194

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM
REPORT 194

TRAFFIC CONTROL IN OVERSATURATED STREET NETWORKS

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM REPORT 194

TRAFFIC CONTROL IN OVERSATURATED STREET NETWORKS

LOUIS J. PIGNATARO,
WILLIAM R. McSHANE, KENNETH W. CROWLEY,
BUMJUNG LEE, AND THOMAS W. CASEY
POLYTECHNIC INSTITUTE OF NEW YORK
BROOKLYN, NEW YORK

RESEARCH SPONSORED BY THE AMERICAN ASSOCIATION OF STATE HIGHWAY AND TRANSPORTATION OFFICIALS IN COOPERATION WITH THE FEDERAL HIGHWAY ADMINISTRATION

AREAS OF INTEREST:
TRAFFIC CONTROL AND OPERATIONS
TRAFFIC MEASUREMENTS

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
WASHINGTON, D.C. 1978

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

Systematic, well-designed research provides the most effective approach to the solution of many problems facing highway administrators and engineers. Often, highway problems are of local interest and can best be studied by highway departments individually or in cooperation with their state universities and others. However, the accelerating growth of highway transportation develops increasingly complex problems of wide interest to highway authorities. These problems are best studied through a coordinated program of cooperative research.

In recognition of these needs, the highway administrators of the American Association of State Highway and Transportation Officials initiated in 1962 an objective national highway research program employing modern scientific techniques. This program is supported on a continuing basis by funds from participating member states of the Association and it receives the full cooperation and support of the Federal Highway Administration, United States Department of Transportation.

The Transportation Research Board of the National Research Council was requested by the Association to administer the research program because of the Board's recognized objectivity and understanding of modern research practices. The Board is uniquely suited for this purpose as: it maintains an extensive committee structure from which authorities on any highway transportation subject may be drawn; it possesses avenues of communications and cooperation with federal, state, and local governmental agencies, universities, and industry; its relationship to its parent organization, the National Academy of Sciences, a private, nonprofit institution, is an insurance of objectivity; it maintains a full-time research correlation staff of specialists in highway transportation matters to bring the findings of research directly to those who are in a position to use them.

The program is developed on the basis of research needs identified by chief administrators of the highway and transportation departments and by committees of AASHTO. Each year, specific areas of research needs to be included in the program are proposed to the Academy and the Board by the American Association of State Highway and Transportation Officials. Research projects to fulfill these needs are defined by the Board, and qualified research agencies are selected from those that have submitted proposals. Administration and surveillance of research contracts are responsibilities of the Academy and its Transportation Research Board.

The needs for highway research are many, and the National Cooperative Highway Research Program can make significant contributions to the solution of highway transportation problems of mutual concern to many responsible groups. The program, however, is intended to complement rather than to substitute for or duplicate other highway research programs.

NCHRP Report 194

Project 3-18(2) FY '71
ISSN 0077-5614
ISBN 0-309-02854-X
L. C. Catalog Card No. 78-65950

Price: $9.60

Notice

The project that is the subject of this report was a part of the National Cooperative Highway Research Program conducted by the Transportation Research Board with the approval of the Governing Board of the National Research Council, acting in behalf of the National Academy of Sciences. Such approval reflects the Governing Board's judgment that the program concerned is of national importance and appropriate with respect to both the purposes and resources of the National Research Council.
The members of the technical committee selected to monitor this project and to review this report were chosen for recognized scholarly competence and with due consideration for the balance of disciplines appropriate to the project. The opinions and conclusions expressed or implied are those of the research agency that performed the research, and, while they have been accepted as appropriate by the technical committee, they are not necessarily those of the Transportation Research Board, the National Research Council, the National Academy of Sciences, or the program sponsors.
Each report is reviewed and processed according to procedures established and monitored by the Report Review Committee of the National Academy of Sciences. Distribution of the report is approved by the President of the Academy upon satisfactory completion of the review process.
The National Research Council is the principal operating agency of the National Academy of Sciences and the National Academy of Engineering, serving government and other organizations. The Transportation Research Board evolved from the 54-year-old Highway Research Board. The TRB incorporates all former HRB activities but also performs additional functions under a broader scope involving all modes of transportation and the interactions of transportation with society.

Published reports of the

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

are available from:

Transportation Research Board
National Academy of Sciences
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

Printed in the United States of America.

FOREWORD

By Staff
Transportation
Research Board

This report will interest not only traffic engineers but also others concerned with traffic engineering solutions to problems of urban traffic congestion. Researchers in the field of traffic characteristics and traffic operations will find several appendices to be of interest. Furthermore, in addition to the presentation of research findings, the report contains "Guidelines for the Treatment of Traffic Congestion on Street Networks." These guidelines will be of primary interest to traffic operations personnel and, apart from the remainder of the report, can be used as both a tutorial aid and a manual when faced with specific local traffic congestion problems.

The research project from which this report resulted began in 1971 as one of two projects investigating improved traffic control measures. Funds for both projects were provided as an outgrowth of the recommendations from *NCHRP Report 84,* "Analysis and Projection of Research on Traffic Surveillance, Communication, and Control." The first project, "Improved Control Logic for Use with Computer-Controlled Traffic" (NCHRP Project 3-18(1)), was conducted by Stanford Research Institute and led to development of the ASCOT package and traffic control system. The second project, dealing with other solutions to congested urban traffic network problems, was undertaken by the Polytechnic Institute of New York (then Polytechnic Institute of Brooklyn) and is concluded with this report.

