlap and appear as only one observation.

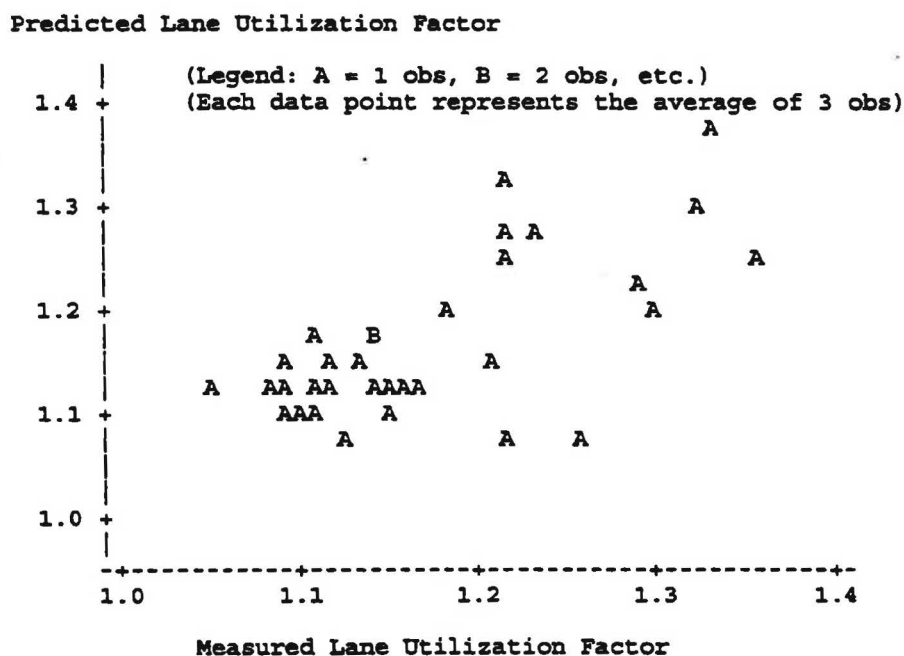


Figure C-28. Comparison of predicted and measured lane utilization factors.

C.4.3 Sensitivity Analysis

The calibrated lane utilization model can be used to examine the relationship between lane utilization, lane group flow rate, and number-of-lanes. Figure C-29 illustrates these relationships for the Random Behavior component of the model (i.e., no prepositioning). As expected, the lane utilization factors increase with number of lanes and decrease with increasing volume. It should be noted that the predicted lane utilization factors exceed the values recommended by the HCM (1) (i.e., 1.05 at two through lanes, 1.10 at three through lanes).

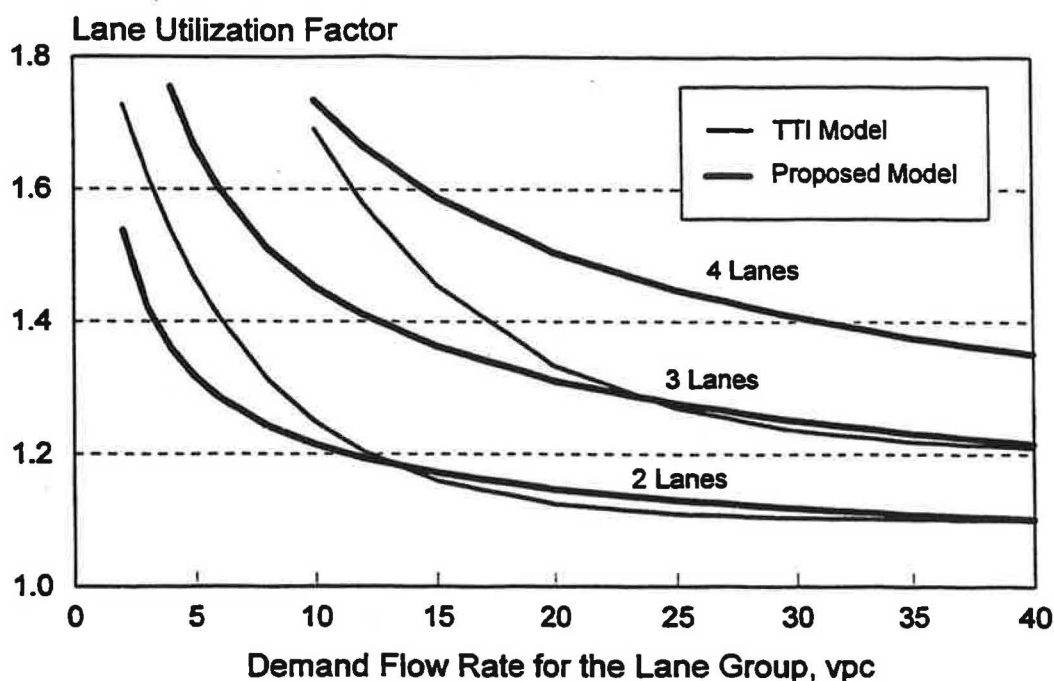


Figure C-29. Effect of flow rate and number of lanes on the lane utilization factor.

Also shown in Figure C-29 are the lane utilization factors predicted by a model developed by Fambro *et al* (7, 8). This model is coined the "TTI Model" in reference to the authors' affiliation. It was calibrated to ten traffic movements at nine signalized intersections. In general, the TTI Model predictions compare favorably with those of the calibrated lane utilization model; however, the agreement is best at the higher flow rates. This agreement is partly due to the fact that both data bases had the majority of their observations in this higher range of flow rates. This agreement suggests that the calibrated lane utilization model may be applicable to all signalized intersections.

C.5 CAPACITY ANALYSIS

This section summarizes the models developed in the preceding sections in the context of a proposed procedure for capacity analysis. These models predict capacity-related traffic characteristics for a range of conditions common to interchanges and closely-spaced signalized intersections. Initially, the models for predicting these characteristics are presented in a summary format that is consistent with the method of presentation used in the HCM (1). Then, two topics that are related to capacity analysis are discussed. These two topics include lane utilization and queue spillback. Next, the ability of the proposed models to predict capacity is demonstrated. Finally, a sensitivity analysis is conducted to illustrate the effect of saturation flow rate and lost time on capacity.

C.5.1 Introduction

The capacity of lane group represents the maximum number of vehicles that can be served by the group's traffic lanes during the time allocated to it within the signal cycle. A lane group flowing at capacity is typically characterized by the continuous discharge of queued traffic for the duration of the signal phase that controls it. To this extent, lane group capacity is dependent on the discharge characteristics of the departing traffic queue. These characteristics include the time lost at the start of the phase due to driver reaction time and acceleration, the saturation flow rate, and the time lost at the end of the phase due to a necessary change interval. The following equation is commonly used to compute the capacity of a lane group:

$$c' = \frac{g s}{3,600} \quad (C-39)$$

with,

$$g = G + Y + R_c - (l_s + l_e) \cap CP \quad (C-40)$$

where:

- c' = capacity of the lane group, vpc;
- g = effective green time where platoon motion (flow) can occur, sec;
- s = saturation flow rate for the lane group under prevailing conditions, vphg;
- G = green signal interval, sec;
- Y = yellow interval, sec;
- R_c = red clearance interval, sec;
- l_s = start-up lost time, sec;
- l_e = clearance lost time, sec; and
- CP = clear period during cycle/phase when subject flow is unblocked (see Appendix D), sec.

Many of the variables defined in Equations C-39 and C-40 represent the basic set of capacity characteristics. The proposed models for predicting these characteristics (i.e., saturation flow rate, start-up lost time, and clearance lost time) are described in the next three sections.

C.5.2 Saturation Flow Rate

The saturation flow rate for a lane group can be predicted using the following equation:

$$s = N \times s_0 \times f_w \times f_{HV} \times f_g \times f_p \times f_{bb} \times f_{RT} \times f_{LT} \times f_D \times f_v \quad (C-41)$$

where:

- s = saturation flow rate for the lane group under prevailing conditions, vphg;
- s_0 = saturation flow rate per lane under ideal conditions, pcphgpl;
- f_w = adjustment factor for lane width;
- f_{HV} = adjustment factor for heavy vehicles;

f_g = adjustment factor for approach grade;
 f_p = adjustment factor for parking;
 f_{bb} = adjustment factor for bus blockage;
 f_{RT} = adjustment factor for right-turns in the lane group;
 f_{LT} = adjustment factor for left-turns in the lane group;
 f_D = adjustment factor for distance to downstream queue at green onset; and
 f_v = adjustment factor for volume level (i.e., traffic pressure).

The first seven adjustment factors listed in this equation represent those described in the HCM (1, Chapter 9). The last two factors were developed for this research and are specifically applicable to interchanges and closely-spaced signalized intersections.

A third adjustment factor f_R was developed for this research that quantifies the effect of turn radius on the saturation flow rate of a left or right-turn movement. This factor is introduced by revising the adjustment factors for protected turn movements provided in the HCM. These revised right and left-turn adjustment factors are:

$$f_{RT} = \frac{1}{1 + P_{RT} \left(\frac{1}{f_R} - 1 \right)} \quad : \text{Protected, Shared Right-Turn Lane} \quad (C-42)$$

$$f_{RT} = f_R \quad : \text{Protected, Exclusive Right-Turn Lane}$$

$$f_{LT} = \frac{1}{1 + P_{LT} \left(\frac{1}{f_R} - 1 \right)} \quad : \text{Protected, Shared Left-Turn Lane} \quad (C-43)$$

$$f_{LT} = f_R \quad : \text{Protected, Exclusive Left-Turn Lane}$$

where:

f_R = adjustment factor for the radius of the travel path (based on the radius of the applicable left or right-turn movement);
 P_{RT} = portion of right-turns in the lane group; and
 P_{LT} = portion of left-turns in the lane group.

Adjustment factors for permissive-only and protected-permissive phasing can also be developed using Equations C-42 and C-43 as a basis.

The ideal saturation flow rate s_0 represents the saturation flow rate of a lane that is not affected by any external environmental factors (e.g., grade), non-passenger-car vehicles (e.g., trucks), and constrained geometrics (e.g., less than 3.6-meter lane widths, curved travel path). In this regard, the saturation flow rate would be equal to the ideal rate when all factor effects are optimum for

efficient traffic flow and the corresponding adjustment factors are equal to 1.0. Based on this definition, it was determined that ideal conditions were represented by an infinite distance-to-queue, on-spillback conditions, a tangent alignment (i.e., an infinite radius), and a traffic pressure representing that typically experienced during peak traffic periods. Under these conditions, 1,000 pcphgpl is recommended as the ideal saturation flow rate for all traffic movements.

The following sections describe the individual adjustment factors developed for this research. One potential factor that was not included in Equation C-41 is that of phase duration. There is some evidence that drivers adopt larger saturation flow rates for phases of short duration than for those of long duration. The evidence was strongest for the left-turn movements and when the g/C ratio was less than 0.27. This effect was also found in the through movements studied, however, it was much smaller in magnitude and not statistically significant (due in part to the fact that the g/C ratios for most through movements were above 0.30). Nevertheless, the data suggest that saturation flow rates can increase by 5 to 10 percent when the g/C ratio for the subject movement is less than 0.20. However, more research will be needed to verify the significance of this trend and its magnitude before it can be recommended for application.

Distance-to-Queue Adjustment Factor. The distance-to-queue adjustment factor f_D accounts for the adverse effect of downstream queues on the discharge rate of an upstream traffic movement. In general, the saturation flow rate is low for movements that have a downstream queue relatively near at the start of the phase; it is high for movements that are not faced with a downstream queue at the start of the phase. Thus, the distance-to-queue adjustment factor is based on the distance to the back of the downstream queue at the start of the subject phase.

The variable "distance to queue" is measured from the subject (or upstream) movement stop line to the "effective" back of queue. The effective back of queue represents the location of the back of queue if all vehicles on the downstream street segment (moving or stopped) at the start of the phase were joined into a stopped queue. If there are no moving vehicles at the start of the phase, then the effective and actual distance to queue are the same. If there are no vehicles on the downstream segment at the start of the phase, then the effective distance to queue would equal the distance to the through movement stop line at the downstream intersection. The distance-to-queue is computed as:

$$D = L - \frac{n_s}{N_d} L_v \quad (C-44)$$

with,

$$L_v = (1 - P_{HV}) L_{pc} + P_{HV} L_{HV} \quad (C-45)$$

where:

D = effective distance to the back of downstream queue (or stop line if no queue) at the start of the subject (or upstream) phase, m;

L = distance between the subject and downstream intersection stop lines (i.e., link length), m;
 n_s = number of vehicles on the downstream street segment (moving or queued) at the start of the subject phase, veh;
 N_d = number of through lanes on the downstream segment, lanes;
 L_v = average lane length occupied by a queued vehicle, m/veh;
 L_{pc} = lane length occupied by a queued passenger car (= 7.0 m/pc), m/pc;
 L_{HV} = lane length occupied by a queued heavy vehicle (= 13 m/veh), m/veh; and
 P_{HV} = portion of heavy vehicles in the traffic stream.

The lane length occupied by a queued passenger car vehicle was measured for this research. The summary data reported in Table B-11 indicate that the average values is 7.0 m/pc. The lane length occupied by a queued heavy vehicle can vary, based on the types of heavy vehicles in the subject traffic stream. An approximate value that can be used when field measurements are not available is 13 m/veh. This value is representative of most buses, recreational vehicles, and inter-city trucks.

The effect of distance-to-queue is also dependent on whether spillback occurs during the subject phase. Spillback is characterized by the backward propagation of a downstream queue into the upstream intersection such that one or more of the upstream intersection movements are effectively blocked from discharging during some or all of their respective signal phase. If spillback occurs during the phase, the saturation flow rate prior to the occurrence of the spillback is much lower than it would be if there were no spillback. Thus, the magnitude of the adjustment to the saturation flow rate is dependent on whether spillback occurs during the subject phase. A procedure for determining if queue spillback occurs is provided in a later section.

The distance-to-queue adjustment factor is tabulated in Table C-10. It can also be computed using the following equation:

$$f_D = \begin{cases} \frac{1}{1 + \frac{8.13}{D}} & : \text{ no spillback} \\ \frac{1}{1 + \frac{21.8}{D}} & : \text{ with spillback} \end{cases} \quad (C-46)$$

where:

f_D = adjustment factor for distance to downstream queue at green onset; and
 D = effective distance to the back of downstream queue (or stop line if no queue) at the start of the subject (or upstream) phase, m.

Table C-10. Adjustment factor for distance-to-queue (f_D)

Distance to Back of Queue at Start of Subject Phase, m	Spillback Condition	
	No Spillback	With Spillback
15	0.649	0.408
30	0.787	0.579
60	0.881	0.734
120	0.937	0.846
180	0.957	0.892
240	0.967	0.917
300	0.974	0.932
360	0.978	0.943

Turn Radius Adjustment Factor. Traffic movements that discharge along a curved travel path do so at rates lower than those of through movements. The effect of travel path radius is tabulated in Table C-11. It can also be computed using the following equation:

$$f_R = \frac{1}{1 + \frac{1.71}{R}} \quad (\text{C-47})$$

where:

f_R = adjustment factor for the radius of the travel path; and

R = radius of curvature of the left-turn travel path (at center of path), m.

Table C-11. Adjustment factor for turn radius (f_R)

Radius of the Travel Path, m	Movement Type	
	Left & Right-Turn	Through
8	0.824	1.000
15	0.898	1.000
30	0.946	1.000
45	0.963	1.000
60	0.972	1.000
75	0.978	1.000
90	0.981	1.000
105	0.984	1.000

This factor was calibrated for left-turn traffic movements; however, a comparison of this adjustment factor to that developed by others for right-turn movements indicates close agreement. Therefore, this factor is recommended for use with both left and right-turn movements.

Traffic Pressure Adjustment Factor. Saturation flow rates are generally found to be higher during peak traffic demand periods than during off-peak periods. This trend is explained by the effect of "traffic pressure." In general, it is believed that traffic pressure reflects the presence of a large number of aggressive drivers (e.g., commuters) during high-volume traffic conditions. They demonstrate this aggressive behavior by accepting shorter headways during queue discharge than would less-aggressive drivers. As these aggressive drivers are typically traveling during the morning and evening peak traffic periods, they are found to represent a significant portion of the traffic demand associated with these periods.

The effect of traffic pressure was found to vary by traffic movement. Specifically, the left-turn movements tended to be more affected by pressure as their saturation flow rates varied more widely than those of the through movements for similar conditions. It is possible that this difference between movements stems from the longer delays typically associated with left-turn movements.

Based on the preceding definition, it is logical that the effect of traffic pressure is strongly correlated with the demand flow rate in the subject lane group. Thus, the effect of traffic pressure, as represented by traffic volume, is tabulated in Table C-12 for each movement type. It can also be computed using the following equation:

$$f_v = \begin{cases} \frac{1}{1.07 - 0.00672 v_l'} & : \text{left-turn} \\ \frac{1}{1.07 - 0.00486 v_l'} & : \text{through or right-turn} \end{cases} \quad (\text{C-48})$$

where:

f_v = adjustment factor for volume level (i.e., traffic pressure); and
 v_l' = demand flow rate per lane (i.e., traffic pressure), vpcpl.

As noted previously, ideal conditions were defined to include a traffic pressure effect representative of peak traffic periods. In this regard, traffic pressures under ideal conditions were defined as 10 and 15 vpcpl for the left-turn and through movements, respectively. These flow rates are conservative in their representation of higher volume conditions as they exceed 80 to 90 percent of all traffic demands that were observed at the field study sites. One consequence of this approach to defining ideal conditions is that it is possible for the traffic pressure adjustment factor to have values above 1.0 when the flow rate is extremely high.

Table C-12. Adjustment factor for volume level (i.e., traffic pressure) (f_v)

Traffic volume, vpcpl	Movement Type	
	Left-Turn	Through & Right-Turn
3.0	0.953	0.947
6.0	0.971	0.961
9.0	0.991	0.974
12.0	1.011	0.988
15.0	1.032	1.003
18.0	1.054	1.018
21.0	1.077	1.033
24.0	1.100	1.049

C.5.3 Start-up Lost Time

The start-up lost time associated with a discharging traffic queue varies with its saturation flow rate. More specifically, start-up lost time increases with saturation flow rate because it takes more time for the discharging queue to attain the higher speed associated with the higher saturation flow rate. The recommended start-up lost times corresponding to a range of saturation flow rates are provided in Table C-13. These values can also be computed using the following equation:

$$l_s = -4.54 + 0.00368 s_f \geq 0.0 \quad (\text{C-49})$$

where:

l_s = start-up lost time, sec; and

s_f = saturation flow rate per lane under prevailing conditions ($= s/N$), vphgpl.

The recommended start-up lost times are applicable to left, through, and right-turn movements.

Table C-13. Start-up lost time (l_s)

Saturation Flow Rate, vphgpl	Start-up Lost Time, sec
1,400	0.61
1,500	0.98
1,600	1.35
1,700	1.71
1,800	2.08
1,900	2.45
2,000	2.82
2,100	3.18

C.5.4 Clearance Lost Time

The time lost at the end of the phase represents the portion of the change interval that is unavailable for traffic service. This time is intended to provide a small time separation between the traffic movements associated with successive signal phases and, thereby, promote a safe change in right-of-way allocation. Clearance lost time can be computed as:

$$l_c = Y + R_c - g_Y \quad (C-50)$$

where:

- l_c = clearance lost time, sec;
- Y = yellow interval, sec;
- R_c = red clearance interval, sec; and
- g_Y = effective green extension into the yellow interval, sec.

Green extension was found to be dependent on the volume-to-capacity ratio of the subject movement and the approach speed limit. Specifically, at high volume-to-capacity ratios, drivers tend to use more of the yellow than at lower ratios. This trend suggests that drivers are more likely to enter the intersection on the yellow when demands (and corresponding delays) are high. In addition, drivers on higher speed approaches were also found to use slightly more of the yellow interval than those on lower speed approaches. This trend is likely due to the undesirably high decelerations often associated with quick stops from higher speeds.

The amount of green extension for a given approach can be estimated as:

$$g_Y = 1.48 + 0.014 SL + 6.40 (X - 0.88) I_x \quad (C-51)$$

where:

- g_Y = effective green extension into the yellow interval, sec;
- SL = approach speed limit, km/h;
- X = volume-to-capacity ratio for the lane group; and
- I_x = indicator variable (1.0 if $X > 0.88$, 0.0 otherwise).

For speeds in the range of 64 to 76 km/h and volume-to-capacity ratios of 0.88 or less, the average green extension is 2.5 seconds. This average value is recommended for use in most capacity analysis. More appropriate values can be computed from Equation C-51 when the speeds are outside the range of 64 to 76 km/h or when X for the analysis period exceeds 0.88.

C.5.5 Lane Utilization

Drivers do not distribute themselves evenly among the traffic lanes available to a lane group. As a consequence, the lane of highest demand has a higher volume-to-capacity ratio and the possibility of more delay than the other lanes in the group. The HCM (1) recognizes this phenomena

and offers the use of a lane utilization factor U to adjust the lane group volume such that it represents the flow rate in the lane of highest demand. The adjusted lane group volume can be computed as:

$$v' = v_g' U \quad (C-52)$$

where:

- v' = demand flow rate for the lane group, vpc;
- v_g' = unadjusted demand flow rate for the lane group, vpc; and
- U = lane utilization factor for the lane group.

There are two reasons for an uneven distribution of traffic volume among the available lanes. One reason is the inherent randomness in the number of vehicles in each lane. While drivers may prefer the lesser-used lanes because of the potential for reduced travel time, they are not always successful in getting to them for a wide variety of reasons (e.g., unable to ascertain which lane is truly lowest in volume, lane change prevented by vehicle in adjacent lane, driver is not motivated to change lanes, etc.). Thus, it is almost a certainty that one lane of a multi-lane lane group will have more vehicles in it than the other lanes, during any given cycle.

The second reason for an uneven distribution of traffic volume among the available lanes is driver desire to "preposition" for a turn maneuver at a downstream intersection. This activity commonly occurs on the arterial cross street at interchange ramp terminals and at any associated closely-spaced intersections. This prepositioning causes drivers to concentrate in one lane at the upstream ramp terminal or intersection in anticipation of a turn onto the freeway at the downstream ramp terminal.

Random Behavior Model. The lane utilization factor for an intersection lane group that does not experience prepositioning (e.g., at an isolated intersection) is dependent on the degree of randomness in the group's collective lane-choice decisions. In addition, it is also based on the group's traffic volume on a "per cycle" basis. In general, lane utilization is more uneven for groups with low demands or a high degree of randomness or both. The recommended lane utilization factors for random lane-choice decisions are provided in Table C-14. These values can also be computed using the following equation:

$$U_r = 1 + 0.423 \left(\frac{N-1}{2 v'} \right) + 0.433 N \sqrt{\frac{N-1}{2 v'}} \quad (C-53)$$

where:

- U_r = lane utilization factor for random lane-choice decisions;
- v' = demand flow rate for the lane group, vpc; and
- N = number of lanes in the lane group.

The recommended lane utilization factors are applicable to left, through, and right-turn lane groups.

Table C-14. Lane utilization factors for lane groups with random lane choice (U_r)

Lane Group Flow Rate, vpc	Number of Lanes in the Lane Group			
	1	2	3	4
5.0	1.00	1.32	1.67	2.08
10.0	1.00	1.22	1.45	1.74
15.0	1.00	1.17	1.36	1.59
20.0	1.00	1.15	1.31	1.51
25.0	1.00	1.13	1.28	1.45
30.0	1.00	1.12	1.25	1.41
35.0	1.00	1.11	1.23	1.38
40.0	1.00	1.10	1.22	1.35

Prepositioning Model. The lane utilization factor for interchanges and associated closely-spaced intersections can be strongly influenced by drivers prepositioning for downstream turns. The magnitude of the effect is largely a site-specific characteristic, depending on the number of vehicles turning left or right at the downstream intersection (or ramp terminal). If this information is available (such as from a turn movement count where downstream destination is also recorded), the lane utilization factor can be computed as:

$$U_p = 1.05 \frac{\text{Max}(v'_{dl}, v'_{dr}) N}{v'} \quad (\text{C-54})$$

where:

U_p = lane utilization factor for prepositioning;

v'_{dl} = number of vehicles in the subject lane group that will be turning left at the downstream intersection, vpc;

v'_{dr} = number of vehicles in the subject lane group that will be turning right at the downstream intersection, vpc; and

$\text{Max}(v'_{dl}, v'_{dr})$ = larger of v'_{dl} and v'_{dr}

Lane Utilization Factor. The possibility of prepositioning must be evaluated to determine whether to use Equation C-53 or C-54 to estimate the lane utilization factor U . This possibility can be determined from the following test:

$$\begin{aligned} \frac{\text{Max}(v'_{dl}, v'_{dr})}{v'} &> \frac{1}{N} : \text{prepositioning} \\ \frac{\text{Max}(v'_{dl}, v'_{dr})}{v'} &\leq \frac{1}{N} : \text{no prepositioning} \end{aligned} \quad (\text{C-55})$$

Based on the outcome of this test, the lane utilization factor is computed as:

$$U = \begin{cases} U_p & : \text{prepositioning} \\ U_r & : \text{no prepositioning} \end{cases} \quad (\text{C-56})$$

The distance to the downstream intersection could effect the propensity of drivers to preposition. In recognition of this effect, the test equation is recommend only for intersections in interchange areas or other closely-spaced intersections where the inter-signal distances are less than 300 meters.

C.5.6 Predicting Queue Spillback

Two conditions must occur to precipitate spillback during a saturated flow state. First, the discharging stream from the subject traffic movement must arrive at the back of a stopped downstream queue. Second, the subject phase must be sufficiently long as to permit the discharge of enough vehicles to fill the available downstream distance under jam density conditions.

Occurrence of the first condition can be predicted by comparison of the actual and ideal signal offset between the subject phase and that of the downstream through movement. In general, if the difference between the actual and ideal offsets is positive, then there is a possibility of spillback. A negative difference indicates that spillback is unlikely. This offset difference can be computed as:

$$\Delta = \text{off}_a - \text{off}_i \quad (\text{C-57})$$

with,

$$\text{off}_i = \frac{L}{u_s} - \left(\frac{n_s}{N_d} \frac{3,600}{s_n} + h_c \right) \quad (\text{C-58})$$

where:

- Δ = difference between the actual and ideal signal offset, sec;
- off_a = actual offset between the subject phase and that of the downstream through movement (phase start time downstream minus phase start time upstream), sec;
- off_i = ideal offset to ensure progression without speed disruption, sec;
- u_s = speed at saturation flow, m/s (say, 14 m/s);
- h_c = clearance headway between the last queued vehicle and the first arriving vehicle, sec (say, 2.0 sec);
- s_n = saturation flow rate for the subject lane under prevailing conditions assuming the "no-spillback" condition, vphgpl;

- L = distance between the subject and downstream intersection stop lines (i.e., link length), m;
 n_s = number of vehicles on the downstream street segment (moving or queued) at the start of the subject phase, veh; and
 N_d = number of through lanes on the downstream segment, lanes.

Occurrence of the second condition is dependent on the first condition. If the first condition occurs (i.e., the offset difference is positive), spillback will occur if the phase is long enough to discharge enough vehicles such that the resulting queue exceeds that of the downstream street segment. If the phase is short, such that the discharged vehicles are insufficient in number to fill the downstream distance-to-queue, then spillback will not occur. The corresponding maximum green interval duration that, if exceeded, would lead to spillback is computed as:

$$G_{\max} = \frac{D}{L_v} \left(\frac{3,600}{s_w} \right) + l_s \quad (\text{C-59})$$

where:

- G_{\max} = maximum green signal interval duration for the subject (or upstream) phase that is allowable without spillback during saturated flows, sec;
 D = effective distance to the back of downstream queue (or stop line if no queue) at the start of the subject (or upstream) phase, m;
 s_w = saturation flow rate for the subject lane under prevailing conditions assuming the "with-spillback" condition, vphgpl;
 L_v = average lane length occupied by a queued vehicle (see Equation C-45), m/veh; and
 l_s = start-up lost time, sec.

Equations C-58 and C-59 were used to develop two figures that can be used to predict the likely occurrence of spillback for the subject phase. Specifically, the relationship between distance-to-queue D and ideal offset off_i (i.e., Equation C-58) is shown in Figure C-30. The relationship between distance-to-queue D and the maximum green interval duration G_{\max} (i.e., Equation C-59) is shown in Figure C-31.

As Figure C-30 indicates, the ideal offset is often negative for most intersection pairs when there is a downstream queue. A negative value indicates that the downstream through phase must start prior to the subject upstream phase to provide smooth traffic progression.

Figure C-30 can be used instead of Equation C-58 to determine whether the first condition is satisfied. Specifically, the actual offset is compared with the ideal offset for a given street segment length and distance-to-queue. If the intersection of the actual offset (entered from the y-axis) and the known distance-to-queue (entered from the x-axis) falls below the line corresponding to the segment length, then the difference in offset will be negative and spillback is not likely to occur. If the intersection occurs above the appropriate segment length line, then the difference in offset will be positive and spillback is likely to occur during saturated flow conditions. In this latter case, the second condition should be checked to verify whether spillback will occur.

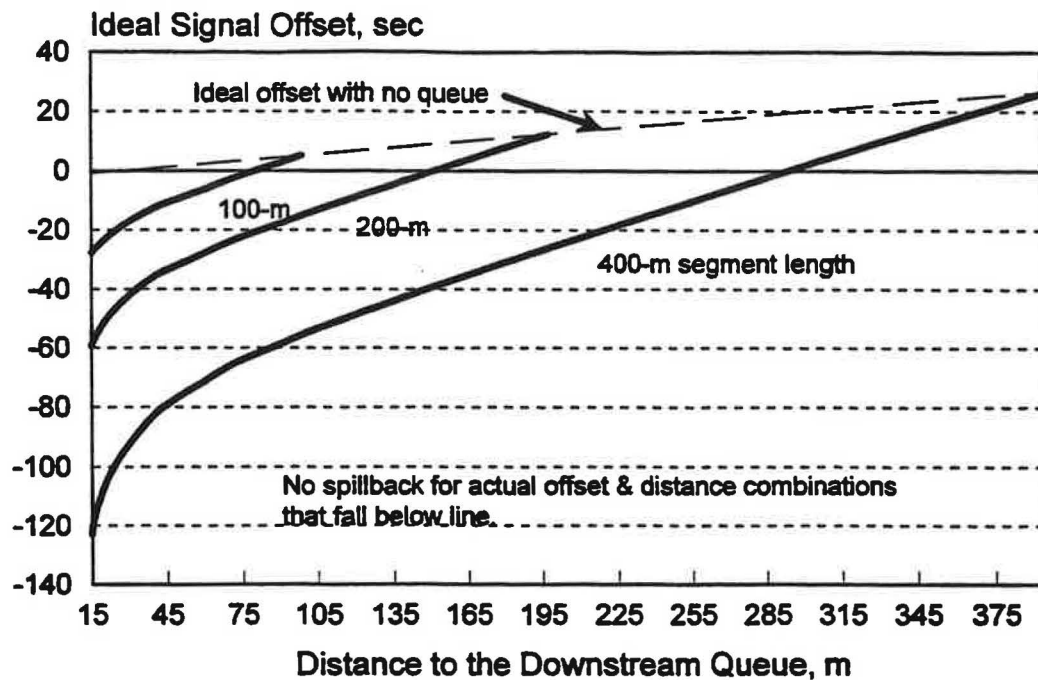


Figure C-30. Relationship between distance-to-queue and ideal signal offset.