The objectives of the project were to examine the seriousness of the traffic congestion problem nationwide, to evaluate the effectiveness of available remedial techniques, and subsequently to prepare guidelines for practicing traffic engineers who can use them to effectively address the problems of congested networks. Among the recommendations for future work that were made, the one which dealt with the need for developing a cost-effectiveness methodology has already been implemented. NCHRP Project 3-18(3), "Cost-Effectiveness Methodology for Evaluation of Signalized Street Network Surveillance and Control Systems," was initiated in May 1975.

CONTENTS

ACKNOWLEDGMENTS

The research reported herein was performed under NCHRP Project 3-18(2) by the Polytechnic Institute of New York, with Louis J. Pignataro, Head of Department and Professor of Transportation Engineering, as principal investigator.

He was assisted by William R. McShane, Associate Professor; Kenneth W. Crowley, Associate Professor; Bumjung Lee, Assistant Professor; and Thomas W. Casey, Research Associate.

Grateful acknowledgment is extended for the aid and support of other Polytechnic staff members, especially: J. J. Starace, for data collection and equipment maintenance in Phase I; Miss B. A. Cicala; Mrs. E. Cummings; and Mrs. C. Devlin.

TRAFFIC CONTROL IN OVERSATURATED STREET NETWORKS

SUMMARY

As stated in the research Project Statement:

> Traffic operations and control techniques that function effectively when street network demands are below saturation deteriorate when severe saturation exists for any length of time. Research is needed to define the scope and magnitude of the problem nationwide; to determine how the problem can best be combatted with existing control techniques; and to begin a systematic research process leading to improved operation and control of oversaturated networks.

This report discusses the scope, magnitude, and root causes of the congestion/saturation problem as determined by a questionnaire survey of traffic engineering professionals and by an extensive set of personal and telephone interviews. In general, the root causes of congestion—as perceived by these practitioners—are founded in lack of alternate routes, in land use policies that generate the traffic patterns, and in vehicle-pedestrian conflicts that aggravate the situation. The engineer is faced with too much traffic, in too little space, in too short a time. Therefore, it must be recognized that the problem should be first attacked at the root causes.

Measures of saturation were identified from a complete review of the relevant literature, a thorough analysis of candidate measures, and a program of field work via time-lapse photography. The results of this work were then used to define a set of appropriate measures and to develop definitions of the various levels of saturation.

A range of treatments was studied via simulation—the UTCS-1 model simulator was used extensively—analytical methods, and some supportive fieldwork. The various candidate treatments and remedies are discussed in terms of three major categories:

1. Signal: Minimal-response signal policies.
2. Signal: Highly responsive signal policies.
3. Nonsignal: Other treatments in a signalized environment.

Two working papers were also developed as part of this effort. One presents a highly responsive queue-activated control policy. The other analyzes the regularity of the vehicular traffic patterns.

This report also contains a set of guidelines developed for the treatment of traffic congestion on street networks. The guidelines provide both a tutorial and an illustrated reference in what techniques to consider and how to consider them systematically.

In general, it was found that the problem of congestion and saturation is widespread, and a broad variety of techniques is being applied in an effort to combat the problem. However, there is no specific direction of attack. On the basis of the project results, the researchers conclude that there are definite measures that can be taken, but preventive action addressing the root causes must be given a high priority. Among the measures that can be taken, those relating to signalization generally can have the greatest impact. There are distinct signal plans for avoiding spillback and for "living with" spillback. The nonsignal remedies

are in no way to be overlooked, particularly those that provide space either for direct productivity increases or for removing impedances to the principal flow.

Recommendations for future work include, among others, emphasis on the cost-effectiveness analysis of situations in which there is competition for street space.

CHAPTER ONE

INTRODUCTION AND RESEARCH APPROACH

PROBLEM STATEMENT AND RESEARCH OBJECTIVES

In early 1971, the National Cooperative Highway Research Program (NCHRP) issued a project statement for Project 3-18(2), which specified the following:

> Severe traffic congestion occurs on the street networks in central districts and other high-activity centers of many urban areas. When streets become "jammed," or oversaturated, capacity and operational efficiency are degraded, resulting in suboptimum utilization of existing facilities. . . . Traffic operations and control techniques that function effectively when street network demands are below saturation deteriorate when severe saturation exists for any length of time. Research is needed to define the scope and magnitude of the problem nationwide; to determine how the problem can best be combatted with existing control techniques; and to begin a systematic research process leading to improved operation and control of oversaturated networks.

Accordingly, this project was initiated with the following over-all objectives:

1. Define the scope and magnitude of the congestion/oversaturation problem.
2. Establish the root causes of such congestion and/or oversaturation.
3. Evaluate and/or define measures of saturation.
4. Evaluate the relative effectiveness of various traffic control measures used to combat the problem.
6. Prepare in a manual form guidelines by which the practicing traffic engineer can effectively address the problem.
5. Undertake and report some limited work on highly responsive signal control.

The research was undertaken in two phases, with the objectives refined between the phases. The emphasis in the latter phase was on minimal-response signal policies and on nonsignal remedies (in a signalized environment).

RESEARCH APPROACH

To define the scope, magnitude, and root causes of the congestion/oversaturation problem, the researchers conducted a questionnaire survey of traffic engineering professionals and an extensive set of personal and telephone interviews. One member of the research team visited a number of city agencies and state offices in California to ascertain the specific techniques used by those organizations to combat congestion. Simultaneously, other members of the research team were conducting in-person interviews with appropriate traffic practitioners in the northeastern part of the United States.

On the basis of these first contacts, it appeared that little specific attention was being paid to the development of signal timing plans aimed primarily at the movement of vehicles in a congested environment. In general, in urban areas, the over-all control philosophy appeared to be to minimize the number of phases, to keep the signal cycles short, and to time signals for progressive movement—recognizing that, in some cases, segments of the system would congest for periods of time. The general feeling was that such an approach provided the best over-all signal timing philosophy. In some instances, traffic engineers were using simultaneous signal timing as a methodology for keeping queues short and thus preventing spillback at intersections upstream of the critical intersection.

Some of the traffic engineers questioned indicated that they use full actuated or volume density type equipment in an effort to make operations at the critical intersection as efficient as possible. Oakland, Calif., for example, has developed its own presence detection signal controller, which terminates a phase as soon as the long-loop detector (40 ft to 60 ft) indicates the absence of vehicles. This form of control was being applied at critical intersections to maximize their capacity.

Over-all, however, these first in-person interviews indicated that little effort was being directed specifically at combatting problems of either congestion or oversaturation. An analysis of the extensive mail survey responses pointed out that these findings were generally valid and that relatively little was being done specifically in the development of operational control plans aimed at combatting congestion or oversaturation. Furthermore, the 40 different

TOPICS reports received from respondents tended to further bear out this fact.

Examination of the TOPICS reports revealed that heavy emphasis was being placed on the upgrading of signal systems and the addition of capacity through intersection approach widening. A limited number of respondents indicated application of operational control procedures that did appear to be directly aimed at combatting congestion. Each of these agencies was contacted and appropriate professionals were interviewed. It is evident from their comments that a broad variety of techniques is being applied in an effort to combat congestion—no specific direction of attack appears, however.

Investigation of the measures of saturation was undertaken by a detailed review of the relevant literature, a thorough analysis of candidate treatments, and a program of field work via time-lapse photography. The results from these efforts formed the basis for defining a set of appropriate measures.

The various candidate treatments, or remedies, were found to fall into three major categories:

1. Signal: Minimal-response signal policies.
2. Signal: Highly responsive signal policies.
3. Nonsignal: Other treatments in a signalized environment.

In studying these treatments, simulation and analytical methods were used in conjunction with some supportive field work. The UTCS-1 model simulator, which was used extensively, had been calibrated and validated as part of the Urban Traffic Control System (UTCS) developed for the Federal Highway Administration. Additional validation was done using data obtained under the present effort.

Two working papers were also developed as part of this effort. One represents a highly responsive control policy, and the other analyzes the regularity and other characteristics of daily traffic volume patterns as observed by detectors.

On the basis of the research results stemming from these studies, it became evident that the approach to developing the guidelines by which the practicing traffic engineer can address the problem of relieving oversaturated conditions should center on the guidelines serving two functions:

1. As a tutorial for the practicing traffic engineer on the "basic concepts" or "basic mechanisms" of congestion/oversaturation.
2. As a reference enumerating candidate treatments and a systematic methodology for considering them.

This systematic methodology is a prime contribution, in that the research showed a lack of knowledge of some basic mechanisms (particularly the importance of block length and the timing of offsets when spillback is unavoidable) and a lack of due, systematic consideration of options. At the same time, the richness of the mix of local conditions and traffic patterns prevented reducing the systematic methodology still further to a set of "pat answers" in regard to a hierarchy of solutions with specified, unequivocal statements on their ranking.

In accord with the project objectives, a set of guidelines is presented in this report in Appendix J. It is a self-contained document for the use of those concerned with the problems of traffic movement—traffic congestion and traffic oversaturation. The guidelines are intended to aid the practicing engineer by reminding him of available options, by uncovering some subtleties that can be overlooked, and by presenting him with quantitative insight into the relative benefits of various options. In this way, an appreciation can be obtained of when various options are effective or necessitated, of how much impact can be expected, and what combinations are most effective.

First and foremost, it is important to understand that the guidelines are intended to aid the traffic engineer in making operational improvements to facilitate traffic movement in an existing situation. They do not address the benefits of land use management to optimize the demand for travel, nor do they address the rescheduling of demand via staggered work hours and other measures, nor do they (extensively) investigate reorganization of the traffic stream via increased transit use. More basically, they do not address the policy issue of whether vehicular traffic movement should be facilitated.

There is no intent to minimize those issues that are not addressed. Indeed, it is the land use pattern that shapes the traffic distribution and is the most basic and most important root cause of traffic congestion. It is strongly recommended that the traffic engineers and the transportation and urban planners stress this basic reality, and plan accordingly. Such actions, or changes, however, tend to have benefits in a different time frame than is of concern in the guidelines developed in this project.

Likewise, the problems of demand management and of transit options are usually the concern of others, not those who are involved with the operational concerns of the guidelines. These techniques have their own intrinsic merit, and whether or not significant improvements can be achieved within the framework of the guidelines, recourse to such other measures in transportation systems management will frequently be warranted. It is, however, beyond the scope of the guidelines to assess or even to address the effectiveness of such measures.

CHAPTER TWO

FINDINGS

SCOPE, MAGNITUDE, AND ROOT CAUSES OF CONGESTION/OVERSATURATION

The findings presented in this section concern the typical extent and causes of the traffic problem as determined from the results of a survey conducted of 569 locations, of which 445 were cities. The cities included all those of population in excess of 50,000 and all those with lower populations but with a traffic engineer. The details of the survey results, including the survey form, are contained in Appendix A.