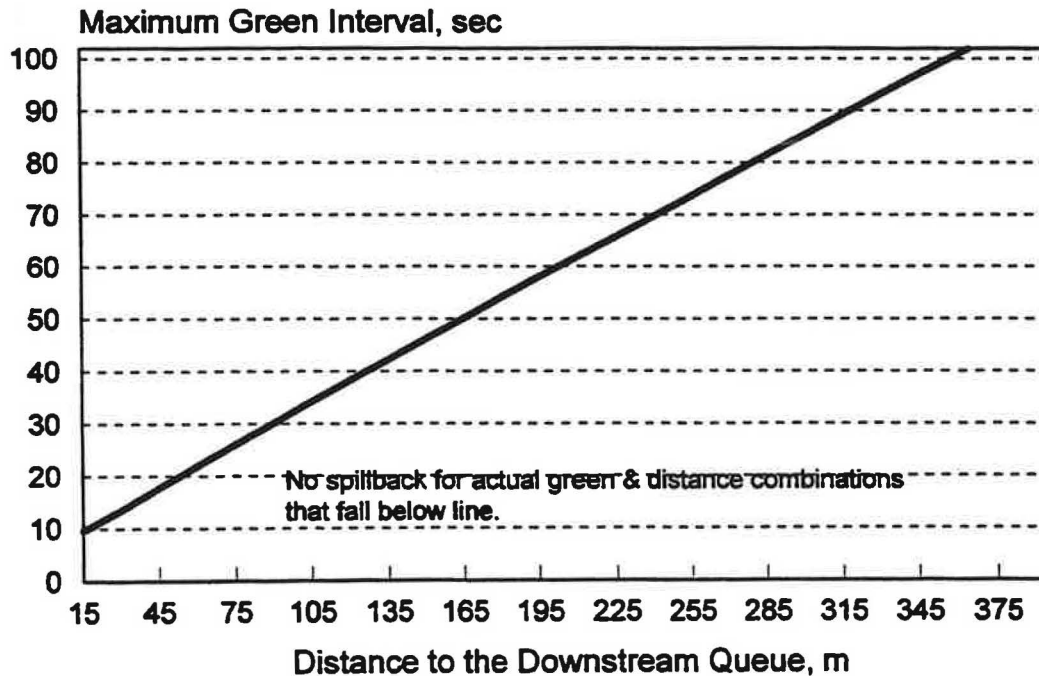


Figure C-31. Relationship between distance-to-queue and maximum upstream green.

Figure C-31 indicates whether the second condition occurs. Specifically, it indicates the maximum green interval duration that would not lead to spillback; green intervals in excess of this maximum would likely lead to spillback. In general, this figure indicates that the maximum green interval increases linearly with increasing distance-to-queue. In fact, a first-order estimate of the maximum green duration is about one second of green for every three meters of distance.

If the offset difference ($off_a - off_i$) is determined to be positive, then Figure C-31 can be consulted to determine if spillback will occur for saturated flow conditions and the known green interval duration. If the intersection of the actual green interval (entered from the y-axis) and the known distance-to-queue (entered from the x-axis) falls below the line, then the green interval is sufficiently short as to prevent spillback. If the intersection occurs above the trend line, then the green interval is sufficiently long as to precipitate spillback whenever traffic demands are high enough to fully utilize the signal phase. Therefore, when the intersection point is above the trend line and demands are expected to fully use the phase, the "with spillback" form of the distance-to-queue adjustment factor f_D should be used.

The following example is presented to illustrate the use of Figures C-30 and C-31 to determine if spillback is likely to occur under saturated flow conditions. Consider a 400-meter street segment bounded by signalized intersections. The average distance-to-queue to queue at the start of the upstream intersection through phase is 135 meters. This phase has 50 seconds of green. The offset to the downstream through movement phase is -10 seconds (i.e., the downstream through phase starts 10 seconds before the subject phase). Consultation with Figure C-30 indicates that an ideal offset for this distance-to-queue is about -45 seconds and, more importantly, that the actual offset-distance combination intersects above the "400-meter" line. This point of intersection indicates that the difference in offset is positive and that the first condition is satisfied. Therefore, spillback is possible, depending on the duration of upstream green.

Consultation with Figure C-31 checks the second condition. Specifically, it indicates that the maximum green interval duration is about 41 seconds (or about 45 seconds using the 1:3 rule) and, more importantly, that the actual green-distance combination intersects above the trend line indicating the strong likelihood of spillback during saturated flow conditions. This finding indicates that the "with spillback" form of the distance-to-queue adjustment factor should be used to compute the capacity of the phase based on the assumption that traffic demands are sufficiently high as to fully use the signal phase.

C.5.7 Predictive Ability of the Proposed Models

As a verification of the calibrated saturation flow rate and start-up lost time models, the predicted and measured discharge times of several last-in-queue vehicles were compared. The predicted discharge time was computed using the following equation:

$$T_{d(n)} = JH + l_s \quad (C-60)$$

where:

$T_{d(J)}$ = discharge time of the J th queued vehicle, sec;
 J = last queue position to discharge;
 H = minimum discharge headway, sec/veh; and
 l_s = start-up lost time, sec.

For this verification, Equation C-60 was used (with the saturation flow rates predicted by Equation C-41 and the start-up lost times predicted by Equation C-49) to predict the discharge time of the last queued vehicle observed during several hundred signal cycles. The results of this evaluation are shown in Figure C-32.

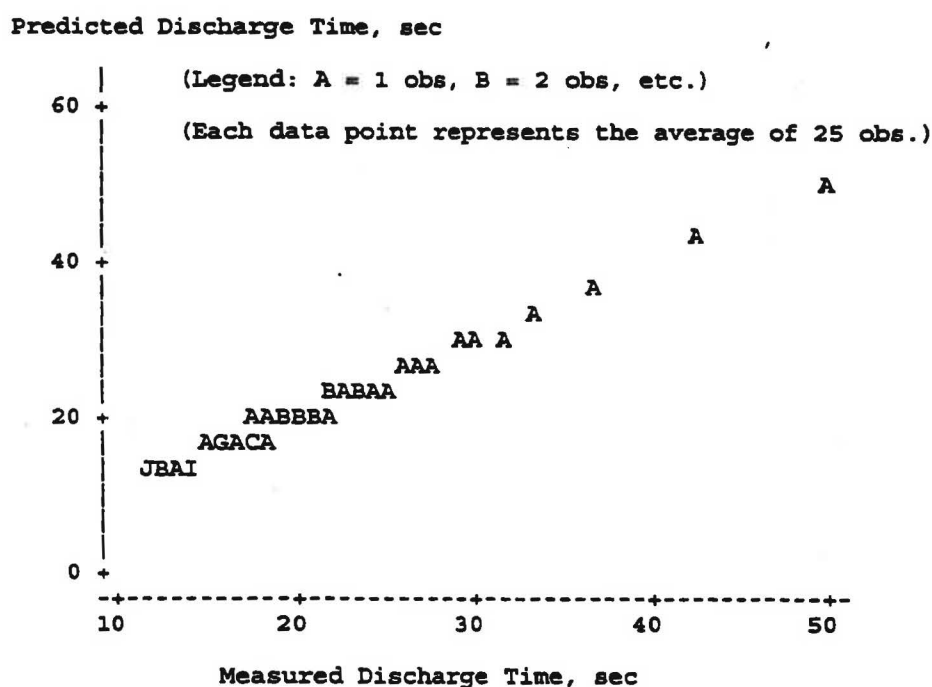


Figure C-32. Comparison of predicted and measured discharge times.

As the trends in Figure C-32 indicate, the recommended models are able to accurately predict the discharge time of the last queued vehicle. Moreover, the fit is equally good over the range of discharge times, indicating no time-based bias in the prediction process. The R^2 for the fit of the predicted to measured discharge times is 0.95.

It should be noted that the database used for this evaluation was the same as that used for the model calibration; thus, this evaluation is a model verification rather than a model validation. Nevertheless, it still provides strong evidence of the accuracy of the recommended models. It should be noted that the data points shown in Figure C-32 represent the average of 25 observations each. This averaging was undertaken to facilitate the graphical presentation of the model's trend-wise fit to the data; the individual data points were used in the assessment of model fit.

C.5.8 Effect of Saturation Flow Rate and Lost Time on Capacity

The ideal saturation flow rate recommended in Chapter 9 the 1985 *Highway Capacity Manual* (9) was 1,800 pcphgpl. Chapter 2 of this manual recommended a start-up lost time of 2.0 seconds. More recently, the 1994 HCM (1) included a recommendation of 1,900 pcphgpl for the ideal saturation flow rate but made no recommendation for a specific start-up lost time value (other than to note that appropriate values for this lost time were in a range of 1.0 to 2.0 seconds). The ideal saturation flow rate recommended in this report is 2,000 pcphgpl which represents an increase beyond the current HCM value. However, the "cost" associated with this increase is an increase in the start-up lost time. Fortunately, the increase in start-up lost time is generally small and does not totally offset the gains in capacity obtained from the higher flow rates.

The relative impact of alternative combinations of saturation flow rate and start-up lost time on capacity are shown in Figure C-33. The capacity shown in this figure was computed using Equation C-39 with a 90-second cycle length and a green extension of 2.5 seconds. In general, there is an 11 percent increase in capacity implied by the increase in saturation flow rate from 1,800 to 2,000 pcphgpl, for the same 1.5-second start-up lost time. However, if a more accurate start-up lost time of 2.82 seconds is used with the 2,000 pcphgpl flow rate, then the increase is reduced by about 20 to 40 percent (corresponding to a true increase in capacity of only about 6 to 9 percent). The point to be made here is that start-up lost time is not a true "lost time;" rather, it is a "cost" of associated with accelerating to the speed at saturation. Hence, the higher the saturation flow rate, the higher the speed will be at saturation flow and the longer the acceleration (or start-up lost) time.

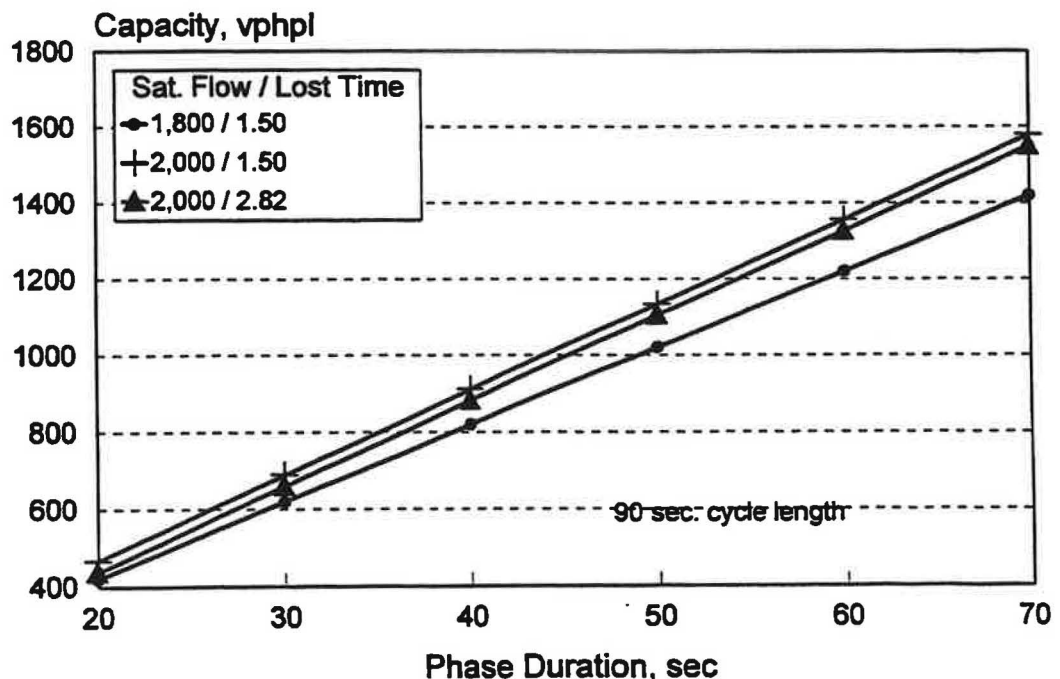


Figure C-33. Effect of saturation flow rate and start-up lost time on capacity.

C.5.9 Conclusions and Recommendations

This section summarizes the conclusion and recommendations resulting from the analysis of capacity characteristics associated with signalized interchange ramp terminals and closely-spaced intersections.

The ideal saturation flow rate recommended in this report is 2,000 pcphgpl. In the context of the factors studied for this research, this ideal flow rate applies to through traffic movements that have an infinite distance to the back of downstream queue, operate under non-spillback conditions, and have traffic volumes that are relatively high (reflecting those found during peak demand periods).

Based on this research, it is concluded that the distance to the downstream queue has a significant effect on the discharge rate of upstream traffic movements. This effect is amplified when the signal timing relationship between the two intersections allows queue spillback to occur. As the distance-to-queue variable is bounded to a maximum value equaling the length of the downstream street segment, the effect of distance-to-queue also includes the effect of spacing between interchange ramp terminals or between a ramp terminal and a closely-spaced intersection.

Turn radius has a significant effect on the discharge rate of a turn-related traffic movement. Saturation flow rates were much lower for turn movements with small radii than they were for turn movements with large radii. In the context of junction type (e.g., single-point urban diamond, at-grade intersection, etc.), the saturation flow rates for the left-turn movements at single-point urban interchanges are more nearly equal to those of through movements because of the large turn radii associated with this interchange type.

Traffic pressure, as quantified by traffic volume per cycle per lane, has a significant effect on saturation flow rate. The saturation flow rates of lower volume movements are much lower than those of higher volume movements.

Other factors were examined for their potential effect on saturation flow rate. These factors include: g/C ratio, junction type, downstream signal indication at the start of the upstream phase, and dual versus single left-turn lane. Of these factors, only g/C ratio was found to be correlated with saturation flow rate in a statistically significant manner. Specifically, the saturation flow rate for left-turn movements with low g/C ratios was found to be higher than the rates of similar movements with larger g/C ratios. This effect was also found in the through movements studied, however, it was much smaller in magnitude and not statistically significant. Therefore, it was determined that more research is needed to verify the significance of this trend and its magnitude before an adjustment factor for this effect can be recommended.

Start-up lost time was found to increase with saturation flow rate. This increase is due to the increased time it takes for the discharging queue to attain the higher speed associated with a higher saturation flow rate. Predicted start-up lost times range from 0.61 to 3.18 seconds for prevailing saturation flow rates of 1,400 to 2,100 pcphgpl.

The average amount of the yellow interval used by drivers is termed "green extension." This characteristic can be used to compute clearance lost time and the effective green time. The recommended green extension value is 2.5 seconds. Other values are possible if the approach speed is outside the range of 64 to 76 km/h or when the volume-to-capacity ratio for the analysis period is larger than 0.88. An equation is provided for these situations.

Lane use is almost always uneven (or unbalanced) in intersection lane groups. The degree of this imbalance is expressed in terms of the lane utilization factor. The lane utilization factor varies depending on the nature of drivers' lane-choice decisions (i.e., to minimize travel time or to preposition for a downstream turn). Lane utilization factors based on travel time minimization tend to be subject to randomness in the lane-choice decision process. The factors stemming from this process range from 1.1 to 2.0, depending on the number of lanes in the lane group and its corresponding traffic volume. Lane utilization factors based on driver desire to preposition can vary widely, depending on the volume of traffic that is prepositioning in the subject lane group.

Capacity is dependent on the prevailing saturation flow rate, start-up lost time, and clearance lost time. The equations provided in this appendix should be used to estimate the saturation flow rate, as affected by distance to downstream queue, turn radius, and traffic pressure. Start-up lost time is not a constant value; rather, it is dependent on the prevailing saturation flow rate. Thus, the equation provided in this appendix should be used to estimate the start-up lost time that corresponds to the prevailing saturation flow rate.

C.6 APPENDIX C REFERENCES

1. *Special Report 209: Highway Capacity Manual*. TRB, National Research Council, Washington, D.C. (1994).
2. Bonneson, J.A. "Study of Headway and Lost Time at Single-Point Urban Interchanges." In *Transportation Research Record 1365*. TRB, National Research Council, Washington, D.C. (1992) pp. 30 - 39.
3. Messer, C.J., Bonneson, J.A., Anderson, S.D., and McFarland, W.F. *NCHRP Report 345: Single Point Urban Interchange Design and Operations Analysis*. TRB, National Research Council, Washington, D.C. (1992).
4. Stokes, R.W., Messer, C.J., and Stover, V.G. "Saturation Flows of Exclusive Double Left-Turn Lanes." In *Transportation Research Record 1091*. TRB, National Research Council, Washington, D.C. (1986) pp. 86-95.
5. *SAS/STAT User's Guide*. Version 6, 4th ed. SAS Institute, Inc., Cary, North Carolina (1990).
6. Kimber, R.M., McDonald, M., and Hounsell, N.B. *TRRL Research Report RR67: The Prediction of Saturation Flows for Road Junctions Controlled by Traffic Signals*. Transport and Road Research Laboratory, Department of Transport, Berkshire, England (1986).

7. Fambro, D.B., Messer, C.J., and Andersen, D.A. "Estimation of Unprotected Left-Turn Capacity at Signalized Intersections." In *Transportation Research Record 644*. TRB, National Research Council, Washington, D.C. (1977) pp. 113-119.
8. Messer, C.J. and Fambro, D.B. "Critical Lane Analysis for Intersection Design." In *Transportation Research Record 644*. TRB, National Research Council, Washington, D.C. (1977) pp. 26-35.
9. *Special Report 209: Highway Capacity Manual*. TRB, National Research Council, Washington, D.C., (1985).

APPENDIX D

CLOSELY-SPACED INTERSECTION FLOW MODELS

D.1 OVERVIEW

The following appendix presents research that addresses traffic operations on the arterial street connections (links) between the traffic signals at signalized interchange ramp terminals, and also on the connecting links to adjacent intersections downstream from the interchange. Various traffic models are applied to the complex traffic operational conditions. The fundamental problems of queue spillback and flow blockage are addressed. Guidelines for identifying "closely-spaced" intersections are provided. A detailed computer-based algorithm, based on the Prosser-Dunn model, for assessing the impacts of queue spillback was developed and applied to a wide range of traffic conditions, including oversaturation.

D.2 LINK FLOW CONDITIONS

Traffic flow on a signalized cross arterial link traveling through an interchange is very complex due to several factors. The downstream traffic signal routinely interrupts the flow forming queues behind the signal which must be subsequently dissipated during the next cycle. The amount of queue formation depends on the amount of upstream turning traffic, the quality of signal progression, and the resulting total traffic demand that arrives on the movement.

Traffic operations on the link ultimately depend on whether the link's arrival demand volume for any movement exceeds the link's capacity to service it. If the link becomes oversaturated, due to demand exceeding output capacity, then the link will initially experience severe queue spillback and soon will become flooded with cars. The degree of resulting traffic flow and congestion depends almost entirely on the ability of the link to discharge vehicles downstream, otherwise total stoppage (grid lock) will occur. During heavily congested conditions, signal coordination primarily determines which feeding movements get to proceed, but usually not the overall quantity of flow on the link.

Assuming that traffic conditions on a link are undersaturated, then flows on the link may be assumed to cycle through a series of states shown in Figure D-1 where the arrival flows from the upstream movements proceed downstream and some are routinely stopped, accumulated in queue behind the downstream signal displaying red, and then subsequently serviced on the following green. Whereas arrival flows to the interchange ramps and some minor cross streets may be random at some average arrival flow rate, most arrival demands to the head of a link are, in reality, the output flow profile from an upstream signal modified by some platoon dispersion, depending on the distance traveled downstream. Platoons are usually not dispersed until they travel about two (2.0) minutes, so platooned flow along the crossing arterial is common for urban interchange operations.

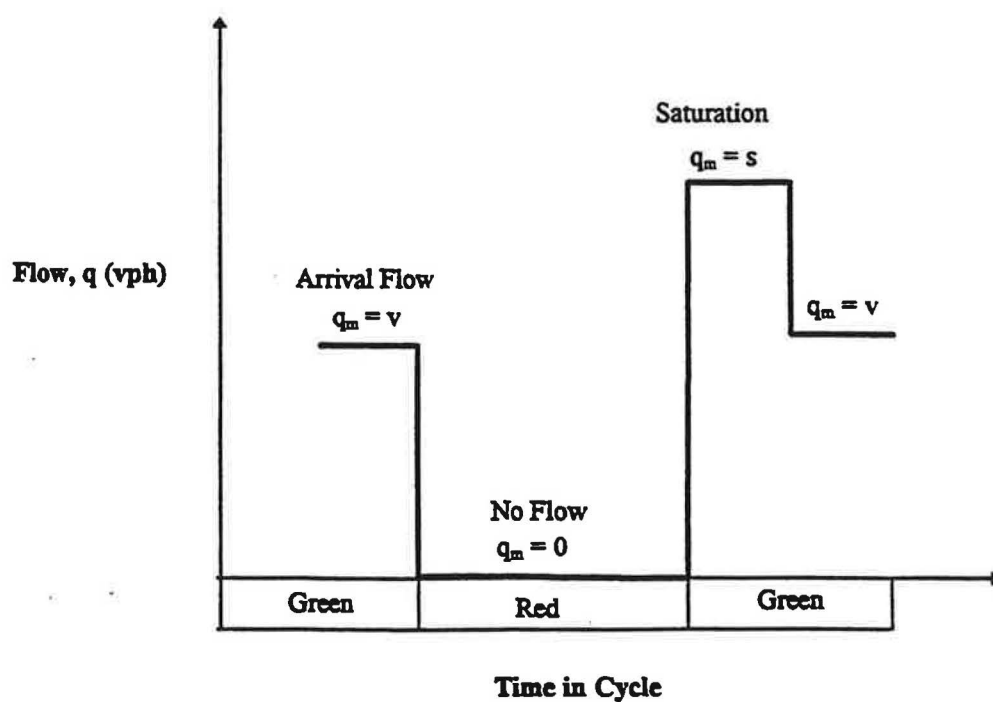
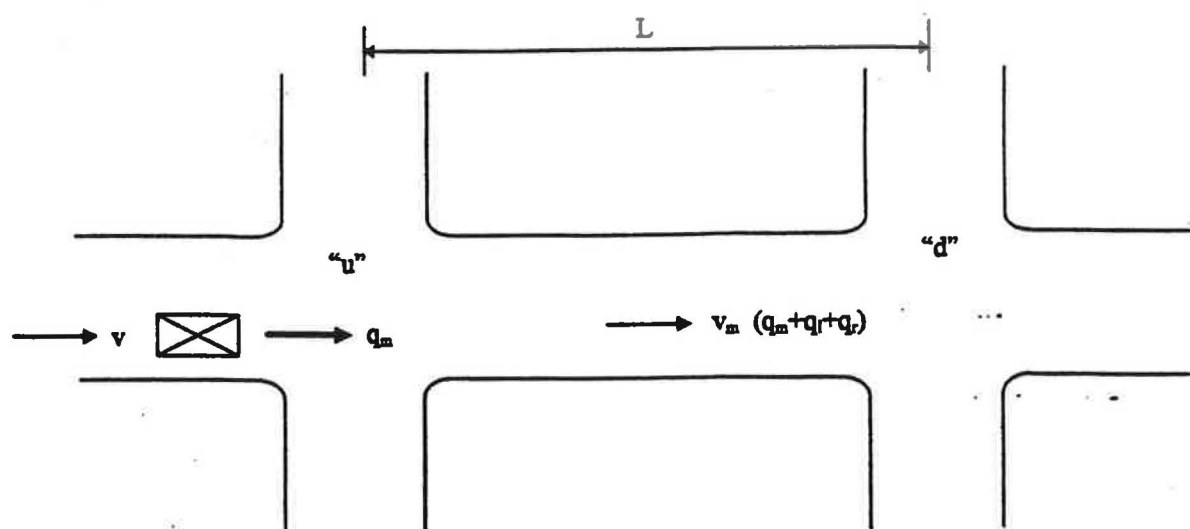


Figure D-1. Traffic Flow Conditions on a Link.

D.2.1 HCM Arrival Flow Profile

Arrival flow to a downstream closely-spaced arterial signal is inherently dependent on upstream traffic and signal conditions which may not be well known. Modeling assumptions may have to be made to provide tractable solutions. For Highway Capacity Manual (1) level of analysis, one assumption usually made is that the arrival flow to an arterial signal is composed of two component flows: one arriving during the downstream green and one during the red, or

$$v_a = v_g(g/C) + v_r(r/C) \quad (D-1)$$

where:

v_a	=	average arrival volume during cycle, vph;
v_g	=	arrival volume on green, vph;
v_r	=	arrival volume on red, vph;
g/C	=	green ratio; and
r/C	=	red ratio.

D.2.2 PASSER II Arrival Flow Profile

The next higher level of arrival flow profile found in traffic signal timing software is in the PASSER II - 90 software developed by TTI for arterial signal timing optimization (2). PASSER II assumes a two-flow model somewhat like the HCM except that one flow, the larger arrival flow, is defined as being the flow of the larger arriving platoon, in time, space, and rate of flow, and all other flows are combined into a single secondary flow region. This flow model is

$$v_a = v_p(p/C) + v_{np}(C-p)/C \quad (D-2)$$

where:

v_a	=	average arrival volume during cycle, vph;
v_p	=	maximum arrival volume in largest platoon, vph; and
v_{np}	=	average arrival volume during remainder of cycle, vph.

Flow profiles depend on the progression and are calculated based on the resulting time-space diagram for the arterial. Delay is calculated as being the area of the queue polygon resulting from the piece-wise integration of the input-output flow profiles over a representative cycle.

D.2.3 PASSER III Arrival Flow Profile

PASSER III is a computer model developed by TTI that is widely used to evaluate and optimized traffic operations at signalized diamond interchanges (3). Its arrival flow model provides the next higher level of sensitivity to variation in predictable arrival patterns. PASSER III models

three flow regions of the cycle, having two flows for each of the three upstream protected phases (i.e., A, B, C). The resulting downstream flow profiles are

$$v_a = v_{Ap} (Ap/C) + v_{Aa} (Aa/C) + v_{Bp} (Bp/C) + v_{Ba} (Ba/C) + v_{Cp} (Ca/C) + v_{Ca} (Ca/C) \quad (D-3)$$

where:

v_a	=	average arrival volume during cycle, vph;
v_{Ap}	=	platoon flow downstream from Phase A, vph;
v_{Aa}	=	average arrival flow to/from Phase A, vph;
v_{Bp}	=	platoon flow downstream from Phase B, vph;
v_{Bp}	=	average arrival flow to/from Phase B, vph;
v_{Cp}	=	platoon flow downstream from Phase C, vph; and
v_{Ca}	=	average arrival flow to/from Phase C, vph.

For diamond interchanges, arrival flows coming from an upstream Phase C signal are normally zero because this phase is the outbound left turn phase to the on-ramp.

D.2.4 TRANSYT 7F Arrival Flow Profile

An even higher-level arrival flow model is employed in TRANSYT 7F, another signal timing optimization and analysis program supported by Federal Highway Administration (4). A discretized output flow profile is developed for each upstream input flow which, when adjusted for platoon dispersion, becomes a component to the arrival flow profile of interest. The individual arrival flow profiles are summed to generate the total arrival flow profile. As in the above programs, the upstream phasing sequence must be defined together with all elements of the time-space diagram, signal progression, and platoon dispersion. The discretization time slice in TRANSYT 7F can be as small as 1/60 of a cycle. Thus, there can be up to sixty discrete arrival flows calculated for a movement, which are the sums of all upstream feeding flows for that time slice $(T+K)$, or

$$v_d(T+K_d) = \sum v_m(T+K_m) \quad (D-4)$$

where T is the travel time between the intersections. This high level of detail can only be achieved by automating the analytics to make the volume estimation process practically feasible.

D.2.5 Arrival Flow Model Recommended

The selection of arrival flow modeling sets the standard for the level of precision for all the level of service modeling to follow. It is presumed that the average flows that would occur during the cycle, peak period, and within the hour of interest can be estimated with sufficient accuracy. For external approaches or isolated approaches that do not have a signal within two minutes travel time to the subject approach, the assumption of random arrival flows (Poisson) is sufficient. For traditional coordinated arterial signal systems having modest turning traffic (say 20%) and generous

intersection spacings, a two-flow model (flow during red, flow during green) as employed in Chapter 11 (Arterials) of the HCM is probably sufficient in most cases. An obvious improvement would be to transition to a two-flow model like used in PASSER II (flow during platoon, flow not in platoon). For interchanges operating with high turning traffic and within a system of closely-spaced intersections, arrival modeling should be at least as detailed as used in PASSER III wherein six arrival flows are possible, two (platoon and non platoon) for each of three feeding movements during the cycle. While coordination in interchange systems is probably not as effective (due to higher turning traffic and more balanced flows) as in arterial systems having predominate through traffic, at least the features of platoons arriving at interchanges could be readily identified so that queuing and delays could be reduced somewhat over solutions assuming random flow.

D.3 MACROSCOPIC FLOW MODELING

Models of traffic flow routinely used to describe continuous flow traffic facilities, like freeways, can also be used to describe the dynamics of flow at traffic signals (5). These models can be used to describe the nature of operational problems experienced. Traffic signal operation routinely creates brief interruptions to continuous flow, forming bottlenecks where the arrival demand exceeds output capacity during the red interval. Shock waves form behind the signal due to the queuing and spillback that occurs (6).

D.3.1 Flow Models

When traffic flow can be assumed to be in a steady-state condition, even for a fairly brief period of time (t) and space (x) such as for random arrival flow or platoon flow at saturation, then the average flow rate during this period, v (v_{pt}) can be thought of as being produced by a traffic stream having an average density k (v_{px}) traveling at an average speed u (x_{pt}), or

$$v = k u \quad (D-5)$$

where:

v	=	average traffic flow rate, v_{pt} ;
k	=	average traffic stream density, v_{px} ; and
u	=	average traffic speed, x_{pt} .