The distribution of cities in the several population size categories used is shown in Figure 1. Figure 2 documents how many of these locales perceive congestion and saturation. (In the questionnaire, *congestion* is defined as a condition in which all waiting vehicles cannot pass through the intersection on one signal cycle. *Saturation* is defined in the survey as that condition when vehicles are prevented from moving freely because of the presence of vehicles either in the intersection itself or because of back-ups (jams) in any of the exit links of the intersection. This condition is defined in this report as *oversaturation*.)

Note from Figure 2 that most cities experience congestion or saturation (oversaturation), or both, for periods in excess of 30 min. Many cities reported that from 6 percent to 18 percent of their affected signals are involved for a period of less than 30 min. It would seem that if a treatment forestalls the onset of saturation for 10 to 15 min, it would have a substantial impact, expressed as a percentage—but only in a limited number of cases.

Figure 3 shows that the duration of congestion varies according to location within the city. Clearly, the central business districts (CBD) and the major arterials experience congestion for substantial periods. The smallest cities —those with less than 50,000 population—have in many respects the greatest problem with saturation (i.e., oversaturation). The largest cities may have the greatest percentage of congested intersections, but they have the smallest percentage of saturated intersections.

Inquiries were also made in the survey as to the extent of queue formation. Figure 4 summarizes these results, showing again that the problem of saturation is not a characteristic of the largest cities; rather, it is a characteristic of the small and medium cities. In this case, cities of population from 250,000 to 500,000 show the greatest numbers of longer queues, extending halfway or more up to the upstream intersection.

It is apparent from the foregoing survey results that some inconsistencies exist—for example, that cities in the 250,000 to 500,000 population range have the lowest percent of intersections experiencing congestion and at the same time have the worst queuing problems. A review of a few of the maps provided with the survey responses, for this particular grouping, shows that such cities have a compact CBD with

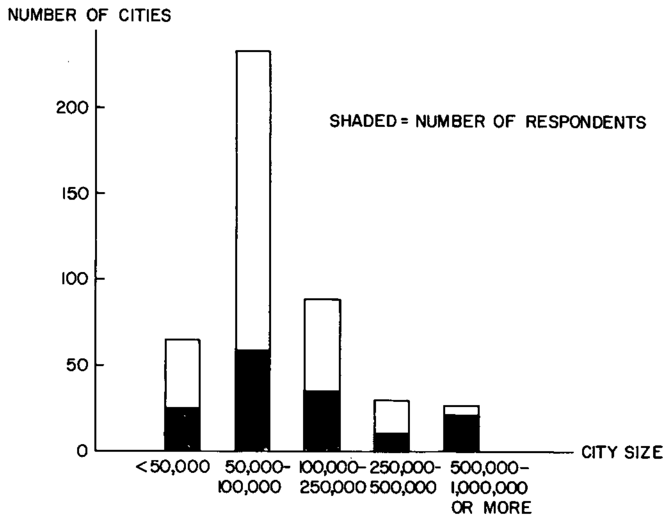


Figure 1. Distribution of city sizes queried.

a number of feeder arterials radiating out from it, experiencing saturation within the CBD and varying degrees of congestion on feeder radials.

This, of course, is true in differing degrees, of all city size groupings. However, at the smaller population levels, the CBD may be comprised of one or two or three intersections and a single major arterial, and at the other end of the population scale (over 500,000) there may be huge CBD's (or even multiple CBD's) and many major approaches. A pattern is reasonably consistent in that problems of saturation and congestion appear to peak in the 100,000 to 250,000 population grouping, and reduce somewhat in the next higher population level, indicating a better ability to cope with the problem (larger CBD, more arterial approaches). The largest cities (over 500,000) show the greatest congestion, but less saturation. The decrease in saturation might be attributed to the fact that the greater city size allows more diffusion of the congested traffic.

In discussing these problems with some representative local professionals, it was found that in some cases there may only be a single major street within 1 mi lateral distance (for instance, Wichita, Kans., Pop. 276,554). The long period of congestion (50 to 60 min) is apparently due to the lack of multiple facilities. The same is true in Kansas City, Mo., which has a population of just over 500,000. Saturation is somewhat more of a problem at this end of the 250,000 to 500,000 category, with Kansas City experiencing saturation up to 15 min in duration.

From this same survey, the researchers concluded that, in general, the root causes of congestion—as perceived by these traffic engineering practitioners—are founded in lack

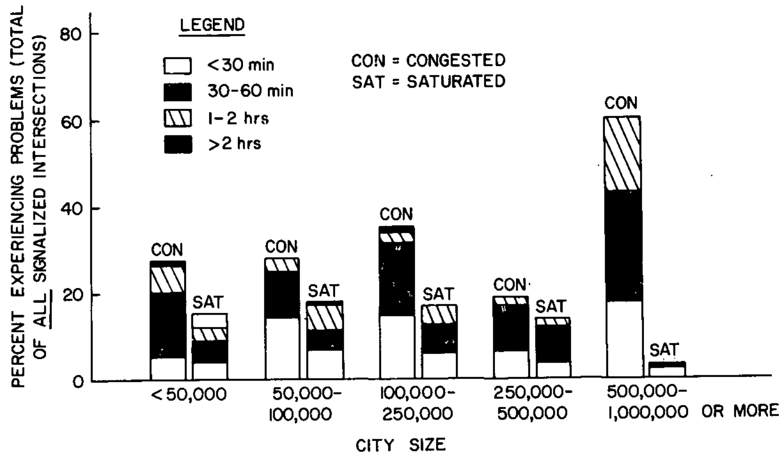


Figure 2. Extent and duration of traffic problems by city size.

of alternate routes, in land use policies that generate the traffic patterns, and in vehicle-pedestrian conflicts that aggravate the situation. The engineer is faced with too much traffic, in too little space, in too short a time.