In undersaturated, signalized arterial operations, the flow rate on a short section of roadway just upstream of a traffic signal is usually in one of three states: $v = v$, $v = 0$, or $v = s$. That is, the flow is characterized as being either the arrival flow, stopped in queue, or the queue has been transformed into a platoon with saturation flow, s . These three flow states were identified in Figure D-1. Whether these changes from one state to another occur almost instantaneously (creating shock waves) or transition over a brief period of time (forming characteristic waves) is more of a theoretical issue and presumed herein to be of minimal importance. For convenience, rapid response to signal change is assumed so that conventional shock wave analysis can be employed.

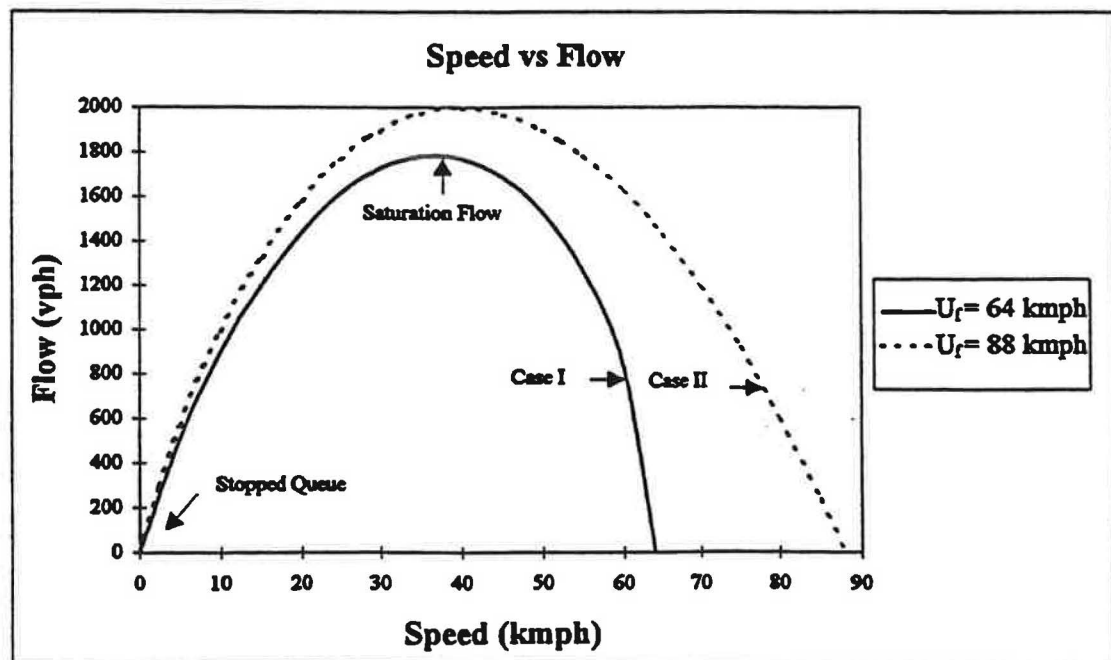
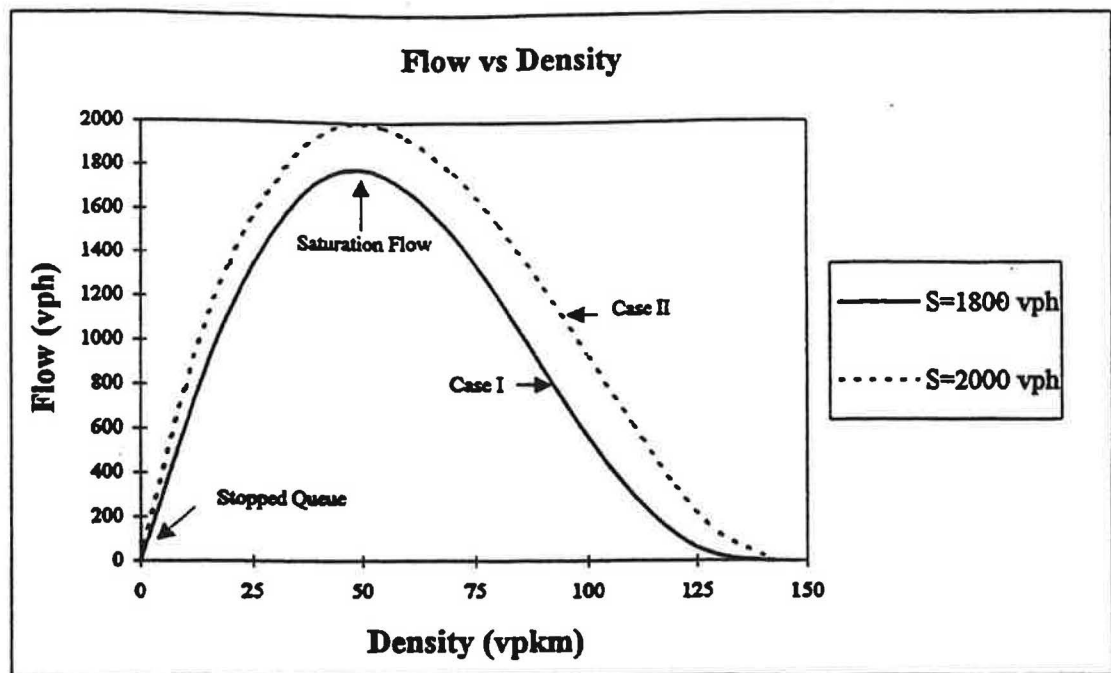


Figure D-2. Characteristics of Traffic Flow for Two Capacity Conditions.

D.3.3 Shock Wave Speed

When arterial arrival flow passing through a green signal is suddenly stopped by the onset of red, the output flow suddenly drops from $v = v$ to $v = 0$. When this change in output flow of the section occurs, the storage begins to queue behind the signal at a storage density k_q , the queuing "jam density" of about 7.0 meters per vehicle (23 feet/vehicle). The speed at which the storing queue propagates (spills back) upstream can be estimated from shock wave theory (6) as

$$W_r = \frac{\Delta v}{\Delta k} = \frac{v}{k_q - k} \quad (D-8)$$

where:

W_r	=	shock wave spillback speed due to red onset, xpt;
v	=	arrival volume (during red), vpt; and
k	=	traffic stream density, vpx.

The speed at which the shock wave propagates upstream increases with increasing arrival volume. For an approach having average flow conditions described as Case I in Figure D-2, and a green ratio of 30% which yields a phase capacity of 540 vphpl, then for arrival volumes of 20, 60 and 100 % of signal capacity, the shock wave speed progressing upstream during red would be estimated by Equation D-8 to be 0.22, 0.67, and 1.12 mps, respectively. However, if the signal became oversaturated or poorly timed such that the start of red ended platoon motion while at saturation flow $v = s$, then the shock wave speed would rise to 5.35 mps. However, as long as the signal is undersaturated, the maximum queue length per cycle would remain the same.

D.3.4 Platoon Wave Speed

When the signal turns green following an extended red time and subsequent queue buildup, the platoon responds and begins to move forward, reaching saturation flow conditions in a few seconds after green onset. For HCM-level of analysis, the platoon is assumed to reach saturation flow almost immediately once the queue-platoon transformation begins at any point in the queue. Under these simplifying assumptions, the platoon's green wave speed would be given by

$$W_s = \frac{\Delta v}{\Delta k} = \frac{s}{k_q - k_s} \quad (D-9)$$

where:

W_s	=	platoon start-up green wave, xpt;
s	=	saturation flow during green, vptpl;
k_s	=	platoon density during saturation flow, vpxpl; and
k_q	=	queue storage density, vpxpl.

The queue's transformation into a moving platoon would propagate upstream following green onset at 5.35 mps (12 mph). If the output saturation flow is unimpeded on the downstream link, then the green wave speed is fairly constant over all arrival volume conditions, simplifying the analysis. The main problem would be to determine how far upstream queue spillback has progressed each cycle before it begins clearing.

D.3.5 Clearing Wave Speed

Once the platoon begins to move forward on green, flow in the platoon is assumed to be saturation until the platoon clears the stopline. The arrival volume, v , is entering the upstream end of the platoon; whereas, saturation flow, s , presumably is occurring downstream to the stopline, or downstream boundary. During this clearing period, the back of the platoon is traveling downstream from its maximum backup at the platoon clearing wave speed of

$$W_c = \frac{\Delta v}{\Delta k} = \frac{s - v}{k_s - k} \quad (D-10)$$

Continuing with the data of Figure D-2 where $k_g = 143$ vpkmpl, $s = 1800$ vphpl, $u_s = 37$ km/h, then $k_s = 48.6$ vpkmpl from $k = v/u$. An examination of the above wave speeds follows.

Figure D-3 presents the resulting wave speeds for the above conditions for volume-to-capacity ratios of 0.2 to 1.0 for a selected g/C ratio of 0.3 and a 100-second cycle. The speed of the shock wave, W_r , is very slow (about 1 mps) and only increases slightly with increasing v/c ratios of the signal. The platoon start-up wave, W_g , is noted to be a constant of 5.35 mps, and the platoon clearing wave is high (about 16 mps) and only decreases slightly with increasing arrival volumes. Thus, because the wave speeds are fairly insensitive to arrival volumes at traffic signals, analyses based on the wave speeds are relatively stable as long as traffic flow on the link is undersaturated.

D.3.6 Queue Spillback

The duration and extent of queue spillback determines whether an upstream intersection will be severely affected by downstream operations. In essence, the characterization of adjacent intersections being too "closely spaced" can be defined for undersaturated conditions where $X \leq 1$. Using the above shock wave theory, the maximum length of queue spillback can be determined in time and space by algebraic solution of the W_r and W_g wave intercepts for a given red time, r . The elapsed time following green onset before W_g catches W_r is

$$T_g = \frac{W_r}{W_g - W_r} r \quad (D-11)$$

where T_g is the elapsed time since the onset of (effective) green.

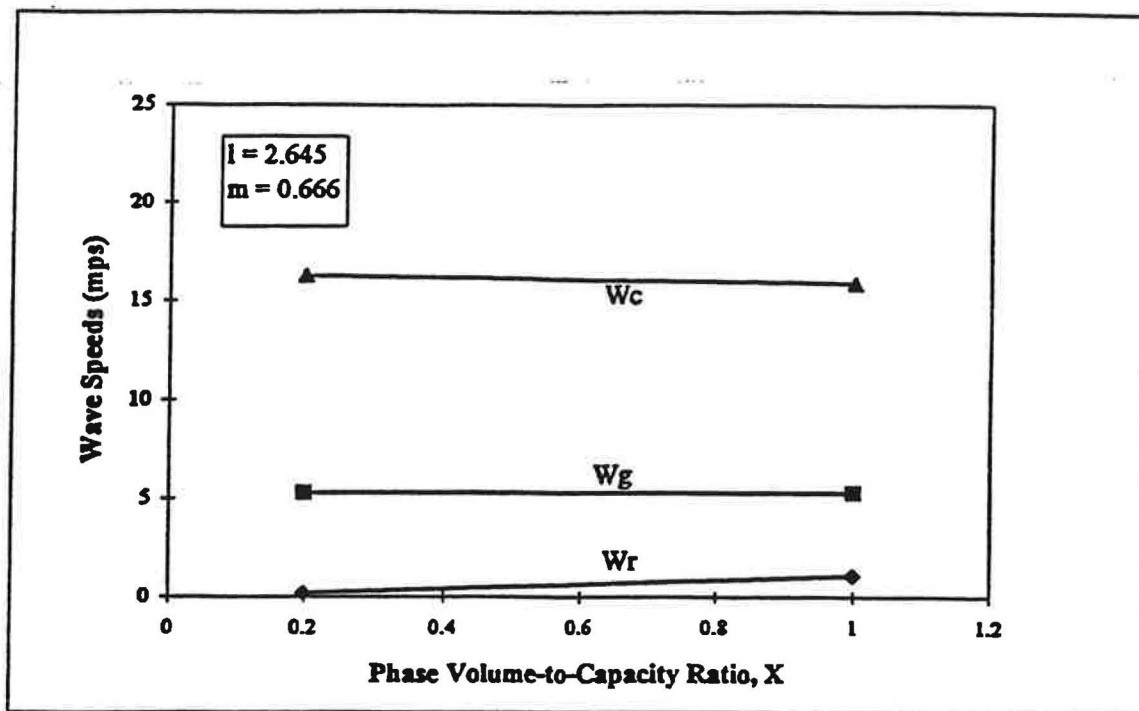


Figure D-3. Wave Speeds at Traffic Signals During Undersaturated Conditions.

The maximum queue backup for undersaturated conditions, L_m is equal to $L_m = W_g * T_g$ or

$$L_m = \frac{W_g W_r}{W_g - W_r} r \quad (D-12)$$

subject to the restriction that the downstream phase is not oversaturated.

Figure D-4 presents the queue spillback for an approach having random flow (uniformly distributed over the cycle) for v/c ratios up to 1.0, or saturation. The traffic and control conditions are as above ($s = 1800$ vphpl, $C = 100$ sec, $g/C = 0.3$). The capacity of the approach is 540 vphpl ($c = 0.3 * 1800$). The maximum queue spillback distance upstream from the stopline would be 52 meters and 99 meters for v/c ratios of 0.6 and 1.0, respectively, using Equation D-12. These results indicate that storage links less than 100 m may often experience spillback problems on entry flows where good progression is not provided, even when the downstream flow is undersaturated.

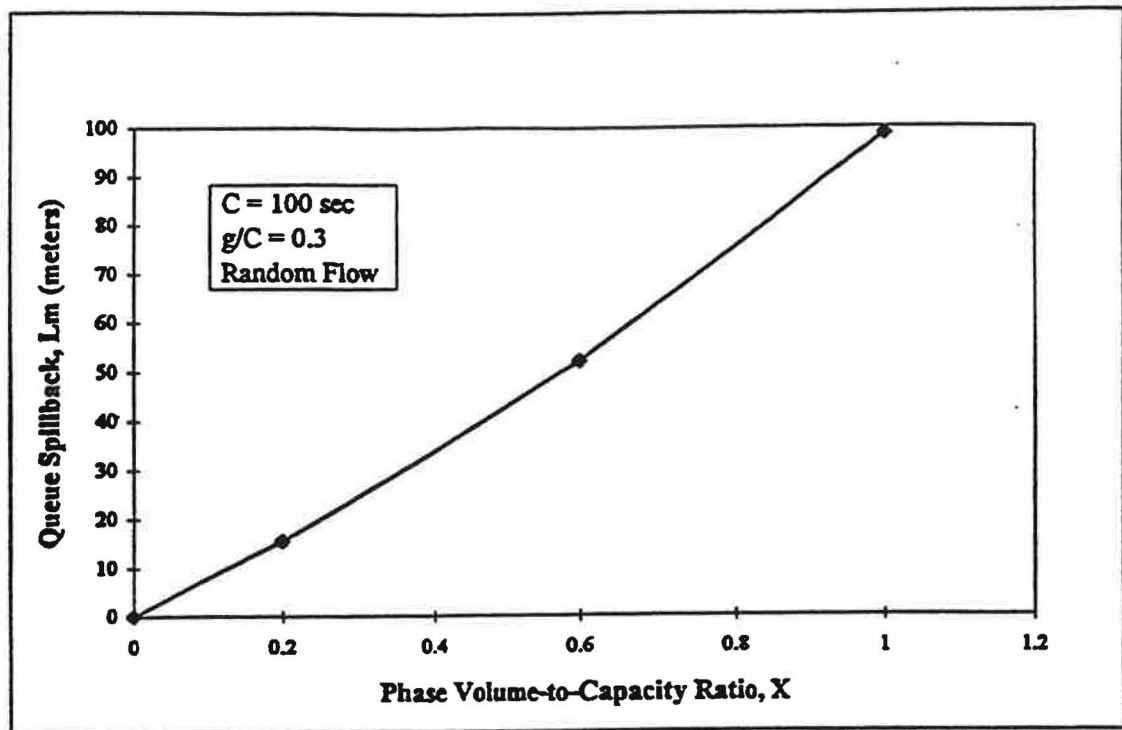


Figure D-4. Affects of Volume on Queue Spillback for Undersaturated Conditions.

D.3.7 Two-Flow Arrival Models

The 1994 Highway Capacity Manual (HCM) and its proposed arterial enhancements (1,7) assume that the arrival volume along an arterial is composed of two arrival flows: a flow arriving on the red, v_r , and a flow arriving on the green, v_g , as noted in Equation D-1. The HCM's two-flow arrival model can also be applied to the above queue spillback equations with little change in form. Defining the arrival volume on green, v_g , to be

$$v_g = R_p v \quad (D-13)$$

where R_p is the platoon ratio, and the arrival volume on red, v_r , to be

$$v_r = \frac{C - g R_p}{r} v \quad (D-14)$$

then the maximum queue backup during undersaturated conditions can be calculated from

$$L_m = \frac{W_{gvg} W_{rvr}}{W_{gvg} - W_{rvg}} r \quad (D-15)$$

where the wave speeds are based on the arrival volumes estimated for the respective red/green signal indications (r , g). Two wave speeds for the initial shock wave, W_r , should be computed: one based on arrival volume v_r and one based on v_g . Should the volume arriving on red be eliminated by great progression, then the shock wave speed on red, W_{rr} , would be zero (0.0) and no queue spillback would occur. This routinely occurs on the outbound phase at diamond interchanges with good four-phase with two-overlap timing. The extremely bad progression case where all the arrival volume arrives on the red is described below.

D.3.8 Maximum Spillback

It is useful to examine the worst-case scenario for undersaturated conditions where all the arrival traffic to a phase arrives on red due to extremely bad progression. The main issue here is "how far upstream would the queue back up, storing at 7.0 meters per vehicle when arrival volumes have reached but not exceeded capacity levels?" Applying the equation of continuity to this fairly simple problem during red results in the following equation for maximum queue backup:

$$L_{\max} = 7.0 v_c \quad v_c \leq c_\phi \quad (D-16)$$

where:

L_{\max}	=	maximum queue backup, meters;
v_c	=	average arrival volume per cycle per lane, vpCpl;
c_ϕ	=	phase capacity per cycle per lane, $s[C/n - 1]/3600$;
C	=	cycle length, sec; and
s	=	saturation flow, vphgpl.

Figure D-5 provides graphs of the maximum spillback for two high-volume cases (400 and 600 vphpl). Also depicted in Figure D-5 is the queue spillback at saturation when the average arrival volume (on red) is equal to the phase capacity. Phase capacity presumes three equal phases each having 4.0 seconds lost time with saturation flows of 1,800 and 2,000 vphgpl. Maximum arrival volumes of 520 and 580 vphpl can be handled for a 100-second cycle at capacity.

Maximum queue spillback is noted to vary with several factors. Maximum spillback increases with increasing arrival volume and cycle length (due to longer reds). Capacity also increases with increasing cycle due to reductions in the proportions of lost time per cycle. A maximum queue spillback of 166 meters (545 feet) is estimated for a 140 second cycle.

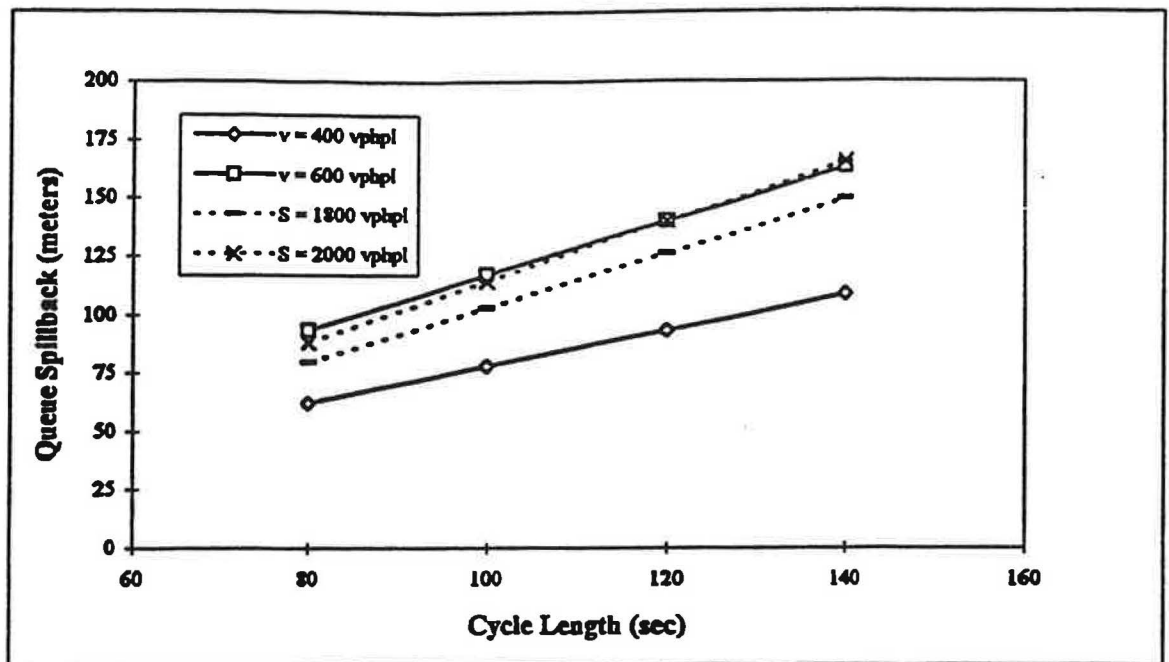


Figure D-5. Maximum Queue Spillback for Capacity Conditions.

D.5 OVERSATURATION

The period of oversaturation is the time when the arrival (demand) volume on a signal phase exceeds its capacity. Oversaturation may occur because the arrival demand has grown to a level greater than the capacity, or the phase capacity may have been reduced due to downstream spillback impeding the saturation flow during green. In any case, queue spillback and congestion on the link are eminent unless control conditions are rapidly changed. Once the signal becomes oversaturated, then traffic operational response to signal control inputs is often the reverse of prior undersaturated response. This error can only feed the growing congestion unless corrections are made quickly.

D.5.1 Transition Time

The time a link L is in transition from undersaturated to oversaturated operating conditions is brief. The time can be calculated from the input-output model derived from the basic equation of continuity of vehicular flow on the link. The input-output model for time periods longer than one cycle is

$$N(L, t) = N_0 + \sum v_m t - \sum c_m t \quad N(L, t) \leq N_{max} \quad (D-17)$$

where:

$N(L, t)$	=	number of vehicles operating on the link of length L at time t , vehicles;
N_0	=	number of vehicles operating on the link at start of period, vehicles;
v_m	=	total arrival volume to head of link destined to movement m , vph;
c_m	=	output capacity of link serving movement m , vph; and
N_{max}	=	maximum number of vehicles that can store on link, vehicles, $k_q L$ with a typical storage density of 143 vpkmpl, or storage spacing of 7.0 m/veh (23 ft/veh).

When it can be assumed for analysis of a phase that the arrival volume and capacity flow are constants for a period of cycles, then Equation D-17 can be solved directly for t . Letting $N(L, T) = N_{max}$ such that T_f is the time to completely fill the link when $v_m \geq c_m$ then T_f is

$$T_f = \frac{N_{max} - N_0}{v_m - c_m} \quad (D-18)$$

Letting v_m be described in terms of the v/c ratio, $X = v/c$, and substituting the above relationships

$$T_f = \frac{(k_q - k) L}{c (X - 1)} \quad (D-19)$$

Examination of typical situations is aided by letting $c = (g/C) s$, $X = 1/PHF$ where PHF is the peak hour factor ($PHF = 0.9$), $s = 1800$ vphpl, $k = k_m = 9$ vpkmpl, and $k_q = 143$ vpkmpl so that

$$T_f = \frac{(143 - 9) L}{g/C \ 1800 (1.1 - 1.0)} = \frac{134 L}{180 g/C} \quad (D-20)$$

where T_f is in hours and L in km. Assuming that the green ratio, g/C , is 0.3 results in link fill times during an $X = 1.1$ oversaturation of 15, 30, and 45 minutes for available storage lengths of 100, 200 and 300 meters, respectively. As Equation D-19 shows, there are some obvious advantages to having long spacings between signals; longer links can tolerate proportionally longer periods of oversaturation before filling and spilling back into an upstream intersection.

By defining L to be the available storage space on the link, it can also be seen for the above example that a 200-meter link having a maximum queue spillback of 100 meters during undersaturated conditions would begin to experience spillback into its upstream intersection 15

minutes into the oversaturation period and become fully saturated and congested 30 minutes into the period. Smart signal progression could forestall the onset of major congestion at the upstream intersection by up to 15 minutes in this case. Obviously, the larger the rate of overload, $X > 1$, the shorter the time before major congestion begins at the impeded upstream intersection, presuming minimal upstream metering of input flow.

D.5.2 Oversaturated Operations

Once the subject link, L , becomes flooded with vehicles due to oversaturation, flow into the link from the upstream signal depends almost entirely on the output capacity of the downstream signal of the subject link, which then may depend on its downstream signal, etc. The dependence increases with the increasing percentage of upstream traffic bound for the downstream bottleneck phase. Essentially, the above equation of continuity (Equation D-17) requires this result because $N(L, T)$ must equal N_o , so that the volume in must equal the volume out, or

$$v_u = c_d \quad (D-21)$$

An example of this constrained input process is shown in Figure D-6 where two upstream flows having a combined demand volume of 1.6 times the downstream capacity load the link. The microscopic traffic simulation program NETSIM was used to develop these observations which represent the average value of 10 replications of the conditions shown. The offset between the upstream and downstream signals was varied while the total link throughput and the individual feeding flows were observed for both upstream movements. The total throughput volume was observed to be essentially constant (the downstream capacity) over all offsets. However, the input flows that did occur were related to the signal offset (θ) of the link, but the sum was constant, or

$$v_{u1}(\theta) + v_{u2}(\theta) = c_d \quad (D-22)$$

In this case when all upstream turning movements are so large that each can keep the link full when the phase ends, then no downstream storage space remains at the end of any phase, $N(L, T) = N_{max}$. Actual cases like this are not rare at urban interchanges where rush hour turning volumes can be high on all feeding movements to the link.

As Figure D-6 shows, the selection of offset determines the share each competing movement receives of the total available downstream stream capacity. That is, the offset allocates the limited resource, output capacity, among competing and excessive demands. What one upstream movement receives, another loses, as this situation becomes a "zero sum" game. As Figure D-6 demonstrates, signal offsets that are based on undersaturated progression analysis, say using PASSER II, are often designed to favor the cross arterial's operation, usually do not provide the same favored status when periods of oversaturation occur. In fact, the PASSER II offsets may actually wrongly favor the cross-street turning traffic in some cases of oversaturation.

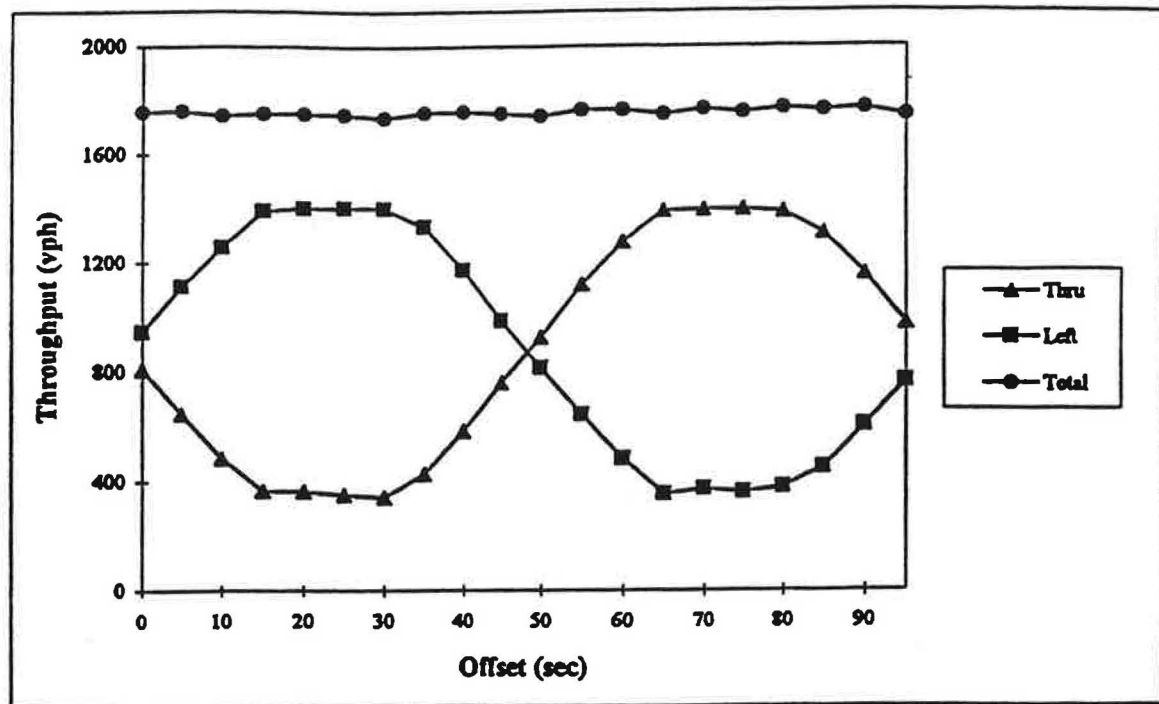


Figure D-6. Traffic Performance on an Oversaturated Link Related to Offset.

In another case of oversaturation, where some input phases at the upstream signal cannot keep the subject link L filled to capacity at the end of the phase, some link storage capacity becomes available to subsequent movements. Offset analysis algorithms, such as the PDX Model to follow should be able to determine how to optimally use this available storage space.

Another problematic case can arise during oversaturation. The condition is called "demand starvation" (8), which seems like an oxymoron during periods of high demand volumes. Demand starvation may routinely occur on short links when an upstream phase feeding the link (plus link storage) cannot keep the downstream bottleneck signal (say Phase c) flowing at saturation during its green. This lost output per cycle cannot be recovered, and the maximum throughput flow is reduced below the phase capacity c_d , even though the total demand on the downstream phase may exceed c_d . Demand starvation can be a problem for some interchanges when the outbound left turn Phase c closes off flow into the interchange. Signal coordination should be provided to ensure that the cut-off phase does not empty when it is (or becomes) oversaturated.

In the following sections, two traffic models will be presented that describe operating conditions during oversaturation. One simple model estimates the maximum delay that might be incurred on an oversaturated link. Another more complex model estimates all performance measures for a link after the model determines whether the link is experiencing spillback and oversaturation.

D.6 DELAY MODEL DURING FLOODED CONDITIONS

D.6.1 Flooded Conditions

In "flooded" conditions of oversaturation, the upstream demands keep the link full at all times. Flow conditions on the link are either saturation flow or stopped in queue, but are never the arrival flow. The spillback constantly engulfs the link in either a stopped queue or a moving fully saturated platoon. Figure D-7 presents both the undersaturated and oversaturated flow conditions upstream of a traffic signal. In oversaturated, fully flooded flow, only two flow states exist (saturation and no flow), resulting in only two "moving" waves (shock wave and platoon wave).