Therefore, it must be recognized that the problem should be first attacked at the root causes; for example:

- The choice of alternate routes.
- The concentrations of movement implied in some land use distribution.
- The use of on-street facilities for goods loading/unloading.
- The multiplicity of access/egress points, and of standing queues on the rights-of-way.
- The concentration of work trips in a short period.
- Unrestricted hours of goods activity.
- Numbers of vehicles, particularly low-occupancy private automobiles.
- Inefficient curb space management: parking, moving lanes, etc.
- Lack of grade separation in areas of extreme intensity.
- The improper design, location, and integration of access/egress points.

The engineer should continually educate other specialists and the public in this need. However, in the time frame of a local, site-specific problem that must be addressed, the guidelines developed from this research and contained in Appendix J are recommended for use.

DEFINING MEASURES OF SATURATION

Terms and Characteristics

The analyses contributing to the findings discussed in this section are contained in Appendixes B and C, and were refined by the field studies described in Appendix D. From the results of this work, it was determined that queue-based measures were the most meaningful indicators of congestion. For this reason, the following definitions are written in terms of queue formation. Throughout, it will be recognized that the problem focuses on an intersection that has become critical. A *critical intersection* (CI) is an inter-

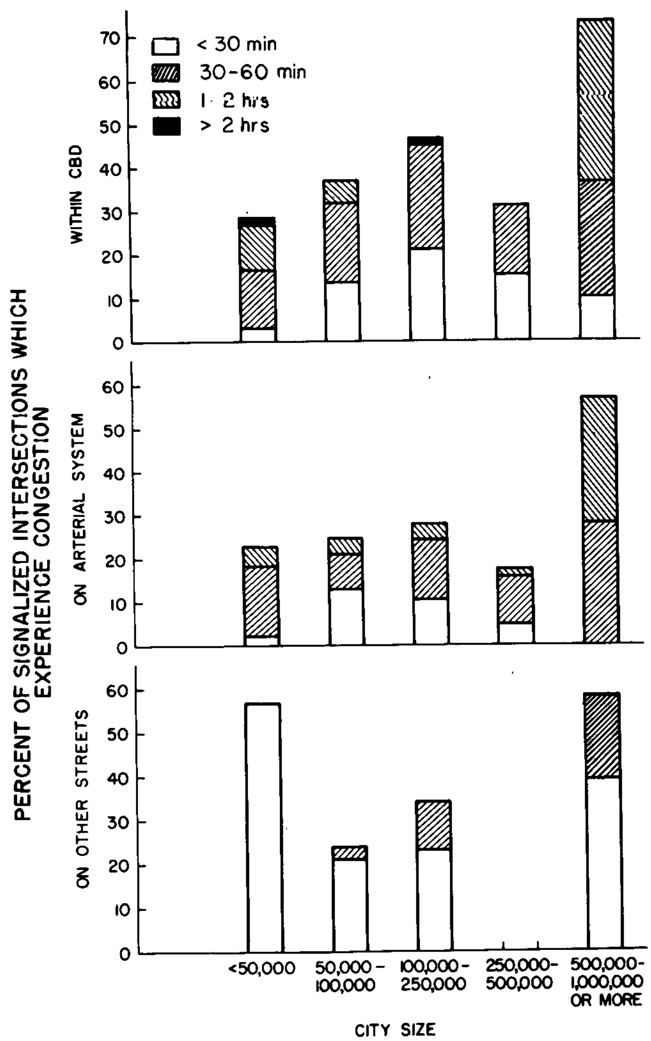


Figure 3. Extent and duration of traffic congestion as function of locale and city size.

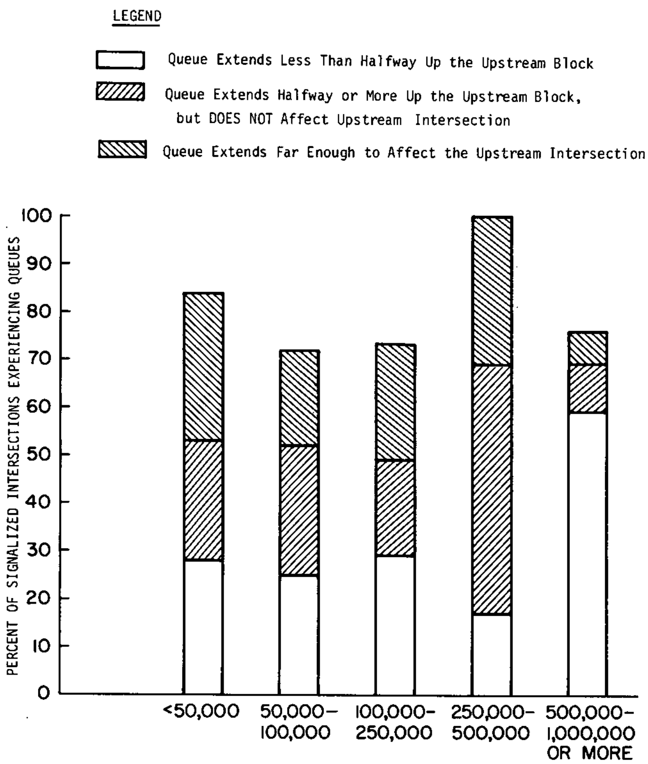


Figure 4. Queue length at congested intersections as a percent of all signalized intersections experiencing congestion, by city size.

section whose capacity is the limiting value of a segment of roadway or of an entire system. There may be more than one critical intersection in a system; it is most common that there is one critical intersection locally that has affected several blocks in one or more directions, giving the appearance (and fact) of areawide congestion.

1. *Congestion-Related Definitions. Oversaturation* represents a condition that occurs when queues of vehicles fill entire blocks and interfere with the performance of adjacent upstream intersections.

A review of the literature, sparse as it is, indicates an imprecision in the use and definition of the terms congestion, saturation, or oversaturation. It was not unusual to find all of the terms used to describe the same level of traffic operations. A more precise definition of the terms is developed herein to allow for a more exact description of the various traffic performance states. The specific definitions used have been constructed around a queue formation mechanism and, as such, they relate to the extent and growth of queues. It is recommended that they be adopted in all uses and classification of congestion.

The following traffic performance definitions are all described in terms of one approach to a signal. As such, they refer to the capacity of one signal's green time for the approach under consideration. The various regions defined are given in Table 1.

The term *uncongested operations* refers to a situation where there is no significant queue formation. Traffic performance may range from very low demand per cycle to conditions where the demand is a substantial fraction of the capacity value. Short queues may occasionally form toward the upper end of this performance range, but do not last for any length of time because the capacity of the signal to process traffic exceeds the demands being placed on it.

Congested operations characterizes the entire range of operations which may be experienced when traffic demand approaches or exceeds, or both, the capacity of the signal. As such, the term describes conditions as diverse as those where demand only slightly exceeds capacity for a time, and minimal queues form, to those where demand for service is so great as to cause segments of the system to jam because of extensive queue development. Since this is not a sufficient definition to describe the problem, the realm of congested operations has been segmented into two major categories: saturated and oversaturated operations.

Saturated operations is a term that describes that range of congestion wherein queues form, but their adverse effects on the traffic in terms of delay and/or stops are local. *Local effects* in this context means that traffic performance is only affected at the intersection at which the queue occurs, and that no other intersection's performance is affected by this queue.

Saturated operations has been further subcategorized into stable and unstable ranges. *Stable saturation* is considered to exist when a queue has formed and is not growing, and delay effects are local. *Unstable saturation* is considered to exist when a queue exists and is growing, and delay effects are still local.

Oversaturated operations is characterized as a situation wherein a queue exists, and it has grown to the point where the upstream intersection's performance is adversely affected. The upstream intersection's performance is considered to be adversely affected when its performance characteristics are determined by the performance of the downstream intersection from which the queue has developed, and not by its own physical and operational limitations.

2. *Types of Oversaturation—Categorized by Cause.* Intersection oversaturation may occur under varying conditions that can be classified into two groups: (1) nonrepetitive and (2) repetitive (the common peak period situations where demand exceeds capacity for periods of time). The *nonrepetitive* cases of oversaturation may be classified by their cause, which is usually apparent: accident, stalled vehicle, illegal parking, etc. In some situations, illegal parking is so common that this classification is somewhat artificial.

The *repetitive*, demand-based oversaturation situation may be classified by five basic types so as to better identify and reference the cause:

a. *Type I*—The critical intersection has a smaller green/cycle ratio than does the upstream intersection. Vehicles that turn in from the upstream cross street make only a minor contribution to the oversaturation of the critical intersection. The critical intersection yielding this type of oversaturation is usually an intersection of two major arterials or an intersection that requires a signal of more than two phases.

TABLE 1

TRAFFIC DESCRIPTIONS

UNCONGESTED	CONGESTED		
	SATURATED		OVERSATURATED
	STABLE	UNSTABLE	
NO QUEUE FORMATION	QUEUE FORMATION, BUT NOT GROWING DELAY EFFECTS LOCAL	QUEUE FORMATION AND GROWING, DELAY EFFECTS STILL LOCAL; (A TRANSIENT STATE MAY BE ONLY OF SHORT DURATION)	QUEUE FORMATION AND GROWING TO POINT WHERE UPSTREAM INTERSECTION PERFORMANCE ADVERSELY AFFECTED

b. *Type II*—The critical intersection and the upstream intersection have the same green/cycle ratio. However, the capacity of the critical intersection is less than that of the upstream intersection because of factors other than the green/cycle ratio, such as turning movements and/or physical conditions. Thus, when demand exceeds the capacity of the critical intersection, queues begin building up at the critical intersection, and cause spillback into the upstream intersection. It should be noted that oversaturation Types I and II may produce congestion conditions that look identical. Separation into the two different forms of oversaturation is useful in that each may respond differently to different potential solution techniques.

c. *Type III*—The critical intersection and the upstream intersection may have the same green/cycle ratio or the same capacity, or both. However, heavy turn-in movements from the upstream cross street fill up the entire link or a significant part of it during a red phase on the arterial and cause spillback on the arterial.

d. *Type IV*—This type of oversaturation of a critical intersection results from the signal offset between the critical intersection and its upstream intersection. It may occur under several different conditions. For example, when two intersections have two different cycle lengths or splits, or both, the varying offset causes queue build-up and spillback when the offset approaches the maximum possible value. This type of oversaturation may also occur simply because of the offset selections, especially on two-way arterials where achievement of "ideal" offsets is frequently difficult.

e. *Type V*—This is defined as being a combination of two or more of Types I through IV oversaturation.

Productivity vs. Quality of Performance

As part of the investigation of the impact of congestion and the relative ability of various measures of effectiveness to depict this impact (see Appen. D), it was established that the quality of performance at an intersection, as measured by intersection crossing times (i.e., speed), degrades substantially before the productivity of the intersection, as measured by its throughput, does.