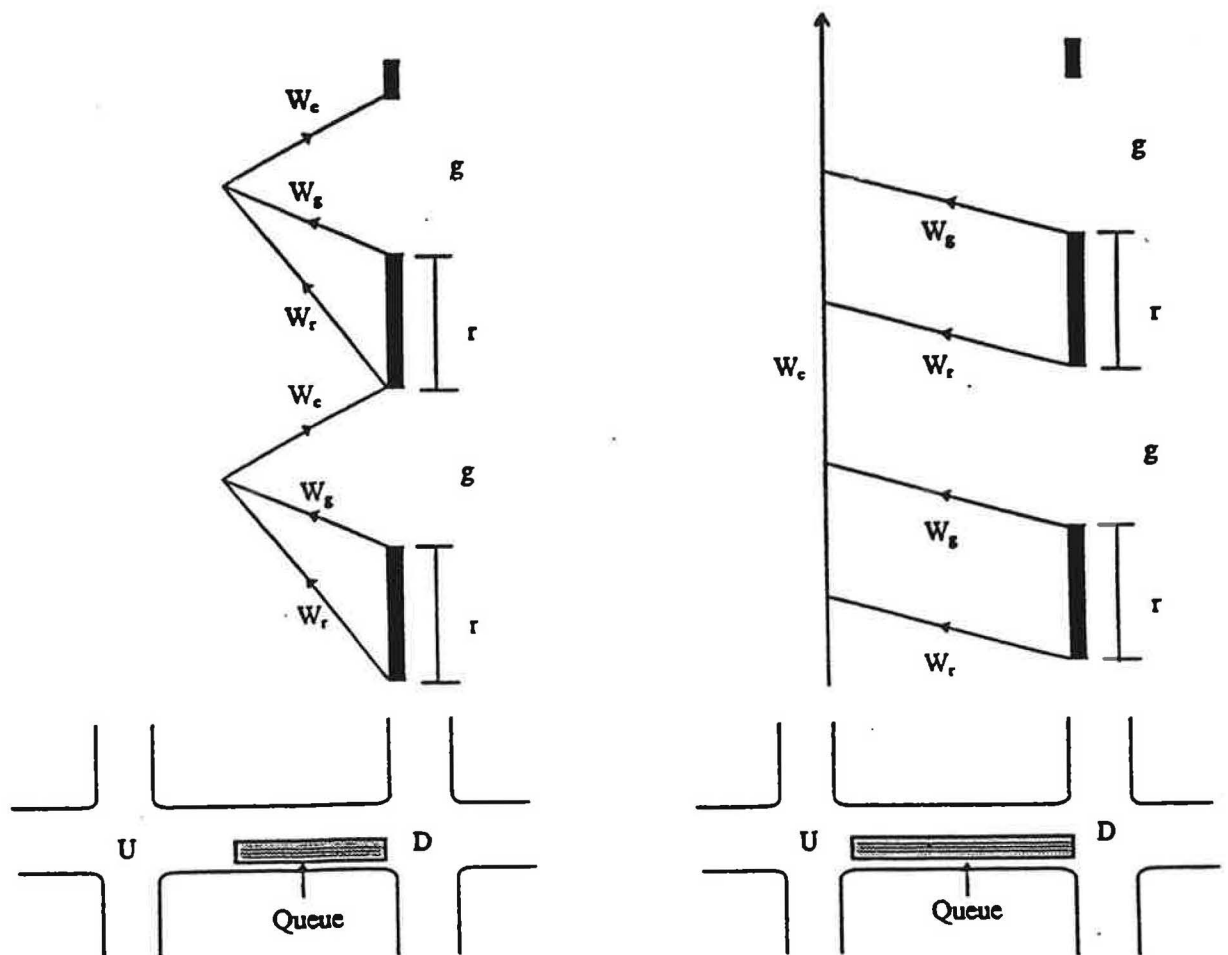


Figure D-7. Traffic Flow Regions for Undersaturated and Oversaturated Conditions.

D.6.2 Delay Model

In the flooded condition having only two flow states, the speed of the shock wave at red onset, W_r , is equal to the speed of the flow wave following green onset, W_g , and they are

$$W_r = W_g = \frac{\Delta v}{\Delta k} = \frac{s - 0}{k_g - k_s} \quad (D-23)$$

The average link density over a representative cycle of time is assumed to be given by the fraction of time any point on the link $p(x,t)$ is stopped or moving. Because the two wave speeds are equal, the fractions are equal to the fractions of time the cycle (C) is effectively (g) green and red (r), respectively, for dedicated movements (lane groups). In this case, the average density becomes

$$k = \frac{g}{C} k_s + \frac{r}{C} k_g \quad (D-24)$$

Assuming that the average flow on the link is given by the average downstream capacity c_d where

$$v = c_d = \frac{g}{C} s_d \quad (D-25)$$

then the average travel speed on the link can be calculated to be

$$u_l = \frac{v}{k} = \frac{(g/C) s_d}{(g/C) k_s + (r/C) k_g} \quad (D-26)$$

Dividing the above equation by g/C and then defining $\alpha = r/g$ for the controlling downstream signal phase, the average travel speed on the link can be found to be

$$u_l = \frac{s_d}{k_s + \alpha k_g} \quad (D-27)$$

Since the speed at saturation flow is $u_s = s/k_s$, then dividing the above equation by k_s/k_s , yields

$$u_l = \frac{u_s}{(1 + \beta)} \quad (D-28)$$

where:

$$\beta = \frac{k_q}{k_s} \cdot \frac{r}{g} \quad (\text{D-29})$$

For example, if $k_q = 143$ vpkmpl, $k_s = 51.3$ vpkmpl, and $s = 1900$ vphpl, so that the nominal saturation speed is 37 km/hr, then for a green split at the downstream bottleneck intersection of 0.46,

$$\beta = \frac{k_q}{k_s} \cdot \frac{r}{g} = \frac{143}{51.3} \cdot \frac{r}{g} = 2.79 \cdot \frac{r}{g} = 2.79 \cdot \frac{54}{46} = 3.27 \quad (\text{D-30})$$

or the average link travel speed, u_l , is

$$u_l = \frac{u_s}{1 + \beta} = \frac{37}{1 + 3.27} = 8.66 \text{ km/hr} \quad (\text{D-31})$$

The link average travel speed during flooded conditions will always be less than the speed at saturation and will be highly dependent on the throughput capacity of the bottleneck phase.

Link delay can now be calculated as the difference between the overall link travel time and the baseline running time at the approach running speed u_a ,

$$d \text{ (sec/veh)} = \frac{L}{u_l} - \frac{L}{u_a} \quad (\text{D-32})$$

Should demand starvation also occur during the cycle, then the delay should be reduced in proportion to the reduction in effective queue length. Assuming that the link length was 100 meters and the nominal link running speed was 56 km/hr (35 mph), then the delay on the oversaturated link would be

$$d = 3600 \left(\frac{0.100}{8.66} - \frac{0.100}{56} \right) = 35.1 \text{ sec/veh} \quad (\text{D-33})$$

D.6.3 Delay Model Results

Several simulation studies were conducted to further test and verify Equations D-31 and D-32. A range of split and offsets were examined using the NETSIM microscopic traffic simulation model. Three downstream green splits were tested so that the over a cycle of offsets to verify the model results. Figure D-8 presents the simulation study for a 100 meter (328 foot) link length.

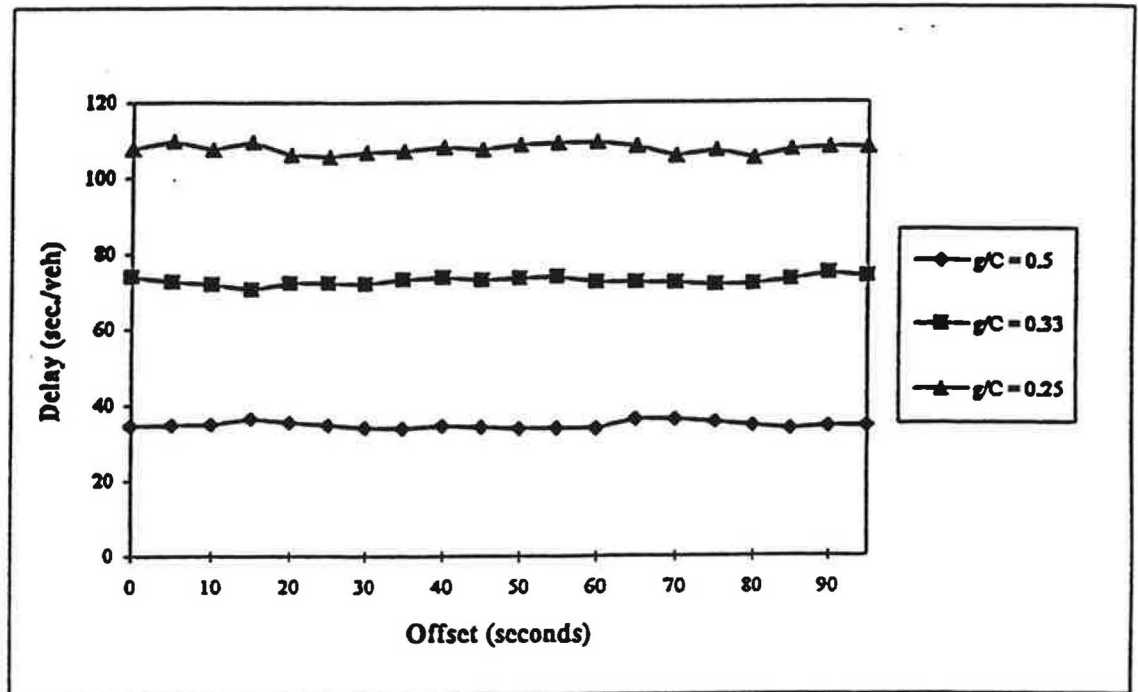


Figure D-8. NETSIM Simulation Results of Oversaturated Delay on 100 m Link.

A summary of these study findings is presented in Table D-10 for the 100 meter (328) foot link length for the three green splits. It should be noted that the projected v/c ratio (degree of saturation) on the downstream is 2.0, 3.1 and 4.3 for the three green splits, but the delays predicted by Equation D-32 are not extremely large. However, large delays would be estimated by the Highway Capacity Manual delay estimation model.

Table D.10 Estimated and Simulated Link Speed and Delay for Flooded Oversaturated Links

Study Case No.	Downstream Green Splits	Travel Speed km/hr	Model Link Delay sec/veh	NETSIM ^a		HCM ^b	
				Speed km/hr	Delay ^a sec/veh	Speed km/hr	Delay ^b sec/veh
1	0.46	8.66	35.1	8.70	34.7	2.22	155.7
2	0.29	4.73	69.7	4.51	72.9	2.13	162.7
3	0.21	3.22	105.4	3.14	107.7	2.09	166.0

^a Averaged over all offsets for 10 iterations.

^b Arbitrarily restricted to $v/c \leq 1.2$; otherwise, HCM model delays would have been extremely large.

Other findings can be observed from these studies for oversaturated, fully flooded conditions from Equations D-31 and D-32. The speed and delay on these links depend on the downstream bottleneck r/g ratio and resulting throughput rate. The delay on the link also depends on the length of the link, L , as given by Equation D-32. Longer links can have proportionally higher delays, but the average travel speed and delay rate per kilometer do not change for oversaturated links. These latter findings suggest that the level of service for an extended section of an arterial based on average travel speed provides a constant assessment of the quality of operations provided. However, methods based on point measures of queuing delay are not likely to reliably estimate the level of service being provided, nor the true congestion experienced.

Additional improvements in traffic performance will arise as the link length increases. The length is not as likely to flood with increasing length, particularly in mesosaturation conditions, and for short-term oversaturation conditions noted earlier. If minor movements feed the link, then the link will partially clear during these phases permitting increased link speed and lower delays for the following phase(s), depending on the relative offset being employed for the subject link.

D.7 PROSSER-DUNNE-EXTENDED TRAFFIC MODEL

This section describes the theoretical framework for a macroscopic traffic model developed to solve operational problems associated with high-volume operations on urban arterials having closely-spaced signalized intersections. As may occur, the vehicle discharge flow rate at an upstream intersection may be seriously affected by operational conditions on the downstream link. Operational impacts may include (1) reduced saturation flow due to limited travel distance to the back of a stopped queue on the downstream link and (2) total signal blockages that reduce the effective green time due to queue spillback into the upstream intersection. These issues are carefully evaluated in this traffic model to provide accurate prediction of the discharge flow rate, capacity and delay at the upstream intersection. Other related outputs from this model include the queue dynamics information throughout the cycle and its interaction with the estimated discharge flow rate.

The underlying theory of this paired-intersection model is based on an analytic traffic model originally presented by Prosser and Dunne in 1994 (9). The Prosser-Dunne Model used a graphical technique to estimate the reduced effective green time based on the assumption that no vehicle can be discharged from the upstream intersection whenever there is a queue spillback blocking the intersection. The Prosser-Dunne Model assumed that the downstream congested link was completely oversaturated, or flooded.

D.7.1 Model Enhancements

Texas Transportation Institute (TTI) has extended this model to a wide range of operating conditions, increased the modeling to three feeding movements, and added the effects of saturation flow variation with travel distance (as described in Appendix C) to the model structure. In addition, TTI has coded the upgraded model into the FORTRAN language for implementation and testing. A second-by-second time-step approach has been implemented to increase the quality of the analysis. Extensive NETSIM simulation studies have been conducted to calibrate and verify the overall model's capabilities with field data collected within this project. The following sections describe the design, logic and test results of the resulting algorithms and computer program, which is referred to as the PDX Model.

D.7.2 PDX Model

The PDX Model addresses the operational problem depicted in Figure D-9. Three upstream turning movements (m) can possibly feed a downstream intersection along the link $u-d$. The link can be of any length, L . These turning movement may be the upstream through, left and right turns onto link L . Both intersections must operate on a pretimed phasing and have the same cycle length for the period of interest. The algorithm seeks to determine the effective green, phase capacity and delay for the upstream movement and the delay on the downstream link, L , for a given set of signal timings and relative offset (θ) between the signals. The current program does not seek to determine an optimal offset θ_{opt} , but this feature could be provided by existing optimization programs like PASSER III. PDX is only an analysis program like the HCM/HCS software.

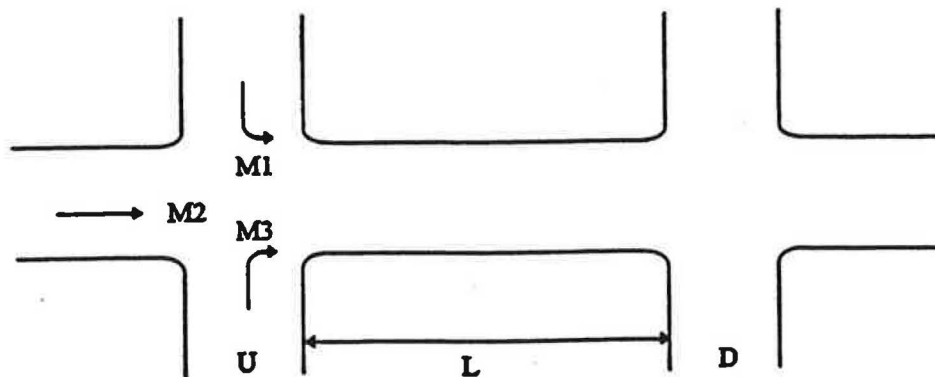


Figure D-9. Paired-Intersection Operational Problem Analyzed by PDX Model.

Figure D-10 presents the basic time-space diagram for oversaturated links after Prosser-Dunne (9). The variables so identified are defined in the following model development. In the PDX Model, however, links can be analyzed that may not be oversaturated by all phases combined, or may not be flooded by some combinations of phase sequence and/or offset, as noted earlier for demand starved links.

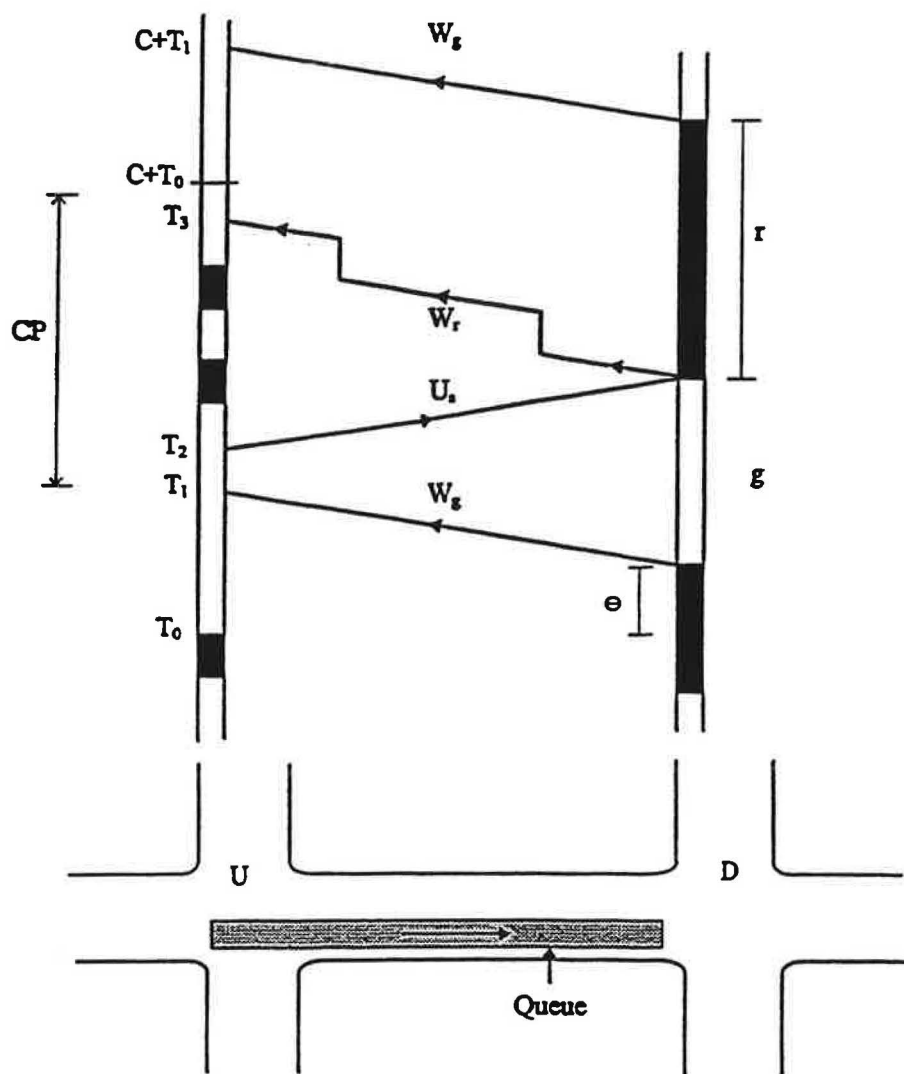


Figure D-10. Operational Time-Space Diagram on Congested Links.

Figure D-11 presents a simplified flow chart for the enhanced PDX Model. Time during the cycle is sequentially incremented in one-second increments in the PDX Model. The Model is coded in FORTRAN computer code. The detailed step-by-step PDX process is summarized in the following paragraphs.

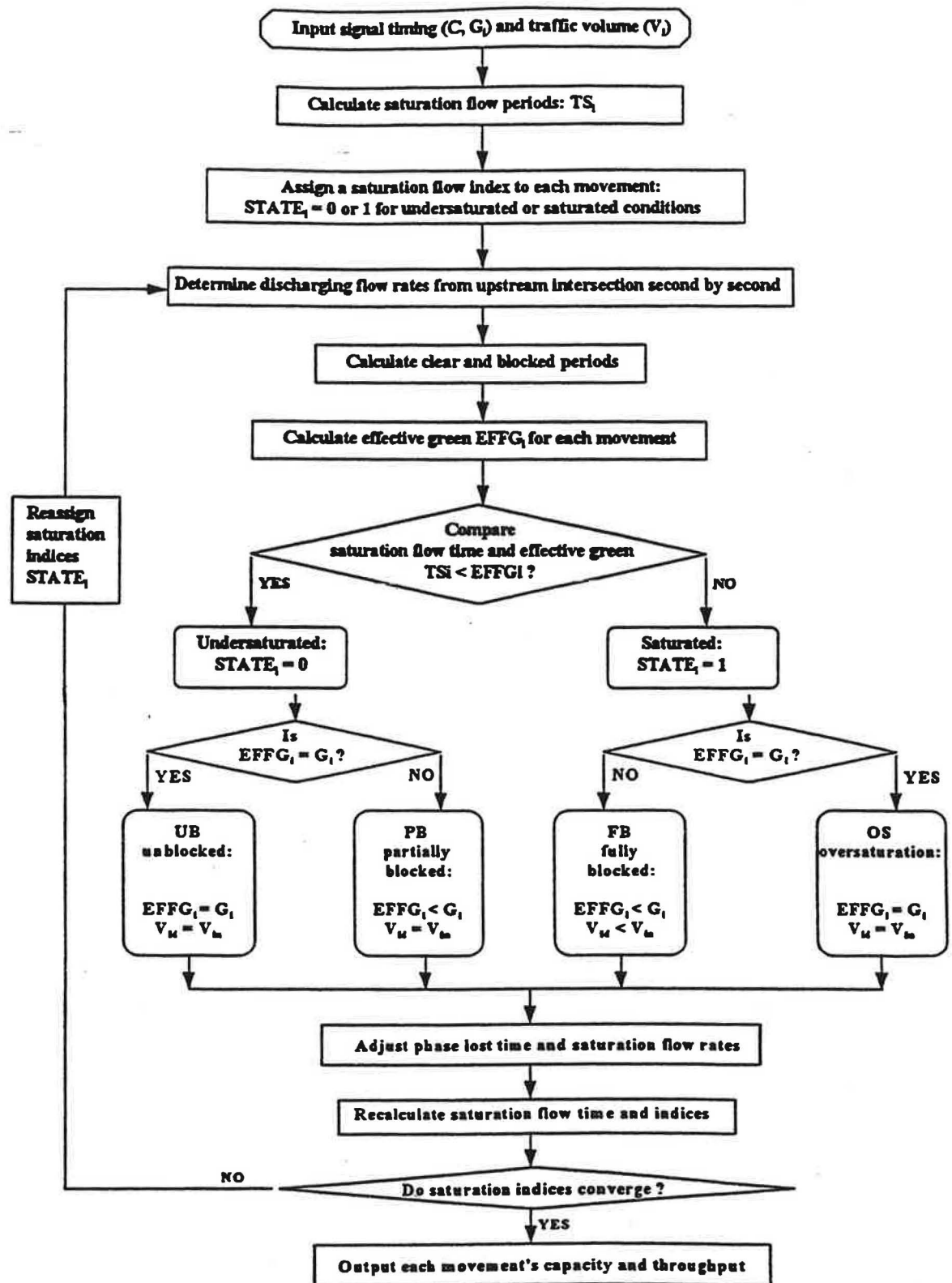


Figure D-11. Flow Chart of PDX Model.

Step 1. Identify Input Parameters

The following parameters are identified as inputs to the program:

G_{um}	=	Green time of upstream movements, sec;
G_d	=	Downstream green time, sec;
C	=	Cycle time, secs;
SL, EU	=	Start loss and end gain for each movement, sec;
θ	=	Offset between intersections, sec;
L	=	Intersignal link length in meters, m;
n_{max}	=	Number of vehicles that can be stored in the intersignal link length; veh;
v_m	=	Arrival flow rate for upstream movement m , vph;
t_q	=	Blocking queue clearance time, sec;
t_f	=	Link travel time during saturation flow, sec; and
s	=	Unimpeded saturation flow rate, vpsg.

The yellow and red clearance times are together taken to be four seconds for both intersections (u, d). Values of t_q , t_f and n_{max} noted in Figure D-10 are calculated by the following formulas:

$$t_q = L * (k_{qd} - k_{sd}) / s_d \quad (D-34)$$

$$t_f = L * k_{sd} / s_d \quad (D-35)$$

$$n_{max} = L * k_{qd} \quad (D-36)$$

where:

L	=	Link length, km;
k_{qd}	=	Queue density of downstream movement, vpkkm;
k_{sd}	=	Saturation flow density of downstream movement, vpkkm; and
s_d	=	Saturation flow rate of downstream movement, vphg.

Conversion factors to the desired units are not shown in the above equations.

Step 2. Calculate Start and End of All Movements

The start and end of any movement m ($m=1,2,3$) at the upstream intersection with respect to time $T_0 = 0$ is referred to as SM_m and EM_m . These times can be calculated as follows:

$$SM_1 = 0 \quad (D-37)$$

$$EM_1 = G_1 \quad (D-38)$$

$$SM_2 = GU_1 \quad (D-39)$$

$$EM_2 = GU_1 + G_2, \text{ etc.} \quad (D-40)$$

where G_m is the displayed green time for movement m , $m = 1,2,3$, at the upstream intersection.

The offset θ between the intersections $u-d$ is considered from the start of the green period of movement 1 at the upstream intersection until the start of the through green at the downstream intersection. Hence, G_d starts a time θ after T_o , equal to the offset with respect to the upstream intersection. The end of the downstream bottleneck phase, EM_d , is

$$EM_d = T_o + \theta + G_d \quad (D-41)$$

where θ = offset in seconds from u to d .

Step 3. Calculate T_1 and T_2

The beginning of the clear period, T_1 , at the upstream intersection starts when the blocking queue clears the upstream movement of interest after the start of through movement green time at the downstream intersection. Note that it has been assumed that a blocking queue exists. This may be determined later to not be true if volumes are not high enough for the length of link involved. If true, then

$$T_1 = T_o + \theta + t_q \quad (D-42)$$

In order to evaluate whether the intersignal link length will be completely blocked under different traffic conditions, some related factors have to be evaluated. One of them is the volume-to-capacity ratio (X_d) at the downstream intersection. Another one is the critical link length (L_c) which determines the occurrence of queue spillback. The critical link length is calculated as follows:

$$L_c = G_d * W_r * W_g * (C - G_d) / ((W_r + W_g) * (C/X_d - G_d)) \quad (D-43)$$

where:

- W_r = Shock-wave speed (mps); and
- W_g = Platoon starting wave speed (mps).

When X_d is greater than 1.0, the downstream link will be oversaturated by vehicles coming from the upstream intersection. If the link length L is greater than critical link length L_c , it is assumed that a proportion of the storage vehicles will not be cleared from the link due to the limited capacity at the downstream intersection. Hence, some residual queue will remain on the link when the green ends. When the link length is shorter than the critical link length, all vehicles stored on the link will be cleared. If X_d is less than 1.0, queue spillback may also occur due to an inappropriate signal timing plan for the downstream green time and offset. This problem can be solved by using the same assumption as the one for oversaturation and by checking the results after the loop calculation in the computer program. If some blockage is found to occur, then saturation flow adjustments are made to account for either impeded or blocked flow using Equation C-46 in the previous appendix.

The time at the upstream intersection, denoted as T_2 , when vehicles discharged from the upstream intersection will first begin queuing at the downstream intersection stop line on red, is calculated as follows:

$$T_2 = T_0 + \theta + G_d - t_f \quad T_2 \geq T_1 \quad (D-44)$$

Step 4. Calculate Input Flow Rates during the Cycle

The potential flow rates from the upstream intersection during the cycle are calculated in this step. Each upstream movement's capacity is computed based on the effective green period of that movement ($SM_m + SL$ to $EM_m + EU$) and the saturation flow rate s . The duration of time, t_s , that every other movement experiences saturation flow is given by

$$t_s = v_i (C - g_i) / (s_i - v_i) \quad (D-45)$$

Thus, from time $SM_m + AL$ to t_s , the flow is saturation and from t_s to the end of movement green time ($EM_m + BL$), the movement flows at its arrival flow rate, v . In the event t_s is greater than the end of green of the movement, the flow continues at saturation flow for the entire period of green for that movement. The flow at all other times in the cycle (e.g., during yellow and red clearance time) is assumed to be zero.

Step 5. Find End of Clear Period, T_3

The next step in the process is to find T_3 , the end of the clear period at the upstream intersection. The time T_3 is defined as the time when the intersignal length would be once again completely filled with vehicles, thus blocking the upstream intersection. The related duration of time t_3 is needed to accumulate enough vehicles to fill the link L after time T_2 . The number of seconds required to completely fill the intersignal length depends on the output flow of vehicles from the upstream intersection after T_2 .

The potential output flow for every second in the cycle is known from the previous steps. The summation of the output flows (vehicles per second) from the upstream intersection over time (seconds), gives the number of vehicles that may enter link L until the link completely fills to its storage capacity, n_{max} , over the duration t_3 , or

$$n(t_3, L) = n_0 + \sum q_{um} \leq n_{max} \quad 0 \leq t_3 \leq C \quad (D-46)$$

The assessment of n_0 is critical to the algorithm. $n_0 = 0$ when $T_1 < T_2$. However, if the link L can store more vehicles than the downstream phase can serve (its capacity) (because the link may be long and/or the phase relatively short) then $T_2 \leq T_1$ and $n_0 = n_{max} - c_m > 0$ (instead of zero, or some other value) at time $T = T_1$. When $T_2 < T_1$, vehicles already on the link cannot clear the next downstream phase. Thus, the most vehicles that can enter the link during the next cycle cannot

exceed the downstream phase capacity, c_m . If it is ultimately found that all the upstream flows can use the link for a time period T_3 longer than $T_1 + C$, then the link does not totally fill during the cycle and the upstream signal is unblocked, but it may still have some reduced output flow. When queue blockage occurs, the end of the clear period for entry into the upstream intersection would be

$$T_3 = T_2 + t_3 \quad T_2 < T_3 \leq T_1 + C \quad (D-47)$$

In the PDX program, a simple DO-loop calculates t_3 . The loop increments the number of seconds after t_2 and for each second adds the number of vehicles entering the link. No departing vehicles need to be considered here. When the link either fills completely or reaches the next cycle ($T_1 + C$), the incrementation process stops. The number of time steps used gives the number of seconds elapsed after t_2 , which is t_3 .

The saturation flow-queue interaction model described in Appendix C has been implemented in the PDX program. As each upstream phase begins to be analyzed during the cycle, the downstream queue length is calculated. The available travel distance to the back of the queue is then determined, knowing L . The adjusted saturation flow is determined from Equation C-8 for the upstream movement of interest. Preliminary analyses based on initial pointers and Equation D-43 estimate whether queue spillback is likely. Should the completed time-step analysis not support the initial assumption, then the alternate equation is selected and the process repeated.