Figure 5 shows a block divided into N segments, each segment being between 50 and 100 ft long. The length of the block determines the number of segments. The most upstream segment is labeled "1." As can be seen from Figure 6, the larger the downstream queue is, the longer it takes the upstream vehicles to cross the intersection; in other words, they go slower. The vehicles go even slower when they see that the signal they are approaching is still red (Fig. 7).

Figure 8 shows that the average time headway does not change as a function of downstream queue position. Combined with the foregoing results, this means that the

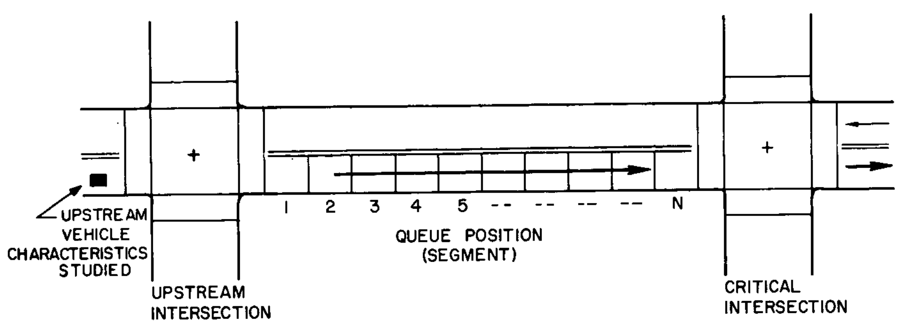


Figure 5. Definition of segments for queue study.

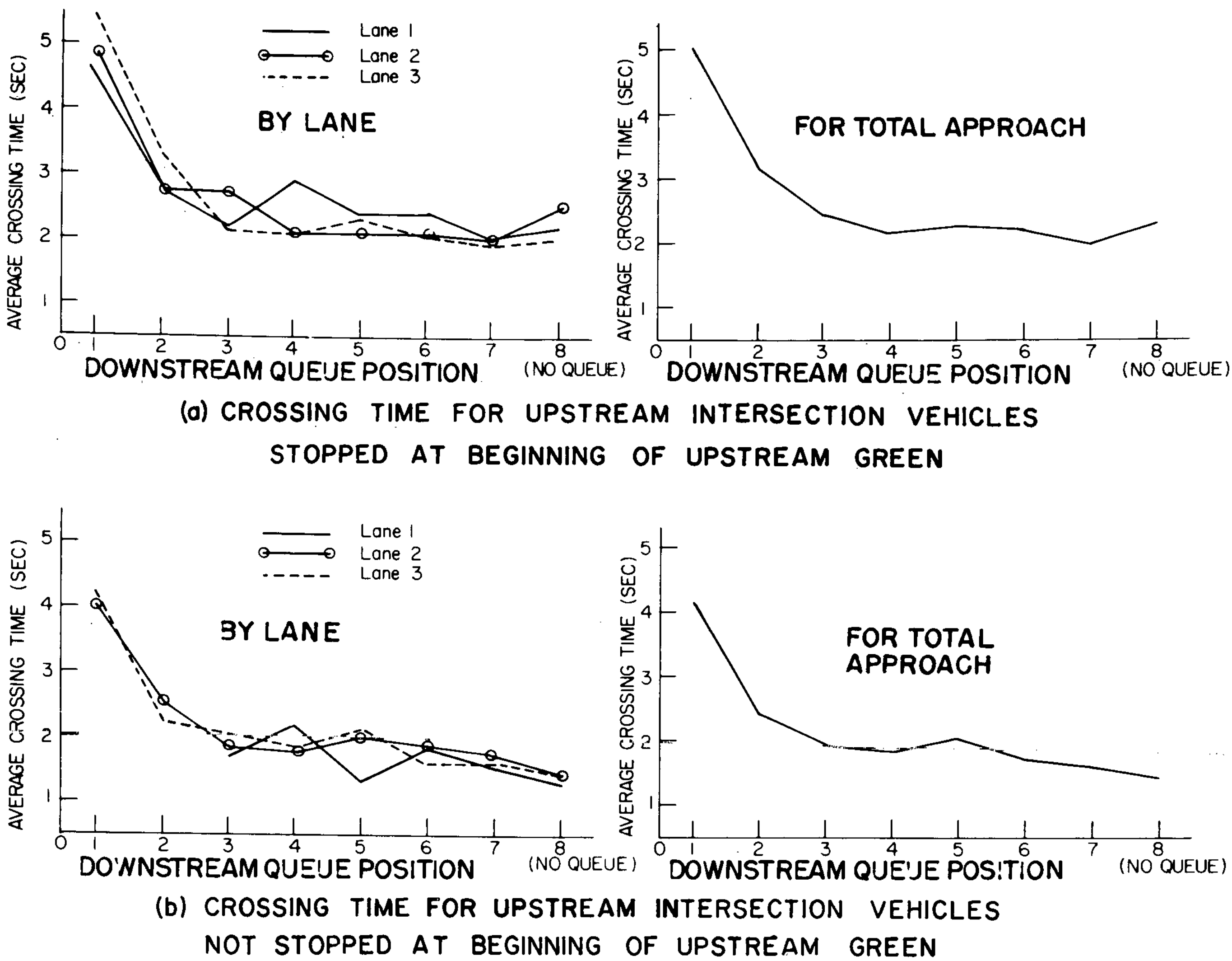


Figure 6. Average crossing time (one site).