Step 6. Identify Clear Period

The clear period is defined as the time in the cycle from the end of queue blocking to the start of blocking in the next cycle at the upstream intersection. The clear period, CP , is the duration from T_1 to T_3 when upstream input flow can occur, or

$$CP = T_3 - T_1 \quad (D-48)$$

The values of T_1 and T_3 have not been calculated modulo C . Thus, they can have values greater than the cycle length. While calculating the clear period, the values of SM_m and EM_m should be adjusted for start loss and end gain, respectively, at some convenient point in the program.

Step 7. Compute the Modified Effective Green Period

The modified effective green period (g_{eff}) during which the upstream movement can discharge vehicles is the time overlap of the unblocked effective green (g_u) of the movement and the clear period. The clear period is calculated in the previous step. Table D-11 shows the modified effective green periods (g_{eff}) for Movement 1 for different positions of t_1 and t_3 with respect to the upstream signal time as calculated in the PDX program. Thus, the real effective green ($g = g_{eff}$) is

$$g_{eff} = g_u \cap CP \quad (D-49)$$

Table D-11. Modified Effective Green Periods for Movement

Value of t_1	Value of t_3	g_{eff}
$SM_1 \leq t_1 \leq EM_1$	$EM_1 \quad t_3 > t_1$	$t_3 - t_1$
	$EM_1 < t_3 > t_1; t_3 \leq C$	$EM_1 - t_1$
	$EM_1 < t_3 > t_1; \text{mod}(t_3, C) < t_1$	$EM_1 - t_1 + t_3 - SM_1$
	$EM_1 < t_3 > t_1; \text{mod}(t_3, C) > t_1$	$EM_1 - SM_1$
$EM_1 \leq t_1 \leq C$	$t_1 < t_3 > C; \text{mod}(t_3, C) < EM_1$	$t_3 - SM_1$
	$t_1 < t_3 > C; \text{mod}(t_3, C) \geq EM_1$	$EM_1 - SM_1$
$t_1 > C;$ $SM_1 < \text{mod}(t_1, C) \leq EM_1$	$t_3 < 2C; \text{mod}(t_3, C) < EM_1$	$\text{mod}(t_3, C) - \text{mod}(t_1, C)$
	$t_3 < 2C; C > \text{mod}(t_3, C) \geq EM_1$	$EM_1 - \text{mod}(t_1, C)$
	$t_3 > 2C; \text{mod}(t_1, C) > \text{mod}(t_3, C) < EM_1$	$EM_1 - \text{mod}(t_1, C)$
	$t_3 > 2C; \text{mod}(t_1, C) < \text{mod}(t_3, C)$	$EM_1 - SM_1$

A similar table can be developed for estimating the modified effective green periods of the other upstream movements.

Step 8. Calculate the Capacity of the Movement

The reduced capacity of the upstream movement is calculated using the updated effective green period by the following formula:

$$c = s_a * (g_{eff}/C) \quad (D-50)$$

The adjusted saturation flow rate s_a includes the queue interaction effects as given by Equation C-8.

The above steps were coded in FORTRAN. The output of the PDX program was entered into a file which could be easily be loaded into a spreadsheet package for further analysis.

D.7.3 Model Testing and Verification

The PDX Model has been tested for various traffic operating conditions. An experimental testbed was designed for this purpose. To verify the test results, the microscopic traffic simulation program, TRAF-NETSIM, was used to provide comparative results under the same operating conditions. The following sections describe the procedures and results of this testing.

Testbed Design. In order to test the above program's applicability, an experimental testbed was designed. An arbitrary paired intersection system was set up with all the traffic and signal timing variables affecting saturation flow being defined. This test system was analyzed using the Prosser-Dunne FORTRAN program for a range of conditions and the movement's capacities were computed. A design scheme of the study paired intersection was shown in Figure D-9.

Required inputs such as traffic volumes, signal timing parameters such as green times, offsets, cycle lengths, and spacing between intersection were carefully prepared. Only a pretimed signal system was considered for this research. The spacing between the two intersections was considered to be 100, 200, and 300 meters, respectively. These spacings were assumed to be representative of most closely-spaced intersections within an interchange environment. Other parameters in the testbed are summarized in Table D-12.

Table D-12. Testing Parameters for the Program and Simulation

V/C Ratios	0.8, 0.9, 1.0, 1.1, 1.2, 1.5
Cycle length	100 seconds
Upstream Phase Splits	50-50,
Downstream Phase Splits	50-50, 60-40, 70-30
Offsets	at 5 seconds Intervals (0 to C)
Spacing	100, 200, 300 meters
Total Cases Studied	900
Number of NETSIM replications	10
Total number of NETSIM simulations	9000

Results and Verification. The computer program was run by using the above operating conditions to get the data base. Each of the above cases was simulated 10 times using TRAF-NETSIM to obtain average simulation results. A total of 9000 NETSIM runs were performed during the testing process. The study results were categorized and evaluated according to different operating conditions and are summarized in the following sections.

The throughput-offset relationship was examined in order to study the outputs from the PDX program. The effects of different volume-to-capacity ratios and intersignal spacings were studied. A very close relationship between PDX program and NETSIM simulation was observed during the comparison process. The results are shown in Figures D-12 through D-15. In each figure, results from the PDX Model and TRAF-NETSIM are shown in the same dimension for comparison purposes. Figure D-16 is the regression plot between NETSIM and PDX Model. The coefficient of regression between the two models was observed to be 0.85. Further research studies are underway to improve the PDX Model.

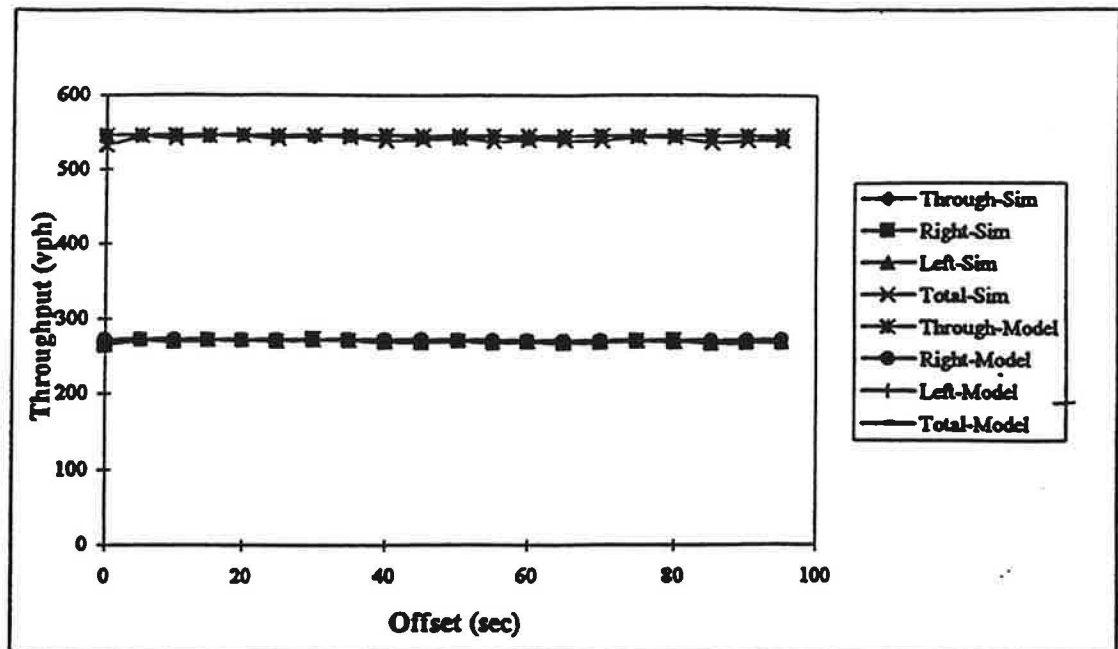


Figure D-12. Throughput-Offset Relationship between NETSIM and PDX Models for a spacing of 100 meters; v/c of 0.8 and Saturation Flow of 1900 vphgpl.

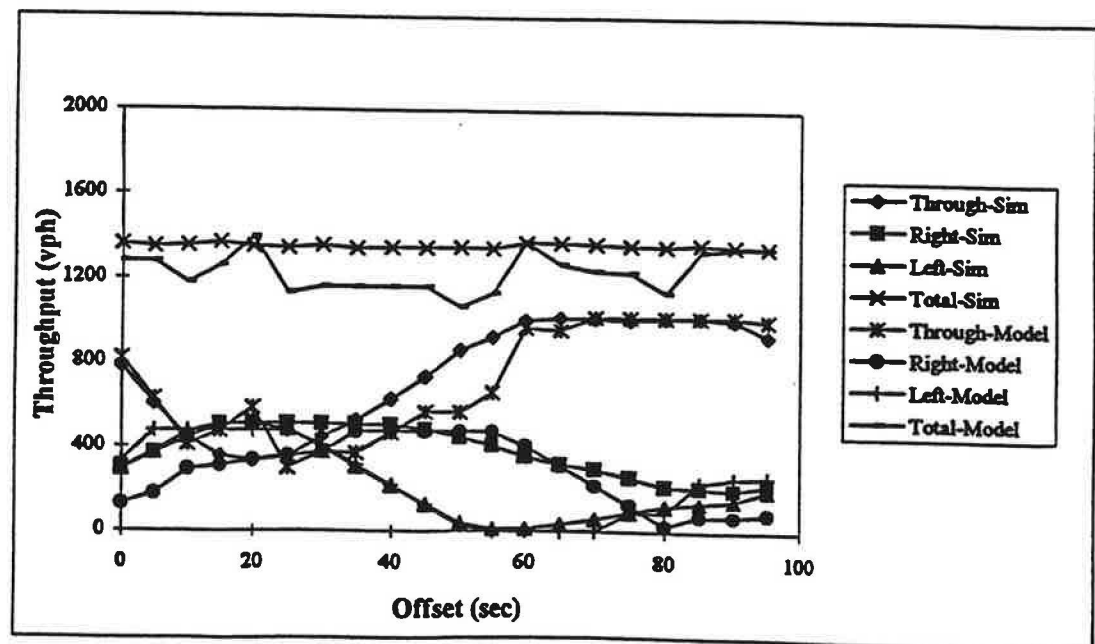


Figure D-13. Throughput-Offset Relationship between NETSIM and PDX Models for a spacing of 100 meters; v/c of 1.5 and Saturation Flow of 1900 vphgpl.

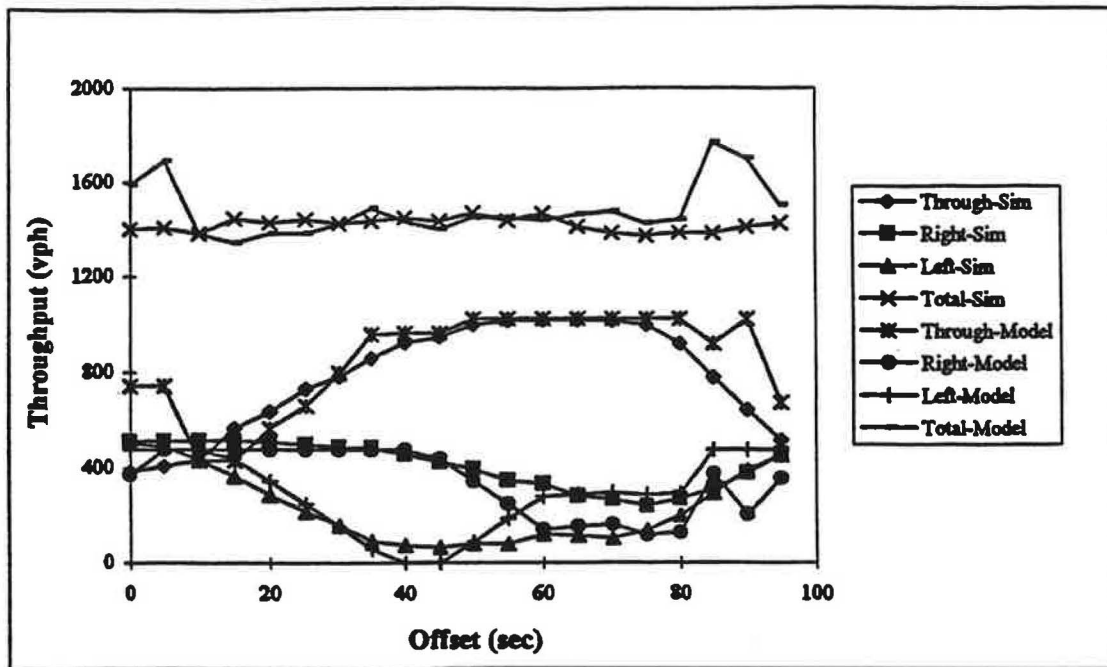


Figure D-14. Throughput-Offset Relationship between NETSIM and PDX Models for a spacing of 200 meters; v/c of 1.5 and Saturation Flow of 1900 vphpl.

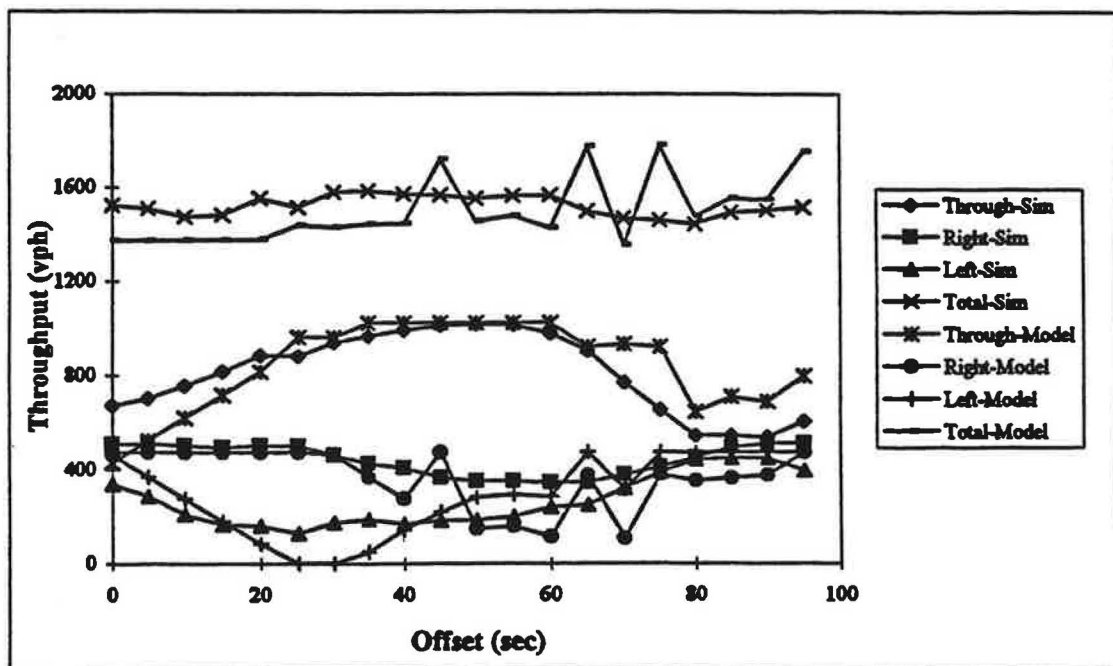


Figure D-15. Throughput-Offset Relationship between NETSIM and PDX Models for a spacing of 300 meters; v/c of 1.5 and Saturation Flow of 1900 vphgpl.

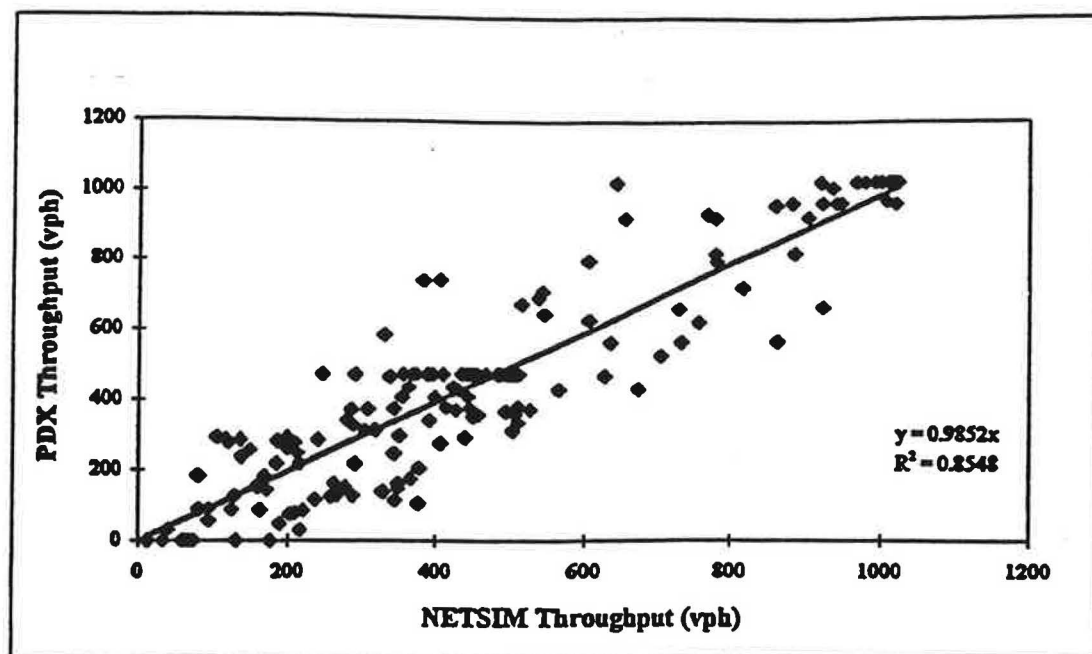


Figure D-16. Linear Regression between NETSIM and PDX Model Throughputs.

D.7.4. Discussion of the PDX Model

This computer program has been extensively calibrated against simulation results obtained from TRAF-NETSIM simulation program. The purpose of the calibration effort was to pinpoint the best fit parameter values used in this computer program to produce reasonable results. Traffic engineering judgement was also exercised during this process. Overall, the computer traffic program demonstrated good flexibility and accuracy in processing different types of traffic conditions based on the comparison results observed by the research team.

It should be noted that in the original Prosser-Dunne Model, traffic operating conditions were always assumed to be oversaturated, therefore, blocking would always occur because of insufficient service capacity at the downstream intersection. It was found, however, that blocking or queue spillback may also occur during undersaturated conditions given the limited storage spacing and bad offsets. Project study results have shown the most important factors that affect the estimation of queue spillback or blocking occurrence are downstream signal intersection's green time, the intersignal spacing (i.e., link length) and the volume-to-capacity ratio. Besides these factors, a critical spacing that defines the boundary condition of the occurrence of queue spillback was identified as a function of the downstream intersection's green time and volume to capacity ratio and other parameters. This new methodology helps define different types of problems based

on the varying nature of different operating conditions and renders the corresponding treatments. After applying these enhancements to the original Prosser-Dunne Model, a wide variety of real-world operating conditions can now be categorized and evaluated systematically by their specific types of problems, such as queue spillback due to inadequate storage spacing and/or oversaturation.

Several key parameters used in the PDX computer program were calibrated extensively under different operating conditions. Sufficient attention has been given to the effects of the selected values of the saturation flow density on the subsequent calculation of other variables. Because of its direct impacts on the estimate of saturation flow speed and interacting traffic wave speed, any change to the saturation flow density would result in different model outputs. So far, the parameter values used in the model have been calibrated to produce reasonable outputs compared with the simulation results from TRAF-NETSIM. Other calibrated parameters include the phase lost time and the unimpeded saturation flow rate. A vehicle unit length of 7.0 meters (23 feet) at jam density was used as the result of our nationwide data collection and analysis effort.

Another major improvement made to the computer traffic model was the introduction of the Queue Estimation Model. As an important part of traffic signal operation, especially during saturation periods, the behavior and characteristics of queuing traffic and its evaluation methodology have been studied by various researchers recently. The study approach used by the research team during the model development and calibration was to obtain queue traffic information dynamically throughout the cycle. A queue calculation submodule was provided to achieve this purpose. Similar to the capacity analysis methodology, the queue submodule performs second-by-second calculations to estimate the changing queue status at the downstream intersection. The useful information provided by this program can be subsequently analyzed or used in other traffic engineering models when studying the traffic operating conditions of a paired intersection.

The effects of available downstream travel distance to the back of queue was also provided in the PDX Model. As presented in Equation C-8, the saturation flow rate on green may be reduced by insufficient clear distance at start of green that permits platoon vehicles from accelerating to nominal saturation flow speeds. Thus, the clear period may have impeded saturation flow, but not blocked flow.

D.7.5 Existing Software Enhancements

Implementation of Queue-Interaction Models into internationally recognized computer programs is highly recommended. Some work toward this objective is known to be already underway (10,11,12).

REFERENCES

1. "Highway Capacity Manual." *Special Report 209*, Transportation Research Board, Washington, D.C., Third Edition, (1994).
2. Chang, E.C. and Messer, C.J. "PASSER II-90 Users Manual." Texas Transportation Institute, College Station, Texas, (1990).
3. Fambro, D. B., Chaudhary, N.A. and Messer, C.J. "PASSER III-90 User's Manual." Texas Transportation Institute, College Station Texas, (1991).
4. Wallace, C.E. and Courage, K.G. "TRANSYT-7F User's Manual." University of Florida, Gainesville, (1988).
5. May, A.D. *Traffic Flow Fundamentals*. Prentice-Hall, Englewood Cliffs, New Jersey, (1990) p.306.
6. Lighthill, M.J. and Whitham, G.B. "On Kinematic Waves: Part II, A Theory of Traffic Flow On Long Crowded Roads." *Proceeding of the Royal Society*, A2239, No. 1178, (1955).
7. Fambro, D.B., Rouphail, N.M., Sloup, P.R., Daniel, J.R., Li, J., Anwar, M., and R.J. Engelbrecht. "Highway Capacity Revisions for Chapters 9 and 11." Report No. FHWA-RD-96-088, Federal Highway Administration, Washington, D.C. (1996).
8. Leiberman, E.B, McShane, M.R., and Messer, C.J. Traffic Signal Control For Saturated Conditions. KLD Associates, Inc. NCHRP Project 3-38(3) Report, Vol 2., (1992) p. 15.
9. Prosser, N. and Dunne, M. "A Procedure for Estimating Movement Capacities at Signalised Paired Intersections." 2nd International Symposium on Highway Capacity, Sydney, Australia, (1994).
10. Rouphail, N.M. and Akcelik, R. "Paired Intersections: Initial Development of Platooned Arrival and Queue Interaction Models." Australian Road Research Board. Working Paper WD TE91/010, Vermont South, Australia (1991).
11. Akcelik, R., Besley, M., and Shepherd, R. "SIDRA (Windows) Input Redesign for Paired Intersection Modelling." Discussion Note to WD 96/008. Australian Road Research Board, Vermont South, Australia (1996).
12. Chaudhary, N.A. and Messer, C.J. "PASSER IV-96, Version 2.1, User/Reference Manual." Texas Transportation Institute, College Station, Texas. (1996).

APPENDIX E

ARTERIAL WEAVING SPEED MODELS

E.1 WEAVING SPEED MODEL

This section describes the development of a methodology for evaluating the performance of selected traffic movements in weaving sections on arterial cross streets in interchange areas. This performance was evaluated in terms of the speeds of both the weaving and non-weaving movements. The weaving maneuver that was examined in this study was the off-ramp right-turn movement that weaves across the arterial to make a left-turn at the downstream signalized intersection. Although several other weaving maneuvers exist in interchange areas, the off-ramp weave maneuver was generally found to have the largest volume and to be the most disruptive to arterial traffic flow.

Henceforth, the weaving problem described in this appendix is referred to as "arterial" weaving. This terminology is adopted to clearly indicate that the weaving studied in this research occurs on streets whose traffic flow is periodically interrupted by traffic signals, as compared to the more extensively studied weaving that occurs on uninterrupted flow facilities (such as freeways).

The following sections provide a brief overview of the weaving problem, a description of the weaving database, and the results of a model development and calibration process. Two maneuver speed models were developed. One model predicts the maneuver speed of weaving vehicles. The other model predicts the maneuver speed of arterial vehicles as they travel through the weaving section. The maneuver speed referred to here is effectively a running speed, as it represents the ratio of travel distance to travel time within the weaving section.

E.1.1 Background

The weaving maneuver studied in this research is somewhat comparable to a two-sided Type C freeway weave, as described in Chapter 4 of the *Highway Capacity Manual* (HCM) (1). As shown in Figure 4-5 of the HCM, the two-sided Type C weave maneuver entails an entry via a ramp, a weave across two or more through lanes, and an exit via a left-hand off-ramp. The off-ramp weaving maneuver considered in this study also enters via a ramp, weaves across two or more through lanes, and exits from the major street by turning left at the downstream intersection. However, there are some fundamental differences between freeway and arterial weaving. Freeway weaving sections operate under uninterrupted flow conditions while arterial weaving sections operate under interrupted flow due to upstream and downstream traffic signals. Another difference is that a freeway weaving section has a fixed length that is based on the distance between its entry and exit points while an arterial weaving section has a varying length as a result of downstream queues.

Weaving Section Classification. Maneuver (or travel path), weave type, and movement type are used in this research to denote different aspects of travel through a weaving section. An arterial weaving section can be classified according to the various maneuvers that pass through it. Maneuver denotes the path a vehicle takes from its point of entry to the point where it first experiences stopped delay or exits the weaving section. In other words, the maneuver begins when

the vehicle enters the weaving section and ends when the vehicle stops in a queue, stops at the stop line, or exits the weaving section, whichever event occurs first. Therefore, the maneuver distance varies from vehicle to vehicle depending on traffic conditions.

An arterial street segment bounded by signalized intersections has ten possible travel paths (or maneuvers), as defined by all possible origin-destination pairs in one travel direction through it. These ten paths are listed in Table E-1.

Table E-1. Classification of an arterial weaving section based on maneuver and weave type

Path No.	Maneuver (Entry to Exit)	Weave Type ¹	Turns Made	Lane Changes ²		Present in Study Sections ⁴	Model ⁵
				2 Lanes ³	3 Lanes ³		
Weaving-Related Paths							
1	Right-Turn to Left-Turn	1	2	1	2	Yes	Weaving
2	Left-Turn to Right-Turn	1	2	1	2	No	none
3	Right-Turn to Inside Through Lane	2	1	1	2	Yes	none
4	Left-Turn to Outside Through Lane	2	1	1	2	No	none
5	Outside Through Lane to Left-Turn	2	1	1	2	Yes	Arterial
6	Inside Through Lane to Right-Turn	2	1	1	2	Yes	Arterial
7	Through to Adjacent Through Lane	3	0	1	1 (or 2)	Yes	Arterial
Non-Weaving-Related Paths							
8	Right to Outside Through (or Right)	—	1 (or 2)	0	0	Yes	none
9	Left to Inside Through Lane (or Left)	—	1 (or 2)	0	0	No	none
10	Through to Same Through Lane	—	0	0	0	Yes	Arterial

Notes:

- 1 - Weave Type is determined by the number of turns and lane changes required (i.e., level of difficulty). Type 1 = maximum difficulty, Type 2 = medium difficulty, Type 3 = minimum difficulty.
- 2 - Number of arterial lanes crossed during the weave (excluding lane changes from acceleration lanes or lane changes to deceleration lanes or turn bays).
- 3 - Number of arterial lanes in the subject direction (excluding acceleration lanes, deceleration lanes, and turn bays).
- 4 - "Yes" indicates the path is present in arterial weaving sections in interchange areas.
- 5 - Applicable maneuver speed model.

The arterial segments considered in this research were bounded by an upstream off-ramp terminal and a downstream signalized intersection. Due to the nature of the off-ramp terminal, only seven of the ten possible paths existed. The seven travel paths that were present in these study sections are indicated by a "Yes" in Column 7 of Table E-1.

Travel paths that include a weave (i.e., change lane) are further categorized by their type. The weave type is defined by the difficulty of the weaving maneuver as expressed in terms of the number of turns and lane changes each weaving vehicle must make. Turns and lane changes increase the difficulty of negotiating a path because they have the most potential for reducing the

vehicle's speed by inducing periods of deceleration, acceleration, and merging. The Type 1 weave is the most difficult because it entails the most turns and lane changes. In contrast, the Type 3 weave is least difficult because it entails no turns and the fewest lane changes. The study sections considered for this research contain all three weave types. However, only the Type 1 weave was studied because of its characteristically high traffic volume in interchange areas.

Further classification of weaving-related maneuvers is possible by including secondary characteristics such as the type of movement required at each end of the path and its associated traffic control, angle of entry, and the frequency of *significant* acceleration or deceleration (due to a large speed change). Movement type describes how a vehicle enters or exits the weaving section (e.g., left, through, right). As shown in Table E-2, the type of traffic control in combination with the angle of entry (or exit) has a different effect on the speed of the corresponding movement. Specifically, the speeds associated with a weaving maneuver path will vary depending on whether the path always, sometimes, or never requires acceleration at its entry or exit.

Table E-2. Classification of weaving-related maneuvers based on movement type

Entry/Exit Movement	Traffic Control ¹	Angle of Entry/Exit ²	Frequency of Accel. at Entry	Frequency of Decel. at Exit
Left	Signal	Sharp	Always	Always
		Flat	Sometimes	Sometimes
Right	Signal	Sharp	Always	Always
		Flat	Sometimes	Sometimes
	Stop	Sharp	Always	
		Flat	Always	
	Yield	Sharp	Always	
		Flat	Sometimes	
	Uncontrolled	Sharp	Always	Always
		Flat	Never	Never
Through	Signal		Sometimes	Sometimes
	Uncontrolled		Never	Never

Notes:

- 1 - Traffic control for the entry or exit movement under consideration only. Independent of the traffic control for the other movements at the associated intersection.
- 2 - Angle of entry or exit is the deflection angle in the travel path. Sharp angles are about 90-degrees (turning radius of about 7.6 meters). Flat angles are about 30-degrees or less (turning radius of about 30.5 meters or more). Through movements are considered to be straight (no angle).

As indicated in Table E-2, some combinations of movement and control do not generally exist. In particular, stop and yield-controlled right-turn exits usually do not exist on arterials (a turning roadway for a right-turn exit is considered to be an uncontrolled exit).

Two maneuver speed models were developed in this research. A weaving speed model was created for predicting Path 1 speeds. An arterial speed model was created for predicting the speeds of Paths 5, 6, 7, and 10 combined. The weaving path, Path 1, is a Type 1 weave from the off-ramp to the arterial intersection left-turn exit. The arterial paths, Paths 5, 6, 7, and 10, are arterial through movement entries and are considered together despite the fact that technically three paths have a weaving component. Paths 3 and 8 were not studied in this research because they were not as problematic as the Path 1 weaving maneuver.

Literature Review. Many researchers have conducted studies of weaving sections on freeways and highways (2, 3, 4, 5, 6), but little research has been done concerning arterial weaving. A weaving section is defined by the HCM (1) as being formed by the crossing of two or more traffic streams where a merge is closely followed by a diverge. While the HCM acknowledges the existence of arterial weaving, it does not explicitly provide analysis procedures for this type of weaving; however, it does suggest that the freeway procedures can be used as an approximation.

In a recent study of single-point urban interchanges, Messer and Bonneson (7) noted the turbulence that occurs on the arterial cross street due to the high volume of weaving vehicles associated with interchange areas. They noted that this turbulence often restricted the capacity of both the cross street and the off-ramp.

Schoppert *et al* (8) devised a methodology for quantifying the level of service in arterial weaving sections. Their methodology was based on the travel times (or speeds) of arterial vehicles passing through the weaving section. They proposed quantifying this level of service using origin-destination tables to quantify the number of lane changes occurring in a weaving section. They hypothesized that the number of lane changes would be directly related to the ramp volume. If so, the number of lane changes could be directly related to the travel times (or speeds) of arterial vehicles. No field data were collected by Schoppert to calibrate this methodology.

Iqbal (9) studied weaving operations on arterial sections that were not bounded by signalized intersections and that had uncontrolled entry and exit maneuvers. He collected width, length, angle of approach, alignment deflection angle, speed, and volume data at twenty arterial weaving sections and used this data to calibrate weave and non-weave speed models similar to those in HCM. He demonstrated how his models yielded better estimates of the weave and non-weave speeds than did the freeway models provided in the HCM, even after he calibrated the HCM models with his data.

E.1.2 Data Collection

The data for this study were collected at six study sites in four states. Each study site consisted of a section of urban arterial located between a freeway off-ramp and a closely-spaced signalized intersection. During each field study, flow rates, travel times, travel distance, and stopped delays were measured for each of the seven paths identified by a "Yes" in Column 7 of Table E-1. Three of the four types of traffic control (i.e., signal, yield, and uncontrolled) are represented in the database for the off-ramp right-turn movement.

A video camera, elevated ten meters above the ground, was positioned at each end of the weaving section. The cameras faced toward each other in order to provide continuous surveillance of each vehicle entering the study section. The upstream camera was used to measure the entry time and movement type of each vehicle entering the study section. It was also used to measure the entry speed of each arterial vehicle. The downstream camera was used to measure the queue position and exit time all vehicles at the downstream arterial intersection. The cameras were synchronized in time to facilitate matching the entry and exit times for each vehicle. The field studies were conducted during the evening peak hour at each of the six sites. The study section was surveyed prior to each study to obtain its length and the width of various cross section elements. The details of the data collection effort are provided in Appendix B.

E.1.3 Model Development

The weaving database was analyzed using the Statistical Analysis System (SAS) (10). In all cases, the speed and flow rate data were aggregated into averages based on 15-minute analysis intervals. The analysis of the aggregated data consisted of using linear regression techniques to calibrate several candidate maneuver speed model formulations. Two models were ultimately identified as having the best fit to the data. One model predicts the weaving maneuver speed (i.e., Path 1 speed) and the other model predicts the arterial maneuver speed (i.e., average speed of Paths 5, 6, 7, and 10). The components of these two models are described below.

Maneuver Speed Models. The speed models developed for this research were based on empirical formulations that adhered to logical boundary conditions. The form of each model is similar; however, there is some variation in the model variables due to the differences in the priority assigned to the two vehicle classes (i.e., major and minor movement). The weaving speed model is:

$$U_{m,w} = b_0 U_a^{b_1} e^{(-b_2(1-P_U)V_w/3600)} \quad (E-1)$$

where:

$U_{m,w}$ = average maneuver speed for weaving vehicles, m/s;

U_a = average arterial speed entering the weaving section, m/s;

P_U = probability of a weaving vehicle being unblocked (i.e., able to change lanes freely);

V_w = average weaving flow rate, vph; and

b_0, b_1, b_2 = regression coefficients.

This model relates the weaving maneuver speed to the average speed of arterial vehicles entering the weaving section. This latter speed was measured as a spot speed at the point of entry to the arterial weaving section. Hence, it represents the "desired" speed of arterial drivers for the given arterial volume conditions when there is no weaving activity.

In theory, the weaving maneuver and arterial entry speeds would be similar when the weaving flow rate is negligible. However, the weaving speed would decrease as the weaving flow rate increases or when the arterial flow rate increases such that crossing opportunities decrease. The

model formulated in Equation E-1 incorporates this speed boundary and arterial flow rate sensitivity (via P_w). The regression coefficients b_0 and b_1 would logically be equal to unity; however, they are included in the model to allow for slight deviations from theory in light of practical differences in the speeds of weaving and arterial vehicles.

The arterial maneuver speed model is similar in form to that of the weaving speed model. Its form is:

$$U_{m,a} = b_3 U_a^{b_4} e^{(-b_5 V_a / 3600)} \quad (E-2)$$

where:

$U_{m,a}$ = average maneuver speed for arterial through vehicles, m/s;

U_a = average arterial speed entering the weaving section, m/s;

V_a = average arterial flow rate entering the weaving section, vph; and

b_3, b_4, b_5 = regression coefficients.

Like the rationale for the weaving maneuver speed model, the arterial maneuver speed model is based on the assumption that weaving speed should equal the arterial entry speed when the weaving flow rate is negligible. The variable for arterial flow rate is included in the model rather than weaving flow rate because it was found to be more strongly correlated with arterial maneuver speed. Logically, the two flow rates are positively correlated such that an increase in the arterial flow rate would likely be associated with an increase in weaving flow rate. Hence, the use of a surrogate variable for weaving flow rate that improved the quality was determined to be acceptable.

Dependent Variable. The dependent variable considered in these models is the maneuver speed. Maneuver speed is defined as the average running speed and represents the ratio of travel distance to travel time within the study section. The distance (and time) are measured from the point of entry to point where the vehicle first stops in a queue, stops at the stop line, or crosses the stop line and exits the study section, whichever is shortest (or occurs first). Therefore, the corresponding maneuver distance (and time) varies from vehicle to vehicle. It also varies among the two maneuver types as the weaving maneuver often has to accelerate from a stopped (or slowed) condition whereas the arterial maneuver generally enters the weaving section at speed. The maneuver speed for a weaving or arterial vehicle is calculated as follows:

$$u_{m,i} = \frac{d_{m,i}}{t_{m,i}} \quad (E-3)$$

$$= \frac{L_w - l_{q,i}}{(t_{exit} - t_{entry})_i - (t_{exit} - t_{stop})_i}$$

where:

$u_{m,i}$ = maneuver speed for vehicle i , m/s;

$d_{m,i}$ = maneuver distance for vehicle i , m;

- $t_{m,i}$ = maneuver time for vehicle i , sec;
- L_w = distance between the off-ramp entry point and the stop line of the downstream intersection (i.e., the length of the weaving section), m;
- $l_{q,i}$ = length of queue joined by vehicle i , m; and
- t_j = time of event j for vehicle i (j = entry, stop, exit), sec.

It should be noted that the maneuver speed is based on the time and location of the vehicle when it stops for the *first* time. No subsequent periods of motion, such as when moving up in or discharging from a queue, are incorporated into the maneuver speed.

The maneuver speed computed with Equation E-3 does not explicitly account for the extra time used during acceleration or deceleration at the beginning or end of the maneuver. This distinction is particularly important for the weaving maneuver. As indicated in Table E-2, the weaving movements considered in this analysis can "always," "sometimes," and "never" experience acceleration (or deceleration) upon entry (or exit), depending on their traffic control and angle of entry. As a result, there is more variability in the weaving maneuver speed data than in the arterial maneuver speed data.

Independent Variables. One of the independent variables used in the weaving maneuver speed model is the "probability of a weaving vehicle being unblocked" P_U . This variable relates to the portion of time that the end of the off-ramp (i.e., the beginning of the weaving section) is not blocked by the passing of the arterial traffic stream or the spillback of a downstream queue. Blockage by the passing stream is represented by the formation of platoons in the arterial traffic stream (induced by signalization or random bunching) which can delay the weaving vehicle.

A distinction is made in the computation of P_U based on the length of the weaving section. Specifically, weaving sections are referred to as "short" or "long" depending on the nature of the weaving maneuver. A "long" weaving section is defined as a section that has sufficient length to allow a weaving driver to weave across the arterial one lane at a time. The minimum length of a "long" section is computed as the product of the length needed for a single-lane lane change (estimated as 90 meters) and the number of arterial traffic lanes being crossed (excluding the first). If this minimum length is not available, then the section is referred to as "short." This designation implies that the weave maneuver requires the simultaneous crossing of all arterial lanes (in the subject direction) using more of a "crossing" than a "lane-changing" action.

Based on the preceding definitions, P_U for short weaving sections is based on there being a gap in the arterial traffic stream that is sufficiently large that the off-ramp driver can cross all arterial traffic lanes at one time. For long sections, P_U is based on there being an acceptable lane-change gap in each lane as it is crossed one lane at a time. An equation was derived to estimate P_U based on the preceding definition using an expected value approach and an assumed random arrival distribution; however, the resulting equation was intractable and felt to be overly complicated. Thus, a simplified form of the model was developed based on a queueing theory representation. This equation is:

$$P_U = \left(1 - \frac{V_a}{s_l N_l} \right)^{N_l} (1 - I_L) + \left(1 - \frac{V_a}{s_l N_l} \right) I_L \geq 0 \quad (\text{E-4})$$

where:

- P_U = probability of a weaving vehicle being unblocked;
- V_a = average arterial flow rate entering the weaving section, vph;
- s_l = saturation flow rate per lane under prevailing conditions ($\approx 1,800$), vphpl;
- N_l = number of arterial through lanes in the subject direction, lanes;
- I_L = indicator variable (1.0 if $D_{m,w} > 90 (N_l - 1)$, 0.0 otherwise);
- $D_{m,w}$ = average maneuver distance for weaving vehicles ($= L_w - L_{q,w}$), m; and
- $L_{q,w}$ = average length of queue joined by weaving vehicles, m.

It should be noted that P_U is equal to 0.0 when the average queue length equals the length of the weaving section (i.e., when $D_{m,w}$ is effectively zero).

In this study, the length of the queue that was joined was measured for each weaving vehicle. In most cases, the queue was joined when it was near its cyclic maximum because most weaving activity occurred after the arterial through movement had passed by the off-ramp. Of course, the queue would continue to grow as other weaving vehicles joined it. For practical applications of Equation E-4, $L_{q,w}$ can be conservatively estimated as the average maximum queue length per cycle at the downstream intersection based on the arterial flow rate entering the weaving section V_a . In this regard, the following equation can be used to estimate the average queue length joined by a weaving vehicle:

$$L_{q,w} = \frac{V_a r}{1 - \frac{V_a}{s_l N_l}} \times \frac{L_v}{3600} \quad (\text{E-5})$$

where:

- r = effective red time for the downstream intersection through movement; sec; and
- L_v = average lane length occupied by a queued vehicle (see Equation C-45), m/veh.

A second independent variable used in the weaving speed model is the average weaving flow rate V_w . This flow rate represents the number of vehicles that perform the Path 1 weave during the analysis period; it is a subset of the number of vehicles that enter the arterial via the off-ramp. The probability of there being a *blocked* condition (i.e., $1 - P_U$) and the average weaving flow rate V_w are multiplied together to obtain the expected number of weaving vehicles that will be blocked by arterial traffic. This product was included in the weave maneuver speed model because the maneuver speed was found to decrease as the number of blocked vehicles increased.

E.1.4 Database Summary

Table E-3 shows the range of values in the weaving database for the independent and dependent variables included in the maneuver speed models. The variables are listed according to the applicable model. This table demonstrates the range over which each model is considered valid. In general, the models were calibrated with sites having two or three arterial through lanes (in the subject direction), closely spaced intersections, and a wide range of arterial flow rates.

Table E-3. Range of independent and dependent variables

Model	Variable	Variable Name	Units	Minimum ²	Maximum ²
Both ¹	N_i	Arterial through lanes in the subject direction	—	2	3
	L_w	Length of the weaving section	m	141.7	265.2
	U_a	Average arterial entry speed	m/s	8.9	21.8
	V_a	Avg. arterial flow rate entering the weaving section	vph	640	1,924
	V_a/N_i	Average arterial flow rate per lane	vphpl	320	676
Weaving	$l_{q,w}$	Queue length joined by a weaving vehicle	m	0.0	160.0
	P_U	Probability of a weaving vehicle being unblocked	—	0.27	0.82
	V_w	Average weaving flow rate	vph	88	270
	$d_{m,w}$	Maneuver distance for a weaving vehicle	m	27.4	265.2
	$t_{m,w}$	Maneuver time for a weaving vehicle	s	4	62
	$u_{m,w}$	Weaving maneuver speed	m/s	2.8	18.7
Arterial	$l_{q,a}$	Queue length joined by an arterial vehicle	m	0.0	209.7
	$d_{m,a}$	Maneuver distance for an arterial vehicle	m	18.0	265.2
	$t_{m,a}$	Maneuver time for an arterial vehicle	s	6	37
	$u_{m,a}$	Arterial maneuver speed	m/s	1.9	23.4

Notes:

- 1 - Variables used by both the weaving and arterial maneuver speed models.
- 2 - All average values are based on a 15-minute intervals.

Table E-3 illustrates the wide range of conditions over which the maneuver speed models are applicable. Specifically, the weaving maneuver speeds range from 2.8 to 18.7 m/s and the arterial maneuver speeds range from 1.9 to 23.4 m/s. These ranges do not contain zero because, by definition, a vehicle could not complete the subject maneuver without having some speed. The low speeds in these ranges were observed during spillback conditions where the subject vehicle creeps forward to join the queue (and thus complete its maneuver).

The maximum speeds in the two aforementioned speed ranges are about equal to the maximum arterial entry speed. This agreement was generally observed during low off-ramp volume conditions. The maximum arterial maneuver speed, a running speed, is slightly higher than the

maximum arterial entry speed, a spot speed, because unimpeded arterial vehicles could enter the study section at speed and then accelerate along the length of the section to a speed higher than the entry speed. The minimum value of the arterial entry speed range was observed during conditions of extensive downstream queueing such that the arterial vehicles were moving slowly as they decelerated to join the back of queue.

E.1.5 Calibrated Models

The model forms developed in a previous section were calibrated using the weaving database, as summarized in Table E-3. The SAS analysis regression procedure was used to determine the regression parameter coefficients for each model. The resulting calibrated weaving maneuver speed model is:

$$U_{m,w} = 3.741 U_a^{0.408} e^{(-11.045 (1 - P_v) V_a / 3600)} \quad (E-6)$$

The resulting calibrated arterial maneuver speed model is:

$$U_{m,a} = 1.986 U_a^{0.717} e^{(-0.634 V_a / 3600)} \quad (E-7)$$

where:

$U_{m,w}$ = average maneuver speed for weaving vehicles, m/s; and

$U_{m,a}$ = average maneuver speed for arterial through vehicles, m/s.

Table E-4 lists several statistics that indicate the quality-of-fit for each maneuver speed model. As these statistics suggest, the weaving maneuver speed model (i.e., Equation E-6) accounts for 16 percent of the variability in the maneuver speed data. Likewise, the arterial maneuver speed model (i.e., Equation E-7) accounts for 22 percent of the variability in the data. In all cases, the independent variables included in the model were found to be strongly correlated with maneuver speed. All tests were conducted with a 95 percent level of confidence. The root mean square error (or standard error) of each model, combined with the number of observations, yields a minimum precision of ± 0.10 and ± 0.16 m/s for estimates of the average weaving and arterial maneuver speeds, respectively.

Table E-4. Maneuver speed model statistics

Maneuver Speed Model	Observations	R ²	Root Mean Square Error	Precision
Weaving	421	0.16	2.02 m/s	± 0.10 m/s
Arterial	324	0.22	2.94 m/s	± 0.16 m/s

The low R^2 values listed in Table E-4 are due partly to limitations of the models and partly to random sources. As discussed previously, these models do not include the acceleration or deceleration component of the maneuver in the maneuver speed calculation. This omission has a larger impact on the weaving maneuver speed model because many weaving vehicles accelerate from the off-ramp (i.e., they arrive on the red indication or have to yield) while others enter at speed. In contrast, arterial vehicles almost always enter the study section at speed. Another factor that affects the R^2 of the arterial maneuver speed model is the inclusion of vehicles on several paths (i.e., Paths 5, 6, 7, and 10). Future enhancements to these models to account for acceleration, deceleration, and individual paths would likely improve the quality-of-fit.

Figure E-1 compares the weaving maneuver speeds predicted by Equation E-6 with those measured in the field. Each data point represents the average of ten observations (sorted by predicted speed) in order to reduce the overlapping density of 421 observations. It should be noted that individual observations were used in the regression analysis whereas the averaged data were used only for the graphical examination of trends. The data in Figure E-1 indicate that there is a general trend of agreement between the predicted and measured speeds over the range of measured speeds.

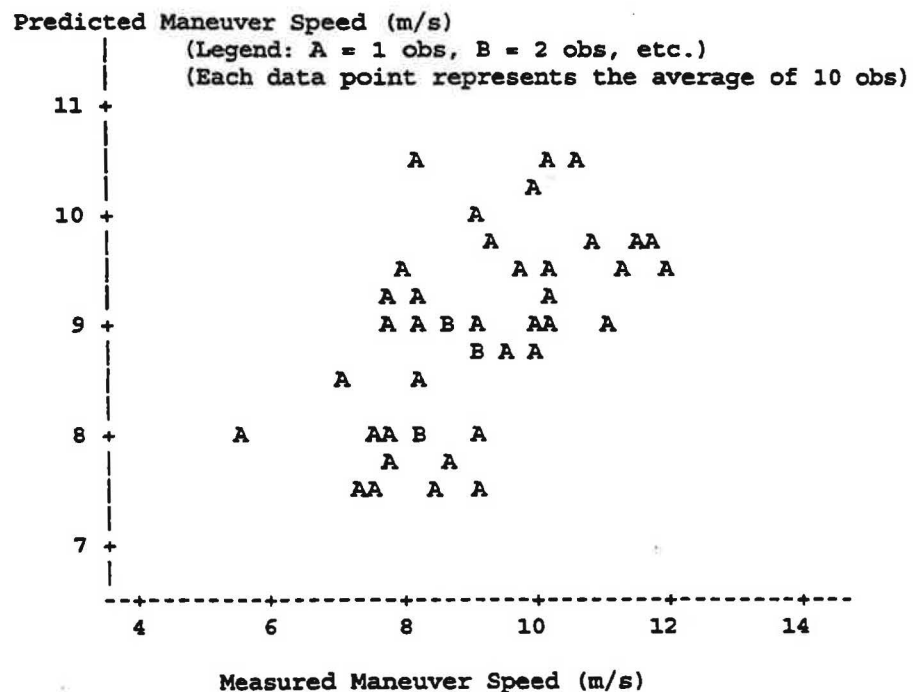


Figure E-1. Predicted versus measured weaving maneuver speed.

Figure E-2 compares the arterial maneuver speed predicted by Equation E-7 with that measured in the field. Each data point represents the average of ten observations (sorted by predicted speed) in order to reduce the overlapping density of 324 observations. As with the preceding figure, there is a general trend of agreement between the predicted and measured speeds. In fact, the trend of agreement is probably a little better than that for the weaving speed model.

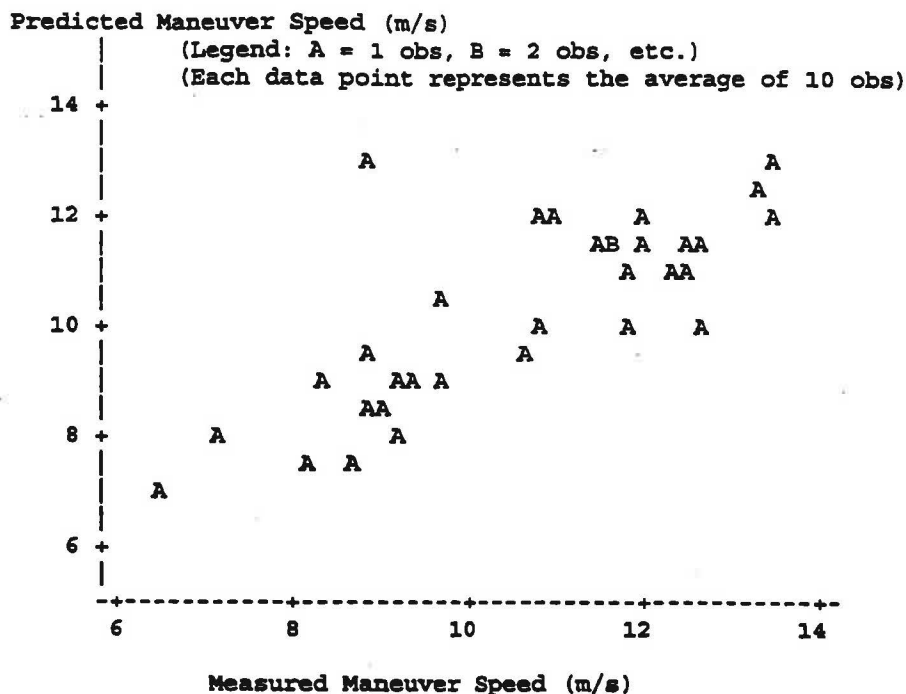


Figure E-2. Predicted versus measured arterial maneuver speed.

E.1.6 Sensitivity Analysis

The effects of arterial flow rate, weaving flow rate, arterial traffic lanes, and weaving section length on maneuver speed are examined in this section. Initially, the two maneuver speed models are compared to one another. Then, a sensitivity analysis is conducted for each model.

Figure E-3 illustrates the effect of arterial flow rate on weaving and arterial maneuver speed. This figure shows the behavior of both models when all other factors are held constant. The values selected for these factors represent their respective average values as found in the database. The range of flow rates over which the two models are compared is larger than the corresponding range in the database. This extension was undertaken to show the overall behavior of each model when extrapolated to extreme (but realistic) values.

The trends in Figure E-3 show that both models predict an exponentially decreasing maneuver speed with increasing arterial flow rate. This trend is somewhat consistent with the traditional speed-flow relationship for uninterrupted traffic streams in uncongested conditions. This figure also shows that the arterial maneuver speed is always higher than the weaving maneuver speed for the same flow rate. This trend is reasonable since the arterial vehicles enter the weaving section at speed while the weaving vehicles often must accelerate from a stopped (or slowed) condition when departing the off-ramp. The trend toward convergence of the two models at higher flow rates is also reasonable as the weaving maneuver speed should approach the arterial maneuver speed as the capacity of the weaving section is neared.

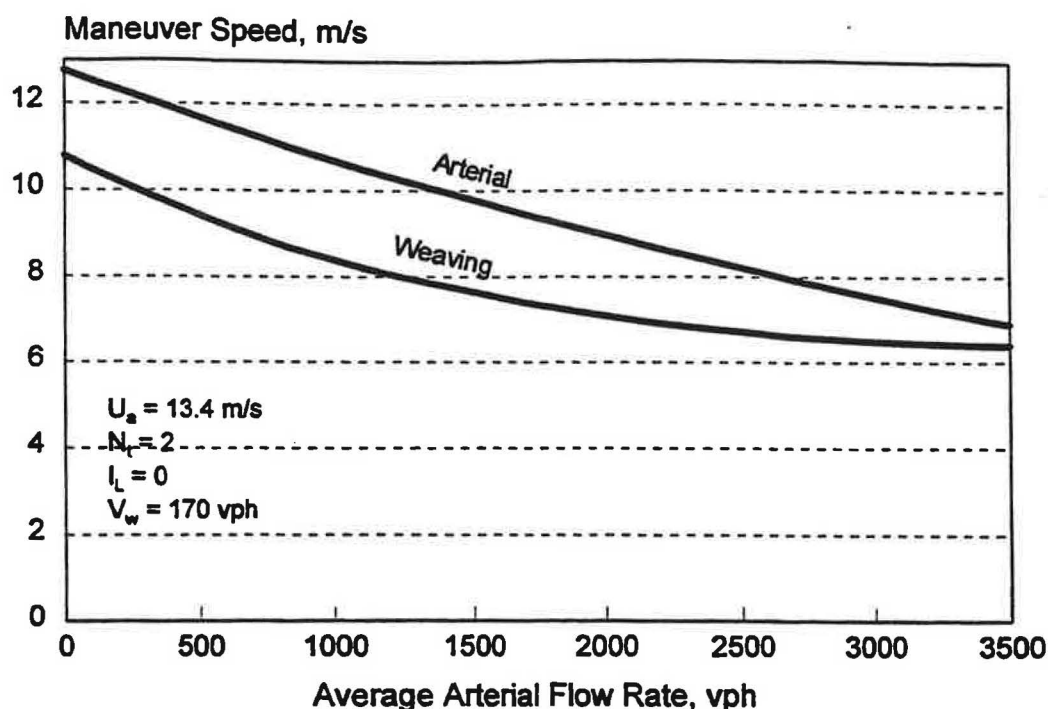


Figure E-3. Effect of arterial flow rate on weaving and arterial maneuver speeds.

Figures E-4, E-5, and E-6 illustrate the effect of arterial lanes, weaving flow rate, and section length, respectively, on weaving maneuver speed. In each figure, one factor is varied while all others are held constant. In all cases, the average arterial entry speed S_a is held constant at 13.4 m/s. Also, the range of flow rates used in these figures is representative of that found in the database.

Figure E-4 illustrates the relationship between weaving maneuver speed and arterial lanes. This comparison applies to the situation where the weaving section can be described as “short” relative to the distance available for the weaving maneuver. In general, the weave in a short section is described as a simultaneous crossing of all arterial lanes at one time.

As Figure E-4 indicates, the weaving maneuver speed is higher for arterials with three lanes in the subject direction, as compared to those with two lanes. This trend is the result of two effects. First, it reflects the effect of the acceleration component relative to the distance traveled. In general, the average speed will be higher when the distance traveled during the weave (i.e., three lanes versus two lanes) is longer such that the effect of the initial acceleration becomes less significant on the overall travel time. Second, it reflects the effect of lower arterial lane flow rates. Presumably, the lower lane flow rates yield larger headways which would be less of a hindrance to a weaving vehicle.

Figure E-5 illustrates the relationship between weaving maneuver speed and weaving flow rate. All other factors being equal, the higher the weaving flow rate, the lower the average weaving maneuver speed. This trend is reasonable and suggests that the capacity of the weaving section is dependent on the combined arterial and weaving flow rates.

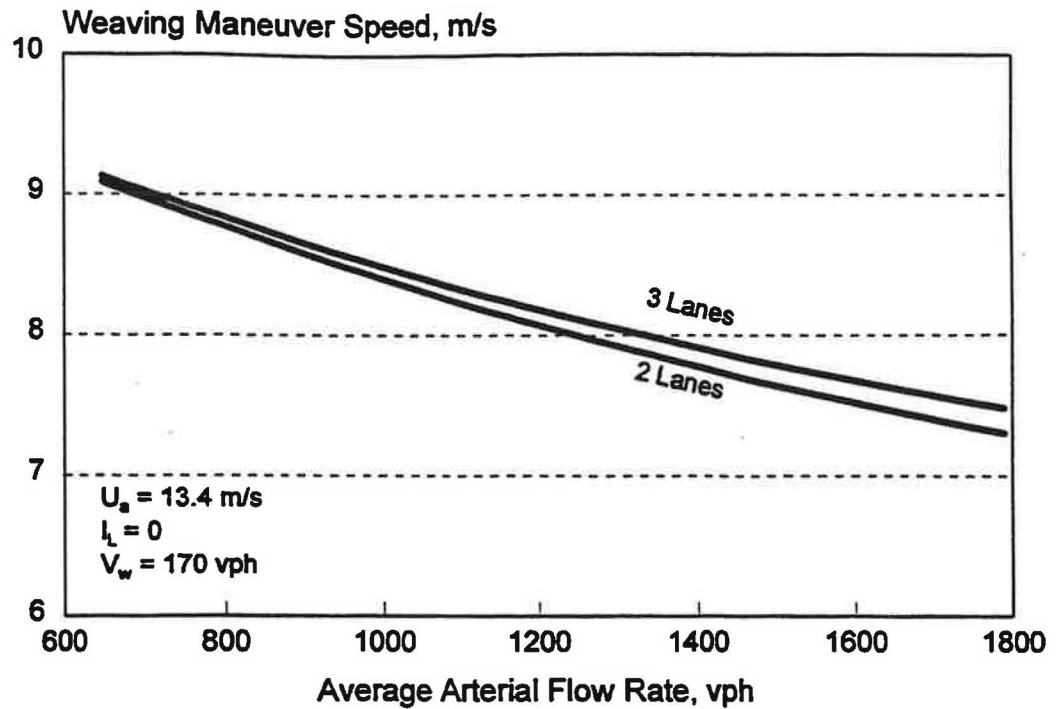


Figure E-4. Effect of number of arterial lanes on weaving maneuver speed.

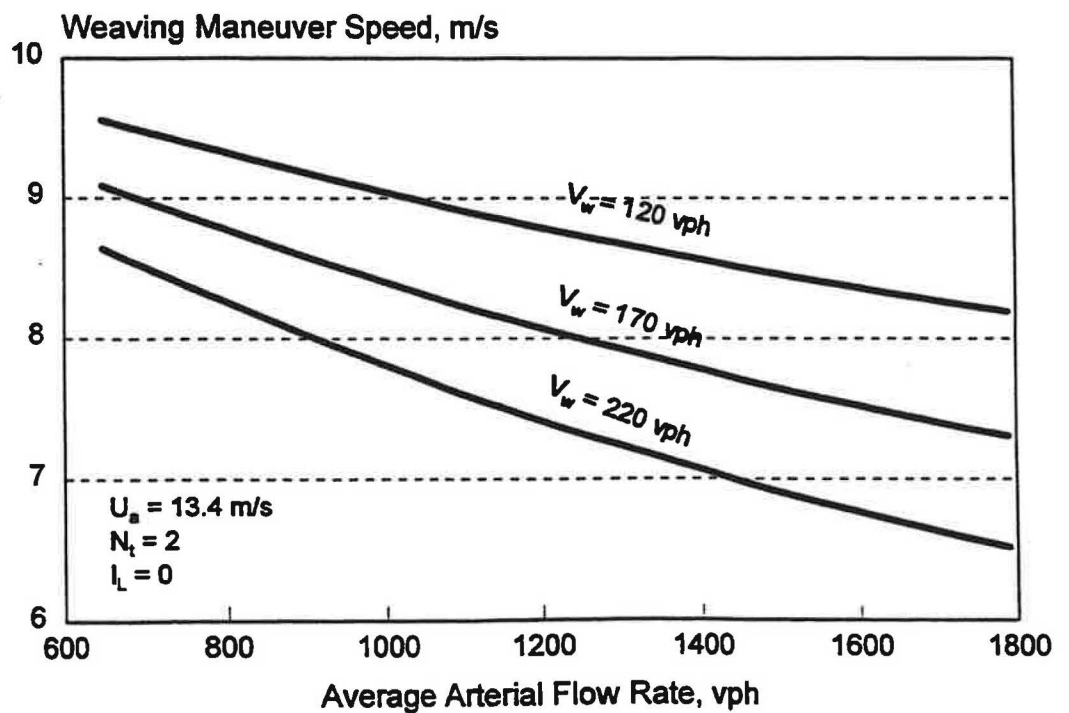


Figure E-5. Effect of weaving flow rate on weaving maneuver speed.

Figure E-6 illustrates the relationship between weaving maneuver speed and the length of the weaving section. All other factors being equal, the weaving maneuver speed is higher in a "long" weaving section than it is in a "short" section. This trend is reasonable since a vehicle can attain a higher speed when it has a longer distance in which to maneuver.

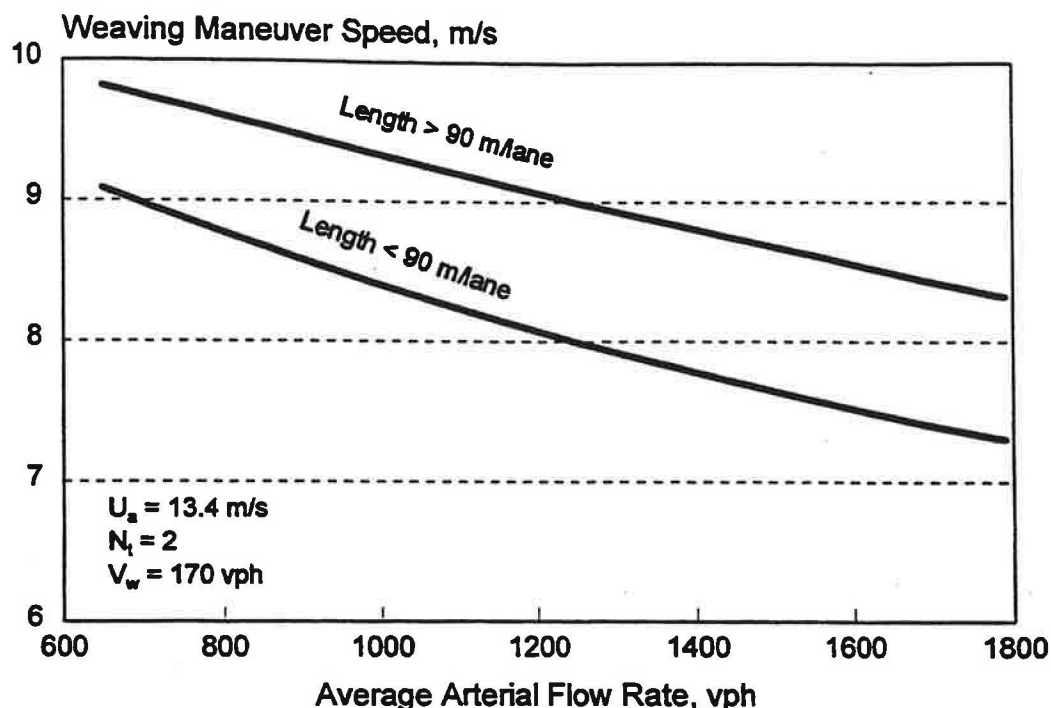


Figure E-6. Effect of weaving section length on weaving maneuver speed.

E.1.7 Summary

An arterial weaving section can be classified by the maneuver (or travel path) through it, the type of (or difficulty associated with) the weave, and the characteristics of entry and exit movements for the weave-related travel paths. These factors can be used to explain the behavior of vehicles in the different travel paths and their corresponding maneuver speed.

The results of the literature search indicated that relatively little research has been devoted to the problem of arterial weaving. While the HCM (1) methods may have application to the analysis of arterial weaving sections with uninterrupted flow conditions, they do not appear to be sensitive to the effects of up and downstream signals on arterial weaving performance. This research was undertaken to quantify the effects of traffic signals on the speed of weaving and arterial traffic streams. The results of this research are focused on a type of weaving that is commonly found on the arterial cross street in interchange areas; however, the modeling approach is also applicable to any arterial weaving section in the vicinity of traffic signals.

Two models were developed in this research to predict the maneuver speed of vehicles passing through an arterial weaving section. Both of these models include the average arterial entry speed as an independent variable. This speed represents the "desired" speed of arterial drivers for the given arterial volume conditions. It is the speed that would be maintained when there are no weaving vehicles. The model calibration revealed that both the weaving and arterial maneuver speeds are dependent upon the arterial flow rate. The model calibration also indicated that weaving maneuver speed is dependent on the weaving flow rate, weaving section length, and the probability of the weaving vehicles being blocked. Separately, both models show strong correlations with the measured maneuver speeds and can be used to estimate the mean maneuver speed with a precision of less than ± 0.2 m/s.

A sensitivity analysis using the calibrated models indicated that both weaving and arterial maneuver speeds decreased with increasing arterial flow rate. This trend reflects the increased turbulence between the arterial and weaving movements associated with an increase in arterial flow rate. The sensitivity analysis also indicated that weaving maneuver speed decreased with increasing weaving flow rate and shorter weaving section length.

These two models can be used to predict the operational performance of other weaving sections in interchange areas with characteristics similar to those studied. While these models are limited to a particular weave maneuver (i.e., Path 1), similar models could be developed for different paths and characteristics. It is recommended that future research be conducted to enhance the models developed in this research to include other travel paths and the effects of acceleration and deceleration at the entry and exit points.

REFERENCES

1. *Special Report 209: Highway Capacity Manual*, 3rd ed., TRB, National Research Council, Washington, D.C. (1994).
2. Ostrum, B.K., Leiman, L., and May, A.D., *Suggested Procedures for Analyzing Freeway Weaving Sections*. In *Transportation Research Record 1398*, TRB, National Research Council, Washington, D.C. (1993) pp. 42-48.
3. Vermijs, R., and Schuurman, H., "Evaluating Capacity of Freeway Weaving Sections and On-Ramps using the Microscopic Simulation Model FOSIM." In *Proceedings of the Second International Symposium on Highway Capacity, Volume 2*, (Ed: Akcelik, R.), ARRB Transport Research Ltd, Vermont South, Victoria, Australia (1994) pp. 651-671.
4. Windover, J.R., and May, A.D., "Revisions to Level D Methodology of Analyzing Freeway Ramp Weaving Sections." In *Transportation Research Record 1457*, TRB, National Research Council, Washington, D.C. (1994) pp. 43-49.
5. Fazio, J., "Geometric Approach to Modeling Vehicular Speeds through Simple Freeway Weaving Sections." *ITE Journal*, Institute of Transportation Engineers, Washington, D.C., (April 1988) pp. 41-45.
6. Wang, M., Cassidy, M.J., Chan, P., and May, A.D., "Evaluating the Capacity of Freeway Weaving Sections." *Journal of Transportation Engineering*, Vol. 119, No. 3, American Society of Civil Engineers, New York, New, York (1993) pp. 360-384.
7. Messer, C.J., Bonneson, J.A., Anderson, S.D., and McFarland, W.F., *NCHRP Report 345: Single Point Urban Interchange Design and Operations Analysis*, TRB, National Research Council, Washington, D.C. (1991).
8. Schoppert, D.W., Kittelson, W., and Shapiro, S., *Quality of Flow in Urban Arterials-Phase I*. Report No. FHWA-RD-78-199, Federal Highway Administration, Washington, D.C. (1978).
9. Iqbal, M.S., "Analytical Models of Weaving Area Operations Under Nonfreeway Conditions." *ITE Journal*, Vol. 65, No. 7, Institute of Transportation Engineers, Washington, D.C. (July 1995).
10. *SAS/STAT User's Guide, Version 6*, 4th ed., SAS Institute, Cary, North Carolina (1990).

APPENDIX F

DEVELOPMENT OF INTERCHANGE ANALYSIS SOFTWARE

F.1 INTRODUCTION

There are various operational software, such as PASSER II - III, HCS, and TRANSYT-7F, which can analyze the performance of different interchanges. Although each software is capable at evaluating the performance of particular interchanges, none are well adapted to making a comparison among interchange alternatives because input must be re-generated for each interchange being evaluated. This appendix describes the proposed development of a prototype interchange analysis software, called INTERCHANGE, which can evaluate the operational performance of various two-level signalized interchanges within a single standalone software package or using an interface with existing software previously mentioned. The program will adapt the recommended operational analysis procedures developed for intersections and interchanges which was described earlier in section 3.3.2 and shown in Figure F-1.

The main benefit of INTERCHANGE is that it only requires traffic input for one interchange configuration and it automatically converts and adapts that input to all other interchanges selected by the user. This feature facilitates the process of comparing different interchange types and allows comparisons of output to be easily made as well. The following sections describe the structure of the proposed software and its potential capabilities. Also discussed is the program's current status of development and future work plans. Finally, an example problem is given to show the practical uses of INTERCHANGE and benefits to the interchange selection process.

F.2 PROCEDURAL DESIGN

As discussed previously in section 3.3.4, two optimal procedural designs of the program were proposed and are shown in Figure F-2. One proposed procedure involves using INTERCHANGE as an input and conversion software to be used in conjunction with existing software. The other procedural design creates a standalone program capable of analyzing all operational aspects of the various interchanges being compared. The left side of the diagram depicts the software design option which uses existing software to perform the analysis. As a first step, the input for one interchange form is entered which includes the turning movement volumes for one interchange form, the geometric and signalization conditions, the type of analysis requested, and the interchange types to analyze. Once these data are entered the program performs a conversion analysis which uses a database conversion algorithm discussed later.

The proposed enhancements to INTERCHANGE would produce data unique to that interchange such as turning movement volumes and required geometry for each alternative form. This converted data would then be automatically input into existing software, such as HCS, PASSER-II, and TRANSYT-7F, for an operational analysis of each interchange being considered.

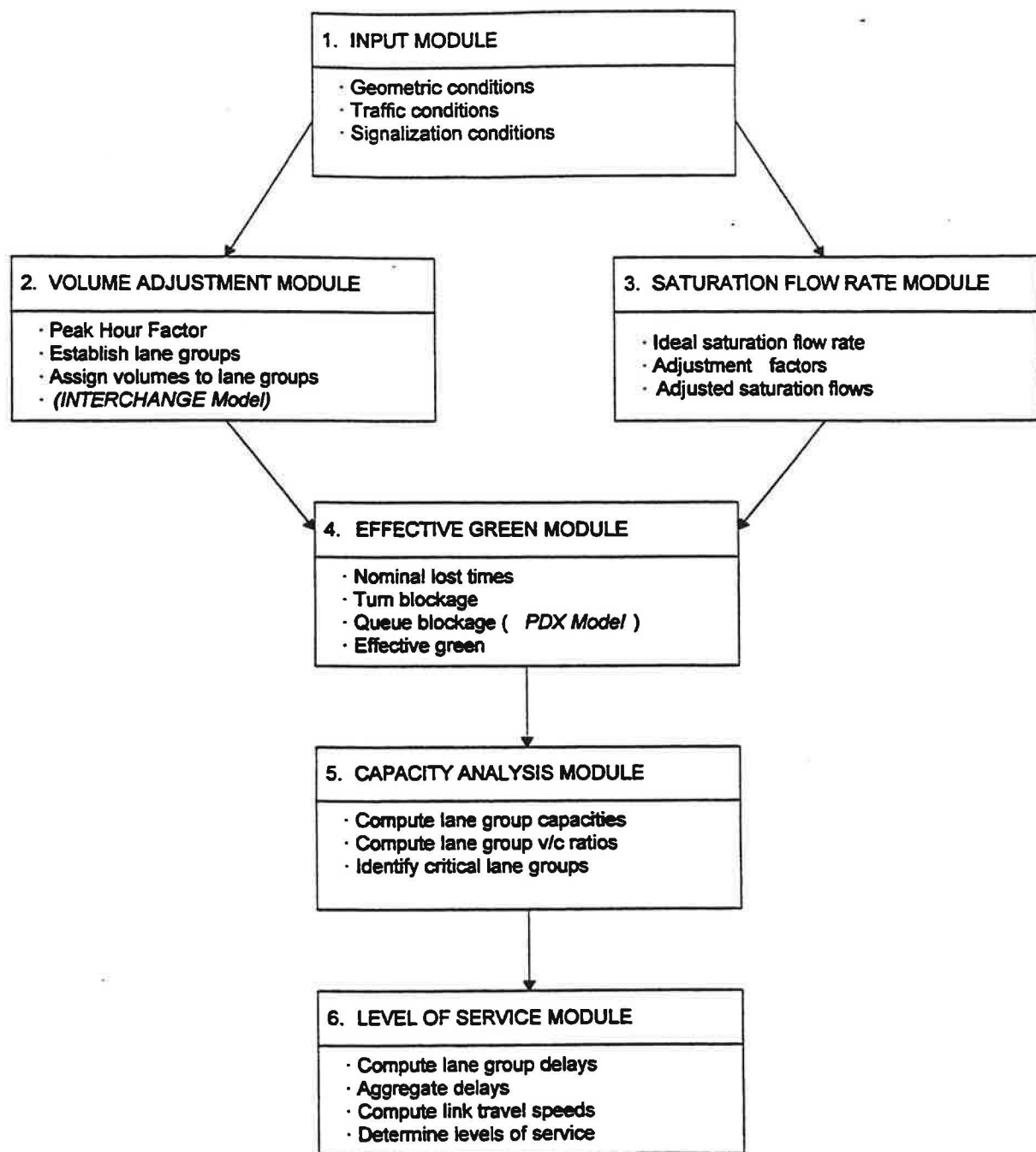


Figure F-1. Recommended interchange operational analysis procedure (1).

INTERCHANGE Procedural Design

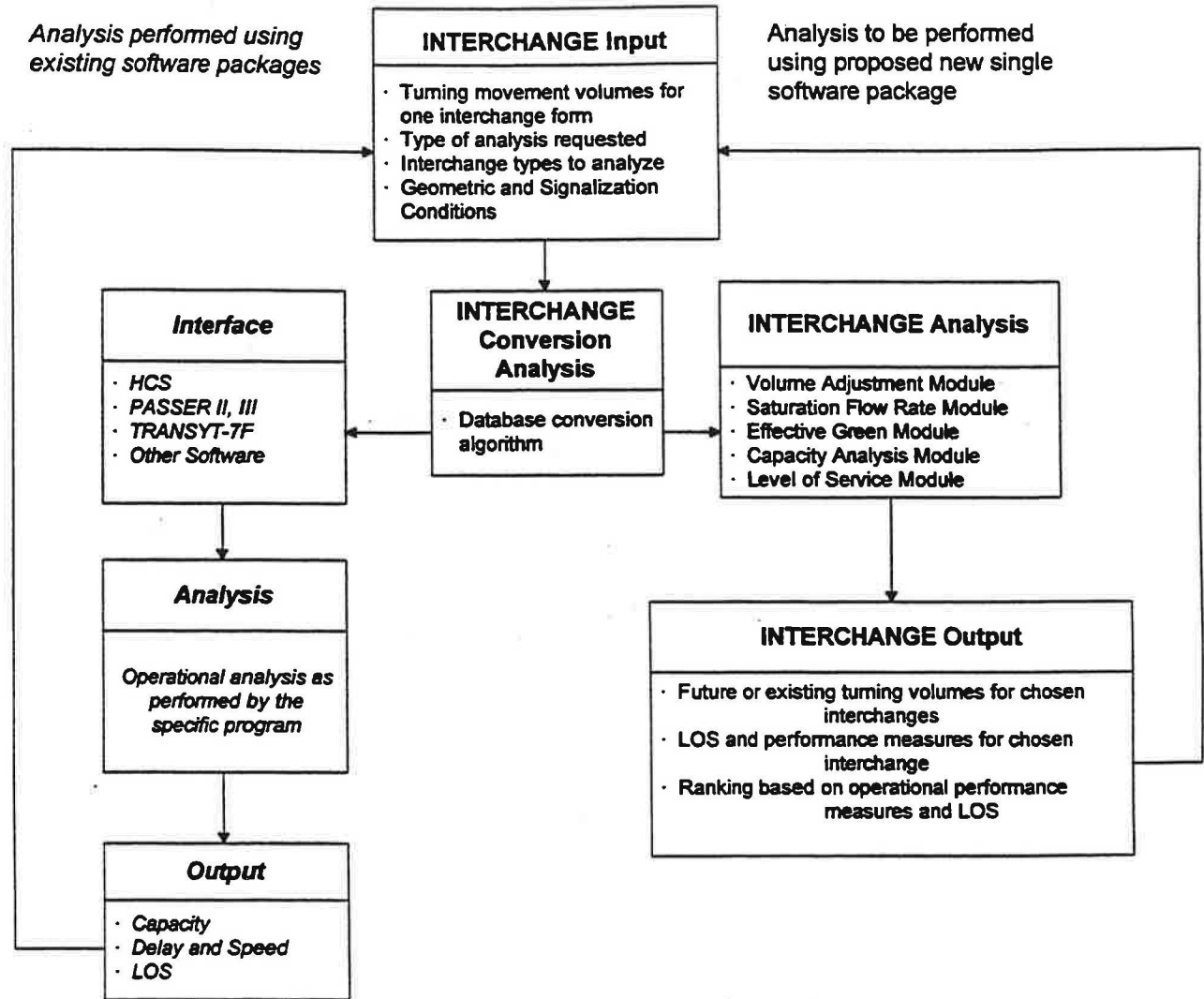


Figure F-2. Flow diagram of optional procedural design for INTERCHANGE.

The output from these existing software would then be viewed through INTERCHANGE to create a common output screen for ease of viewing and analyzing the performance measures. A common database of output values could be stored within INTERCHANGE to facilitate the comparisons. As a last step, the output could be used to revise the initial design assumptions to further optimize the initial design.

As a standalone program, the proposed software would perform the same input and conversion analysis; however, the analysis would be performed within the program itself. The volume adjustment module, saturation flow rate module, effective green module, capacity analysis module, and level of service would all be computed within INTERCHANGE. After performing the analysis on each interchange selected, the program could produce outputs displaying performance measures for each selected configuration individually or together for making easy comparisons. The output could then be tailored to produce useful results for different users.

F.3 TURNING MOVEMENT VOLUME CONVERSION

Whether performing an operational analysis of intersections or interchanges, one of the first steps is inputting turning movement volumes into the existing operational software. To efficiently analyze and compare the capacity and performance of various proposed interchange alternatives with an existing interchange or intersection, a methodology needed to be developed for converting volumes from the existing condition to the proposed alternatives. Current practice involves converting the volumes manually and inputting them into operational programs, such as PASSER-II, HCS, and TRANSYT-7F. No software exists which specifically converts volumes among the various interchange types.

Figure F-1 will be used to illustrate the methodology for converting the turning movement volumes. After the existing traffic conditions have been entered in the input module, part of the volume adjustment module can be used to convert these traffic or turning movement volumes to other interchange forms for further capacity and measures of performance comparisons. Other input data such as geometries and signalization can be converted as well; however, turning movement volumes will be used for this example. To convert the volumes, the software utilizes a database conversion algorithm that only requires the user to enter the turning volumes for one interchange form. This initial form can be either an existing at-grade intersection which may be upgraded to an interchange or an existing interchange which needs modifications to cope with growing traffic congestion.

In the next step, the converted turning movement volumes can then be input into various operational software packages such as PASSER II, HCS, and TRANSYT-7F or analyzed using a standalone program. The advantage to the latter option is the ease of use to the user and efficiency of operation. A single software package would also have the capability of creating a feedback loop to re-adjust certain parameters and re-run the analysis. Future versions of INTERCHANGE can have the capability of performing either option.

A flow diagram of the database conversion architecture is shown in Figure F-3. As shown in the diagram, the core turning movement volumes are the central database of the program. Any volume data input into the peripheral interchanges automatically gets converted to the core module and then may be reconverted to any of the other interchange configurations once selected by the user. The benefit of this structure is that the user only needs to enter the turning movement volumes for one of the peripheral interchange configurations, and the effect of that input is available to all the other interchange forms. Any changes made to the inputs are immediately converted to the core database and ready to be converted to the other interchange types as they are selected by the user.

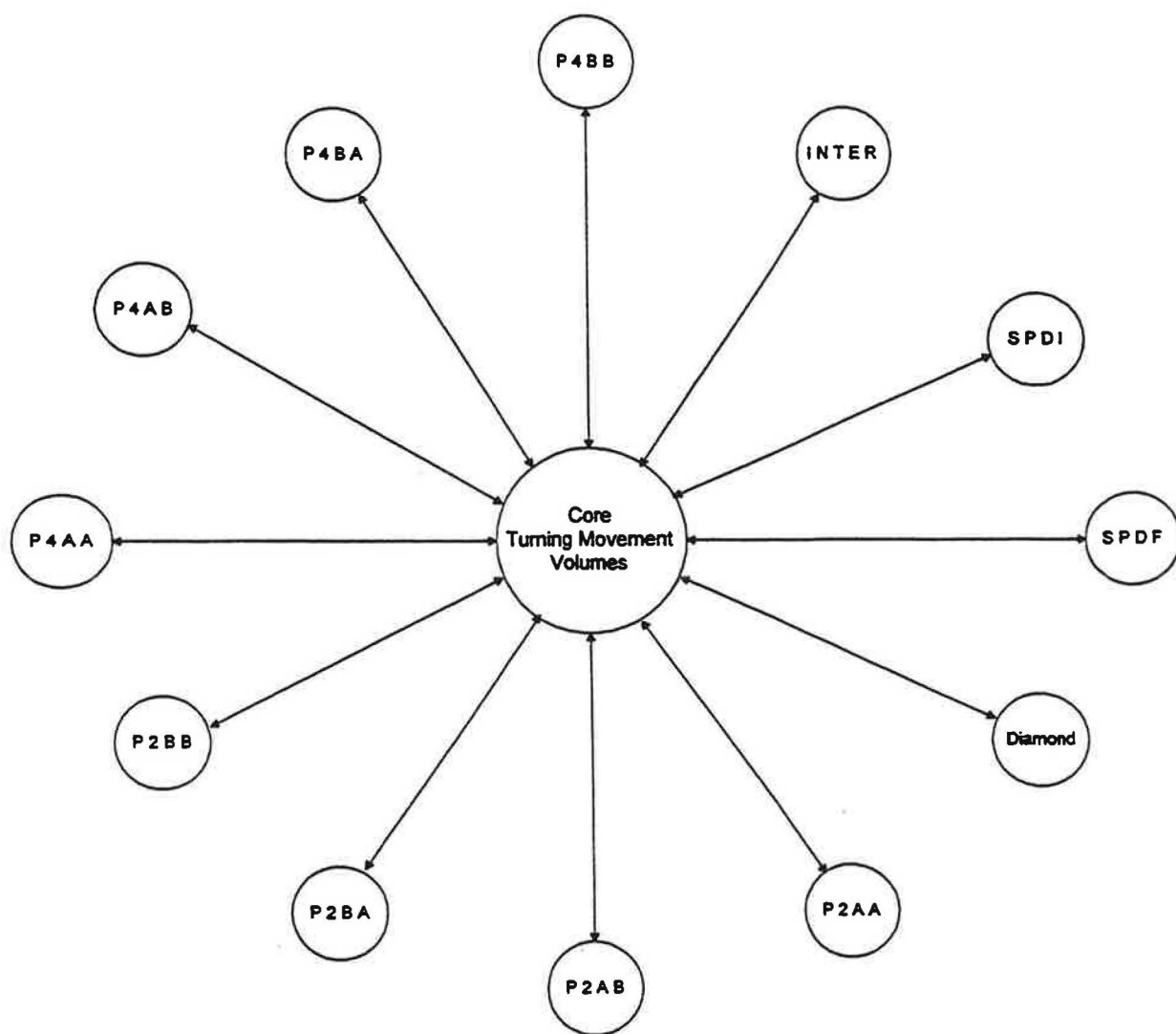


Figure F-3. Database architecture of INTERCHANGE software.

F.4 CURRENT STATUS OF DEVELOPMENT

The prototype for INTERCHANGE has been developed to convert turning movement volumes among 10 different interchange types including an at-grade intersection: namely, the diamond, single-point diamond, parclo 2-quad and 4-quad configurations. The conversion database architecture described earlier was used to perform the conversions. Input screens similar to Figure F-5 were developed using the Visual Basic programming language which is adapted for use with Windows environment. Particular attention was paid to developing a graphical user interface that would be understandable and easy to use. Two different input screens were developed as described in the following section to adapt to different user needs.

In addition to the turning movement conversion, the program has also been developed to provide a simple lane analysis for demonstration purposes. Based on the turning movement volumes, the program calculates the required number of lanes for each approach movement by using specific guidelines from Chapter 9, Appendix I of the 1994 Highway Capacity Manual (HCM). For instance, an exclusive left-lane is provided when left-turn volumes exceed 100 vph and a double left lane is provided for left-turn volumes above 300 vph. The user is allowed to interact with the analysis by entering in existing lane configurations and by overriding certain lane analysis calculations. This feature provides the user with a certain amount of control over the analytical procedure as well as flexibility in the use of the program.

The program structure was created to allow further enhancements to be made to the program. New modules or additional items can be easily added in the future. Possible near-term tasks include coding in parts of the volume adjustment, saturation flow, effective green, and capacity analysis modules for particular interchange types. The input screens can be programmed to accept additional data such as geometric and signalization conditions. Finally, the level of service module can be coded to display the performance measures and provide a means of displaying the interchange comparisons.

F.5 EXAMPLE PROBLEM

An example of the program is shown in the following graphics which show the conversion of turning movement volumes between an existing at-grade intersection and two proposed alternative designs: namely, a partial cloverleaf 2-quad AA, and a diamond configuration. At the end of this section, an example of the lane analysis procedure is shown. To demonstrate the program's interfacing capabilities, a hypothetical at-grade intersection with future turning movement volume was created and is shown in Figure F-4. The north and south through movements have very high volumes of 1100 vph and 1300 vph, respectively. Delays of 288 and 258 sec/veh, respectively, were output by PASSER II and are shown in Table F-1. Removing the through movements would improve the delay and level of service considerably. Therefore, a grade-separated interchange is envisioned to eliminate the heavy through movements. Due to hypothetical ROW, safety, and cost constraints, it was presumed that the two most favorable alternatives are the partial cloverleaf 2-quad AA and the conventional diamond configurations.

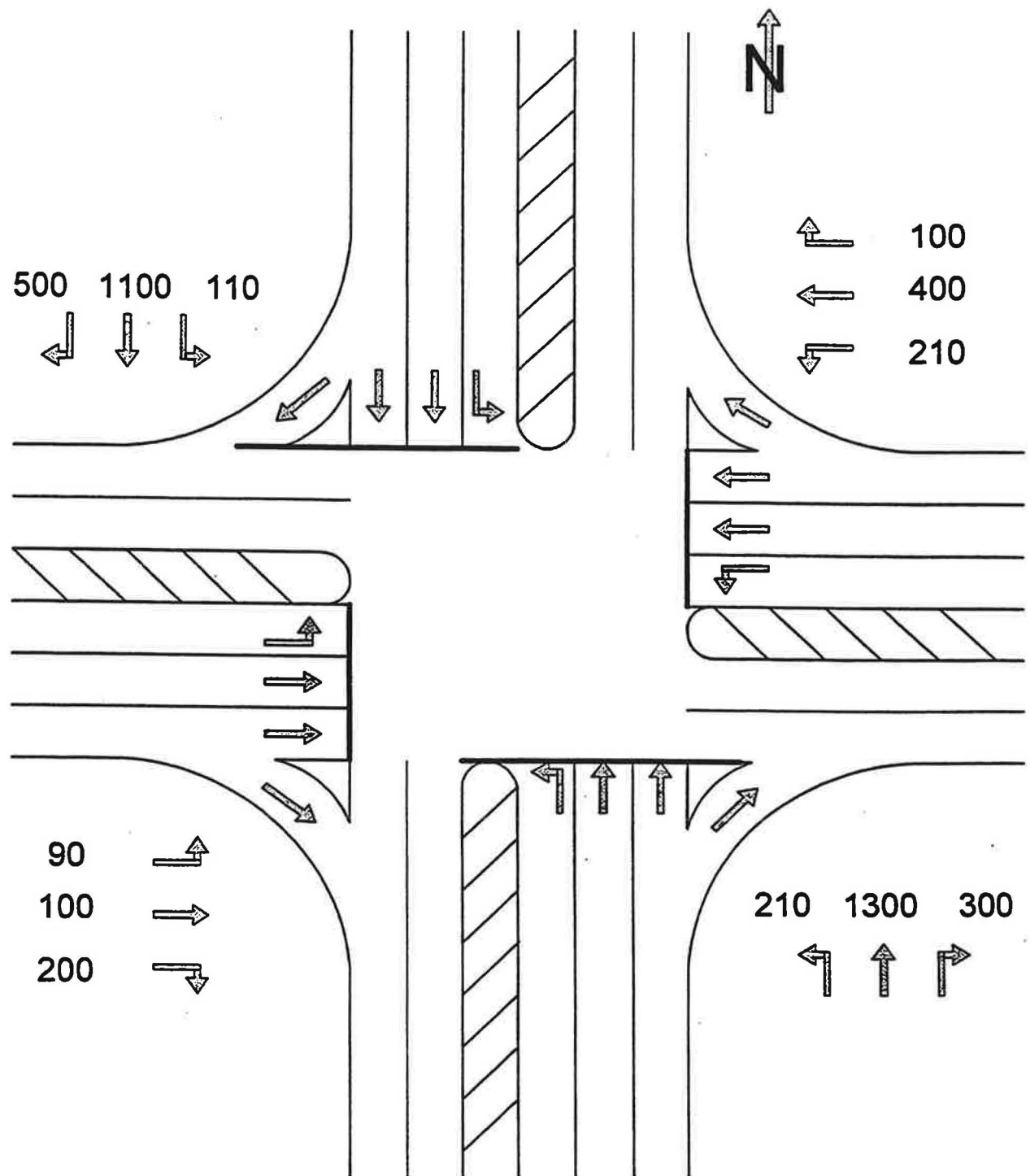


Figure F-4. Existing at-grade intersection with future turning movement volumes.

A step-by-step coding of the hypothetical example into INTERCHANGE will now be described. To begin, the user starts the program with the opening screen, shown in Figure F-5, entitled "Input Screen" as most of the data are input here. A small picture of the existing at-grade intersection is shown at the right of the screen; however, any interchange configuration can be chosen as the existing condition to begin entering turning movement volumes. To enter turning movement volumes, the user clicks on the appropriate square in the first column entitled "turning movement volumes." Future versions of the program will allow other information, shown as column headings, to be input. Clicking on the "turning movement volumes" column causes a small form to open, as shown in Figure F-6, which is used for entering turning movement volumes. As the values are entered, the directional arrows, which correspond to that movement, enlarge in the small picture of the interchange. This indication eliminates any confusion as to which movement is being coded. In addition, directions to the user are continuously being provided at each step through a small screen in the upper right-hand corner.

Table F-1. Performance of At-Grade Intersection with Future Turning Movement Volumes Using PASSER II

Performance Measures	Phase Movement Designation(Nema) (2) Note: Movement 6 & 2 correspond to N-S movements.							
	5	6	1	2	3	4	7	8
Volumes	210	1600	110	1600	210	300	90	500
V/C ratio	0.58	1.33	0.31	1.3	0.45	0.5	0.22	0.78
Delay (sec/veh)	13.8	288.3	11.2	258.5	14	22.8	13	28.8
Level of Service	B	F	B	F	B	C	B	C
Volumes *	210	500	110	300	210	300	90	500
V/C-RATIO *	0.51	0.88	0.19	0.53	0.45	0.5	0.22	0.78
DELAY (SECS/VEH)*	12	32.5	7.6	16.6	14	22.8	13	28.8
LEVEL OF SERVICE*	B	D	B	B	B	C	B	C

Note: * indicates values without N-S through movements.

INTERCHANGE

Input Screen

Turning Mov't of Vol Lanes Sat. Flow vphgpl Green Time sec Red Time sec

Choose One

- ☐ Diamond Interchange
- ☐ Parclo 2-Quad AA
- ☐ Parclo 2-Quad AB
- ☐ Parclo 2-Quad BA
- ☐ Parclo 2-Quad BB
- ☐ Parclo 4-Quad AA
- ☐ Parclo 4-Quad AB
- ☐ Parclo 4-Quad BA
- ☐ Parclo 4-Quad BB
- ☐ Single-Point Diam.
- ☒ Core Intersection

Directions
Click on the appropriate column to input values

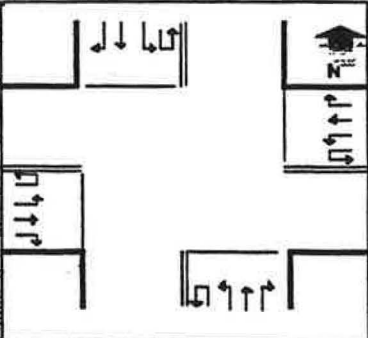


Figure F-5. Input screen with small picture of an at-grade intersection.

INTERCHANGE

Input Screen

Turning Mov't of Vol Lanes Sat. Flow vphgpl Green Time sec Red Time sec

Choose One

- ☐ Diamond Interchange
- ☐ Parclo 2-Quad AA
- ☐ Parclo 2-Quad AB
- ☐ Parclo 2-Quad BA
- ☐ Parclo 2-Quad BB
- ☐ Parclo 4-Quad AA
- ☐ Parclo 4-Quad AB
- ☐ Parclo 4-Quad BA
- ☐ Parclo 4-Quad BB
- ☐ Single-Point Diam.
- ☒ Core Intersection

Core Intersection

Turning Movement Volumes

SB	500	WB	100
ST	1100	WT	400
SL	100	WL	200
SU	10	WD	10
EB	200	NR	300
ET	100	NT	1300
EL	80	NL	200
EU	10	NU	10
EBR		NR	

Directions
Click on the appropriate column to input values

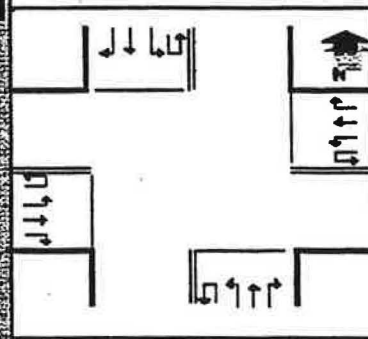


Figure F-6. Small picture of at-grade intersection with input form.

The proposed freeway will run north-south and eliminate the heavy north-south through turning movements volumes: namely, the 1100 vph south and 1300 vph north volumes. When coding the future at-grade intersection turning movement volumes into INTERCHANGE, the north-south through movement volumes are entered in normally as is shown in Figure F-6. An enlarged picture of the intersection can be viewed by clicking the "Max" button. Figure F-7 shows the screen of the enlarged at-grade intersection. The same volumes that were coded in the input boxes in Figure F-6 are now shown in their respective approaches directly on the picture of the intersection. This enlarged view may be more appropriate for some users since turning movement volumes can be changed within the picture and little confusion exists as to the location of the turning movement volumes.

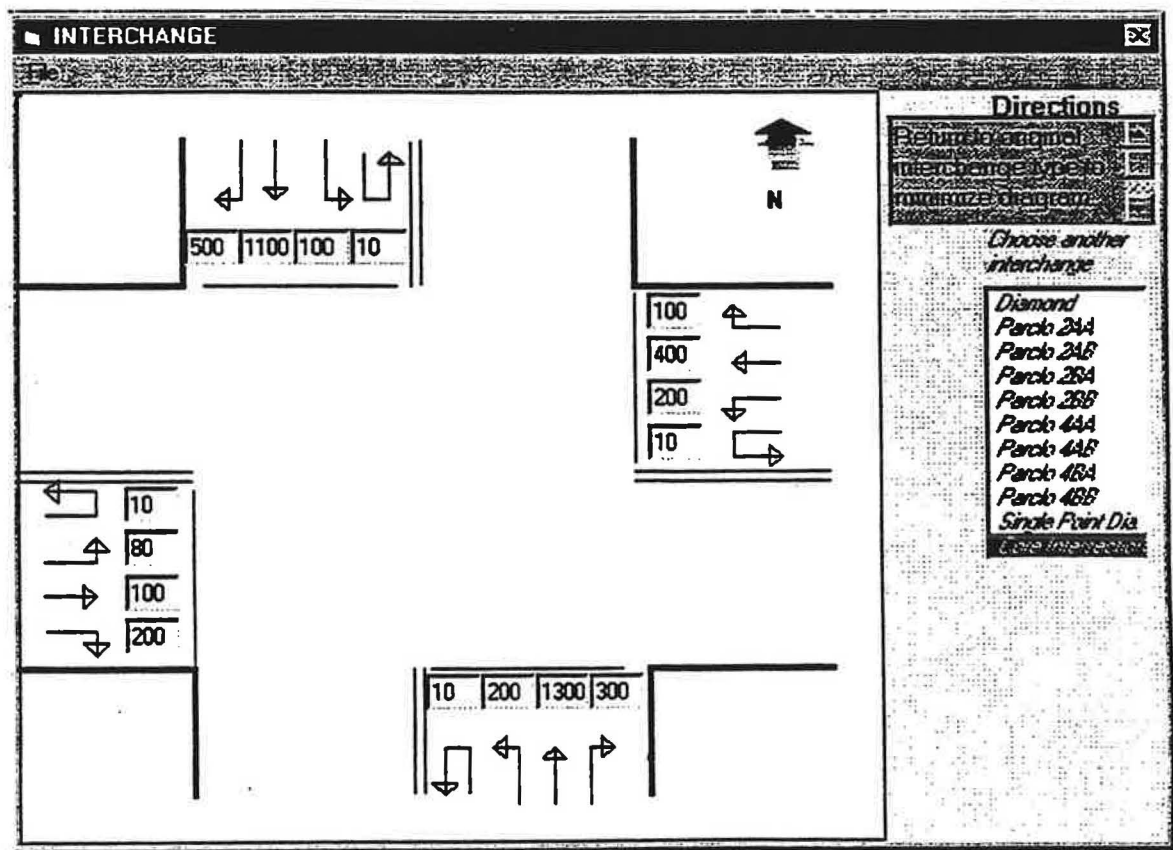


Figure F-7. Enlarged picture of at-grade intersection.

To compare the existing at-grade intersection with the partial cloverleaf 2-quad AA form the user can either return to the original screen with the smaller pictures or choose an enlarged picture of the 2-quad AA form from the drop down menu box at the top right of the screen. Figure F-8 shows the results of choosing the latter. The enlarged partial cloverleaf 2-quad AA picture

the converted turning volumes already automatically provided on the interchange approaches. The north and south through and U-turn movements have been eliminated since it was assumed that these movements are not served properly by this interchange form and therefore their use is discouraged.

The user can readily compare another interchange form to the original intersection or minimize the enlarged interchange drawing by clicking on the "Min" button at the lower right. Figure F-9 demonstrates the results of choosing to compare the existing at-grade intersection and partial cloverleaf 2-quad AA alternative with a second interchange alternative, in this case a diamond interchange. As with the cloverleaf, all the movements are automatically provided on the diamond interchange approaches. Note how the east and west right and left turning movements have exchanged volumes during the conversion between the two interchange configurations. The north-south through movements present in the at-grade intersection, namely 1300 and 1100 vph, have been replaced by default frontage road through volumes (0 vph in this case) in the diamond interchange configuration. The through movements for both of these forms will remain mutually exclusive throughout the program.

The results of a comparison of measures of performances generated by PASSER II are shown in Tables F-2 through F-4. Table F-2 shows a comparison of system delay and average intersection delay. The diamond interchange performed slightly better than the parclo 2-quad AA with system delays of 15.65 sec/veh as opposed to 18.9, respectively. Delays associated with both the left and right side intersection of each interchange are shown in Table F-3. On average, the left-side intersection of both interchanges provided better performance for the future turning volumes than the right-side. This can be attributed to the higher turning movement volumes that used the left side; namely, the high west-bound left turning volumes, 200 vph, and the high southbound right movements, 500 vph. Significant fluctuations in performance measures were seen in the parclo 2-quad AA design. Very high delays of 36.3 and 30.3 sec/veh were seen on the left-side intersection's east-bound left and west-bound through/right movements respectively, and very low delays of 7.9 sec/veh were experienced on the east-bound through. The parclo 2-quad AA also experienced similar fluctuations in volume-to-capacity ratios for each turning movement as shown in Table F-4.

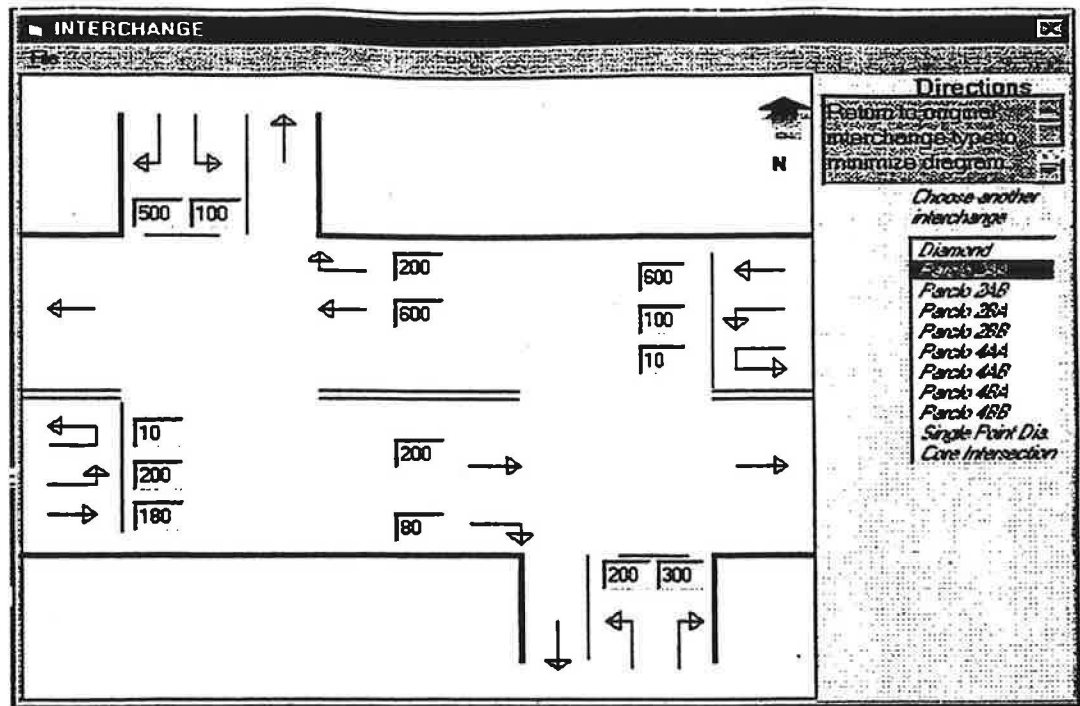


Figure F-8. Enlarged picture of a partial cloverleaf 2-quad AA.

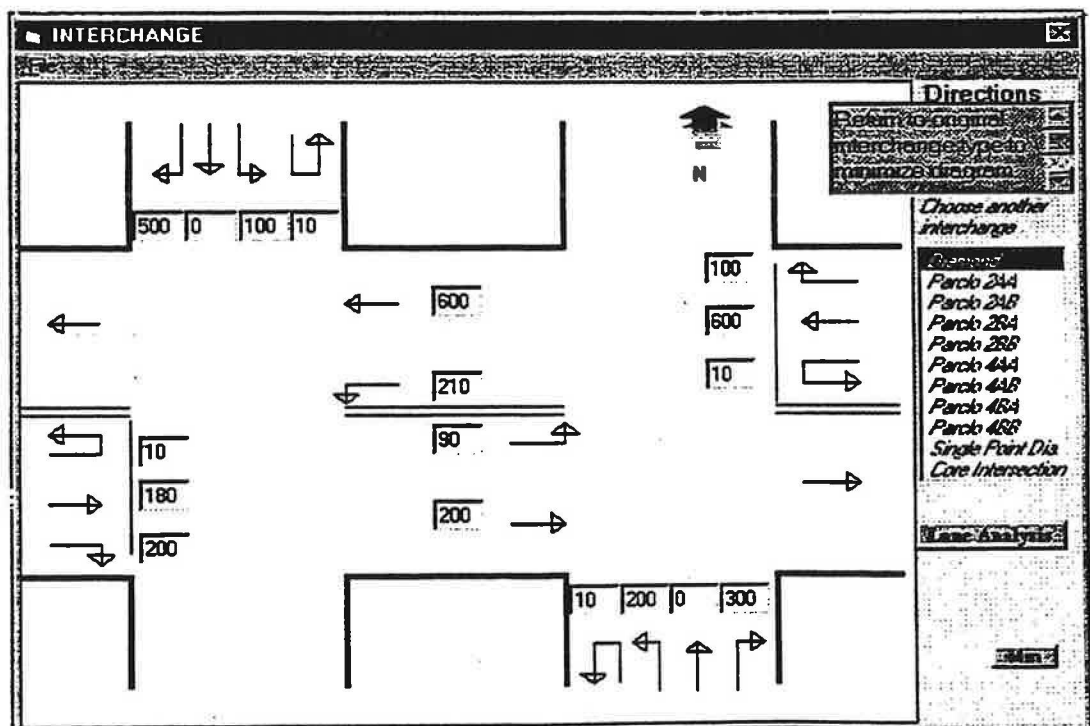


Figure F-9. Enlarged picture of a diamond interchange.

Table F-2. Comparison of System Delay and Average Intersection Delay

Interchange Type	System Delay (sec/veh)	Avg. Intersection Delay(sec/veh)
Diamond	15.65	16.7
Parclo 2 AA	18.9	20.7

Table F-3. Delay Analysis for Left and Right-Side Intersections Using PASSER II

Interchange Type	PASSER II Movement Designation							
	5	6	1	2	3	4	7	8
Left-Side								
Diamond	*25.1	11.5	24.2	23.3	0	20.2	11.3	0
Parclo 2 AA	36.3	30.3	0	7.9	0	27.3	12.8	0
Right-Side								
Diamond	7.6	14.2	*24.8	13.3	16.5	0	0	19.6
Parclo 2 AA	0	8.1	20.8	16.2	14.8	0	0	17.2

Note: * indicates u-turn traffic with no-bay

Table F-4. V/C Analysis for Left and Right-Side Intersections Using PASSER II

Interchange Type	PASSER II Movement Designation							
	5	6	1	2	3	4	7	8
Left-Side								
Diamond	*0.06	0.56	0.59	0.59	0	0.76	0.15	0
Parclo 2 AA	0.78	0.81	0	0.1	0	0.84	0.15	0
Right-Side								
Diamond	0.31	0.51	*0.06	0.14	0.38	0	0	0.6
Parclo 2 AA	0	0.31	0.31	0.31	0.33	0	0	0.55

Note: * indicates u-turn traffic with no-bay

The previous analysis assumed that the lane configuration was the same as that shown in Figure F-4. To evaluate various lane configurations, the user can employ the lane analysis features of INTERCHANGE which was designed to allow the user a certain amount of interaction in the analysis process. To begin the analysis, the user clicks the "lane analysis" button shown at right in Figure F-9. In the lane analysis screen, shown in Figure F-10, the user can have the program immediately perform a lane analysis based on the turning movement volumes by clicking the "calculate" button or the user can evaluate an existing lane configuration. This latter choice allows the existing condition to be compared with the optimal lane configuration provided by INTERCHANGE. Figure F-11 shows the results of the program's lane analysis, whereby the lane analysis results are shown in their respective boxes.

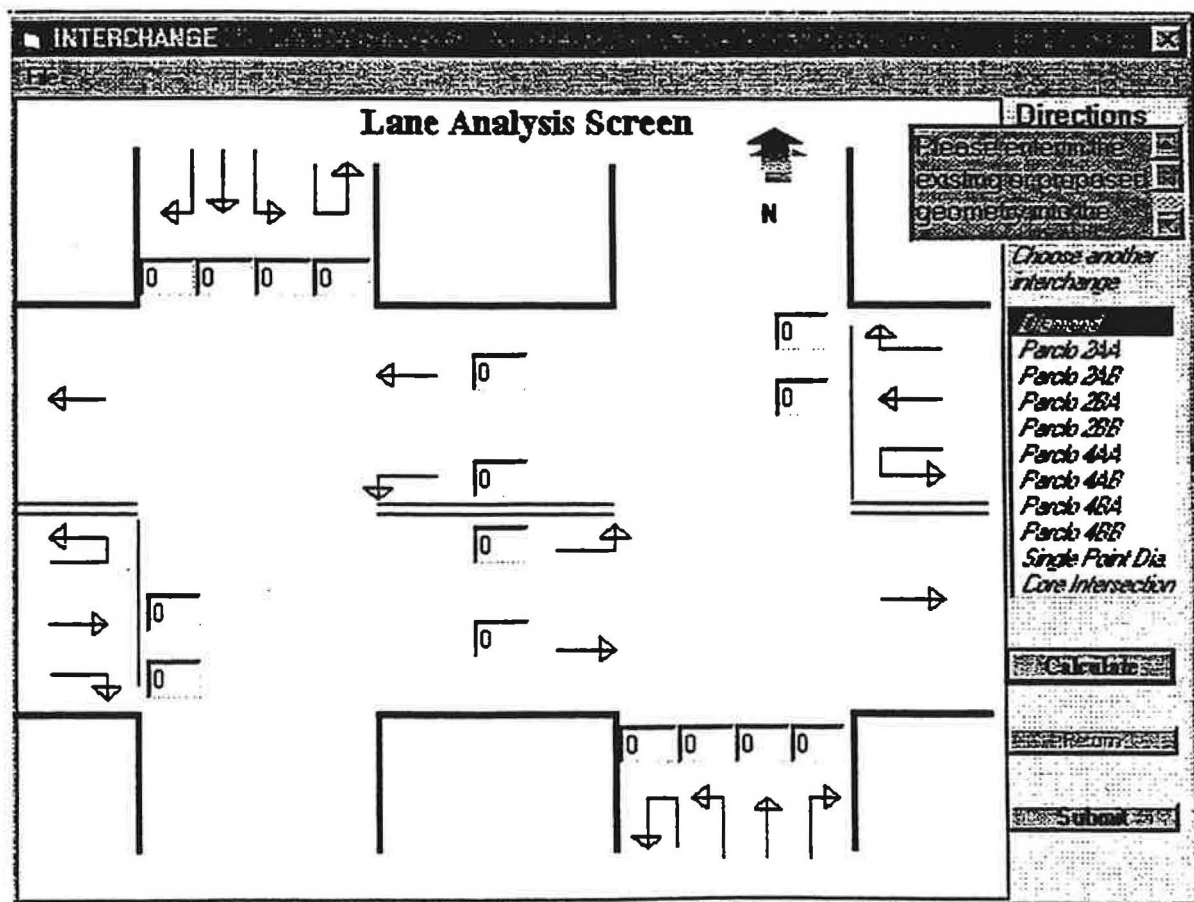


Figure F-10. Enlarged picture of diamond interchange with lane analysis screen.

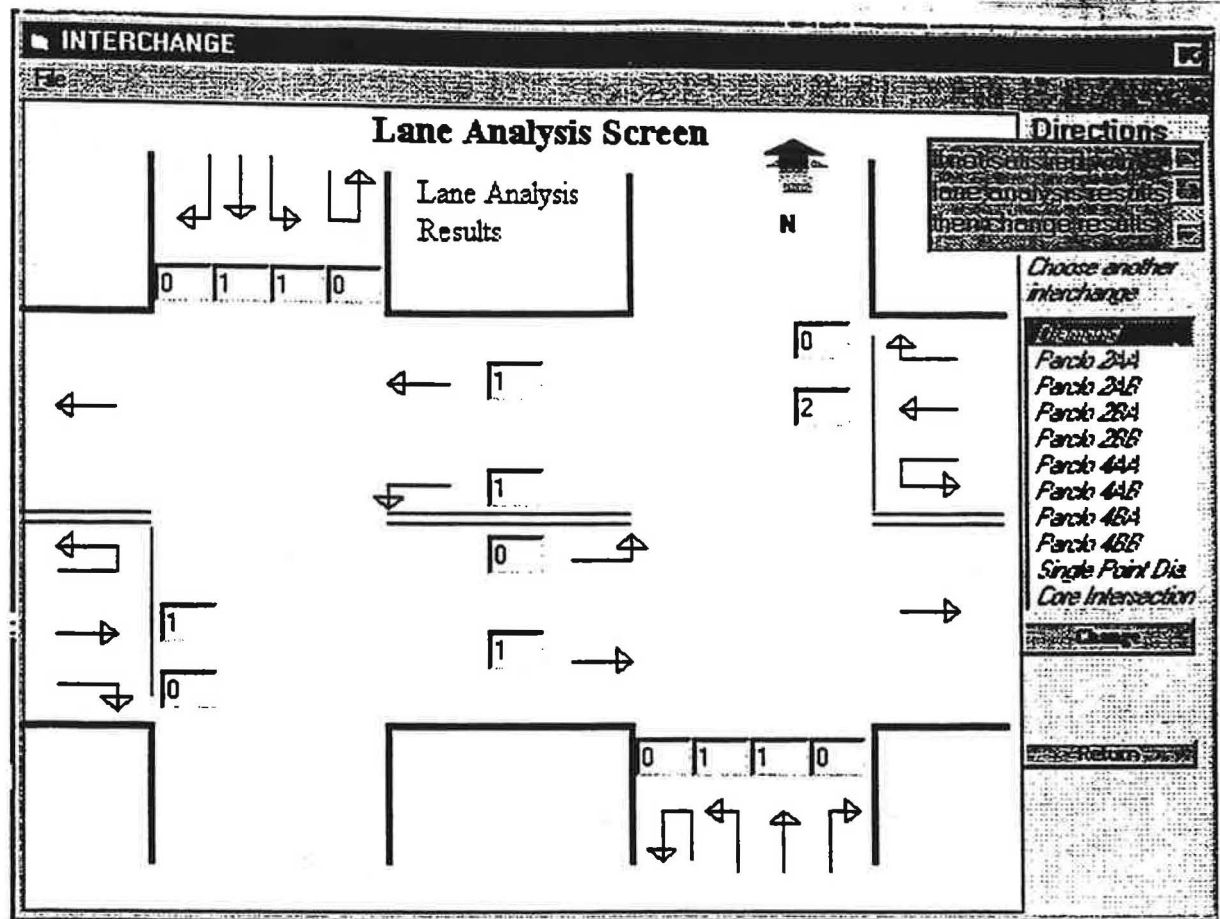


Figure F-11. Enlarged picture of a diamond interchange with lane analysis results.

The user still has the option of re-entering a desired lane configuration after the lane analysis results have been shown and running the analysis one more time. If the user changes certain minimum requirements, for instance, removing a left turn bay, the program will then prompt the user, as shown in Figure F-12, that certain minimum requirements, such as minimum number of exclusive left-lanes, has been violated and whether to continue to override the minimum values. Dependent on the response of the user, the program will either keep the minimum number of lanes or use the user inputted value. Any additional lanes required are added to the through lanes as shown in the final lane analysis results in Figure F-13. For instance, an additional lane has been added to Westbound through movement at the left-side intersection as a result of the user choosing the "no lane" option as demonstrated in Figure F-12.

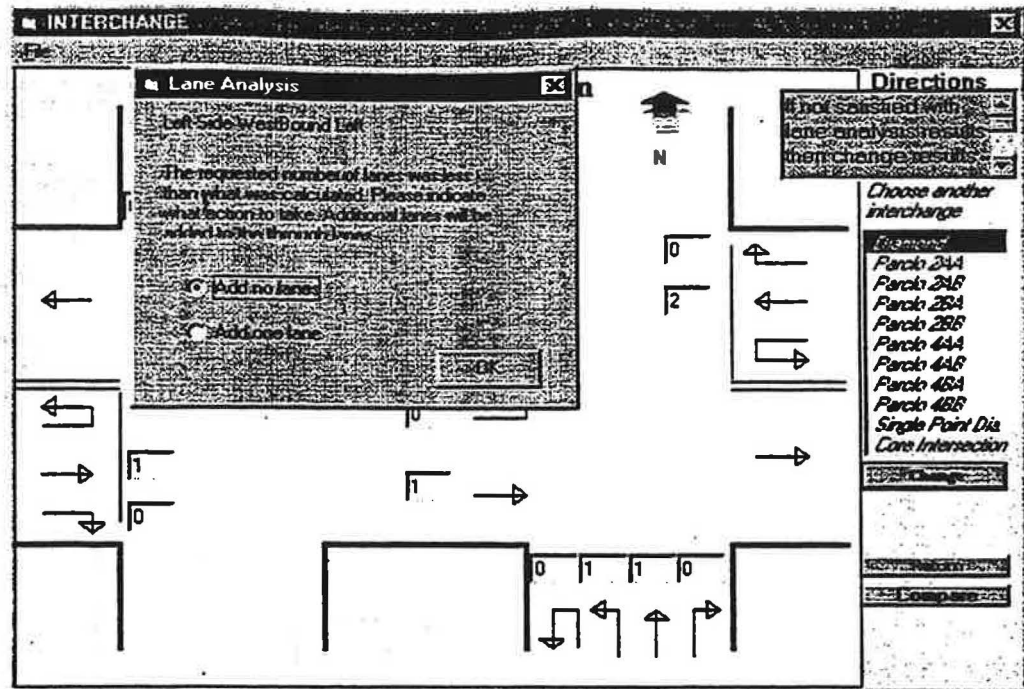


Figure F-12. Enlarged picture of a diamond interchange with lane analysis user prompt.

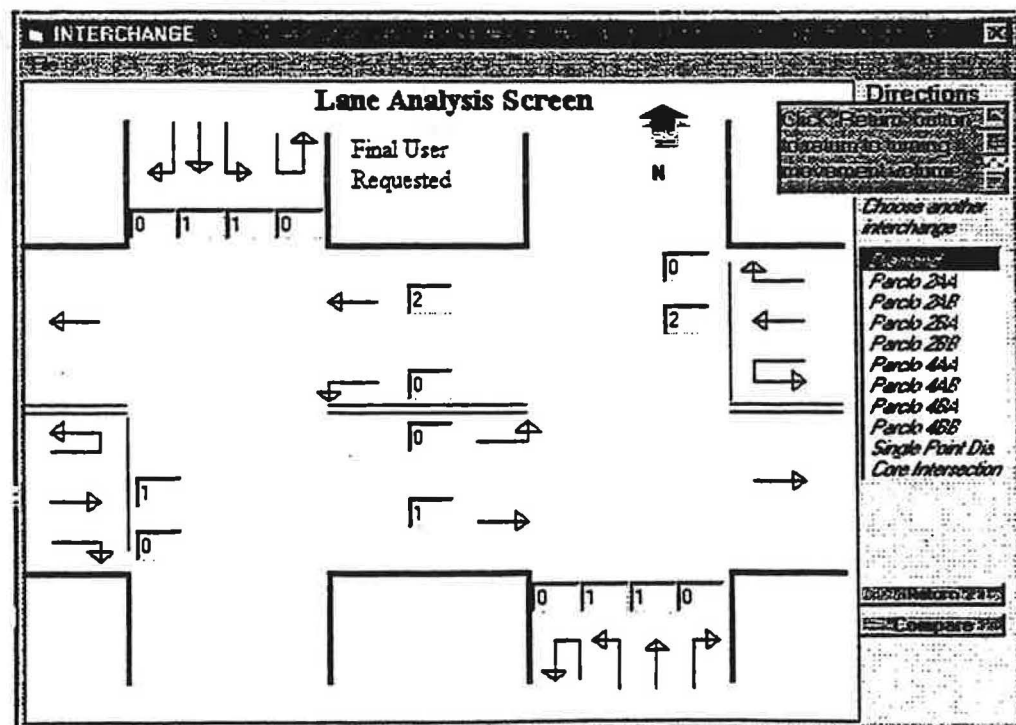


Figure F-13. Enlarged picture of a diamond interchange with final lane analysis.

As demonstrated in this hypothetical example, INTERCHANGE can be a very useful tool for interchange analysis and comparison. Having a standalone program, as suggested in Figure F-2, can greatly improve the efficiency of the analysis as shown by using PASSER II results. Performance measures can be calculated within the program and be output for individual interchange analysis or for comparison purposes. For instance, once each interchange alternative has been analyzed, a central database of performance measures can be accessed at any time to produce comparison reports, thereby simplifying the selection process.

REFERENCES

1. Highway Capacity Manual." Special Report 209, Third Edition, Transportation Research Board, Washington, D.C. (1994).
2. Messer, C.J., and C.P. Chang. "Arterial Signal Timing Optimization Using PASSE II-90." Report 467-2F. Texas Transportation Institute, College Station, TX. (1991).
3. Van Arendonk, J. "Development of an Interchange Analysis Software." M.S. Thesis. Texas A&M University, in press, College Station, May 1997.