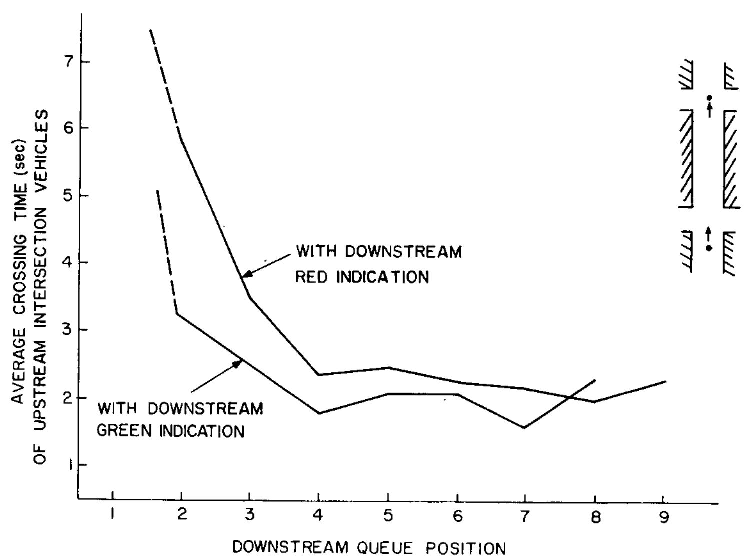


Figure 7. Average crossing time by downstream signal indication.

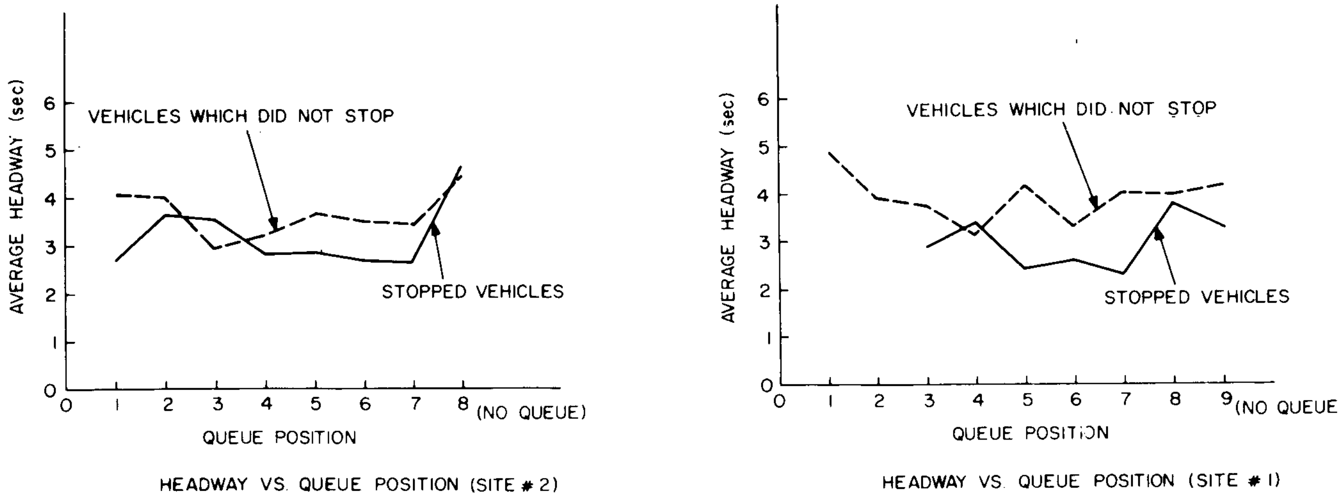


Figure 8. Headway vs. downstream queue position.

space headways (i.e., vehicle spacings) must be shorter because the vehicles are moving slower.

Clearly then, upstream drivers respond to downstream queue size by going slower when crossing the upstream intersection. Thus, they perceive the congestion. Moreover, these vehicles are more closely spaced than they would otherwise be. Still, the productivity of the intersection is not necessarily harmed.

From these studies, one may establish a rule-of-thumb: when downstream queues come within 200 to 250 ft of the upstream intersections, the upstream drivers will sense that congestion exists. This is quite apart from productivity considerations and even from delay as experienced in the downstream link. It is desirable to avoid this change in performance.

For completeness, it should be noted that the foregoing discussion is based on the queue that is present at the initiation of upstream green. Attempts were made to correlate to the average queue over the cycle. For vehicles stopped at initiation of green, it was an adequate substitute. Although instantaneous queue (that is, exact queue measurement) is preferred as a measure of effectiveness (MOE), average queue can be used as a proxy.

EVALUATING EFFECTIVENESS OF CANDIDATE TREATMENTS

Minimal Response Signal Control

Cycle Length and Block Length

In evaluating the effectiveness of signal techniques, two questions must be addressed: (1) Do long cycle lengths have any virtue in their own right? (2) Does block length enter into the cycle length determination?

One of the most prevalent, erroneous beliefs in the traffic engineering community is that the capacity of an intersection increases substantially as the cycle length is increased. This view has been a point of contention in the past (*1*), and studies (for example, *2, 3*) have provided data to support such questioning.

The lack of substantial capacity increases with increasing cycle length is rooted in at least three factors: the loss time per cycle is not that severe because of use of the amber and start-up delays lower than often assumed; the inefficient use of longer greens because of increasing headways; the inability to provide a demand to fill rather long green times. The last item is just another manifestation of the storage problem, addressed later.

The minimum cycle length is frequently expressed as a function of the known demand. The formula that expresses this relationship is

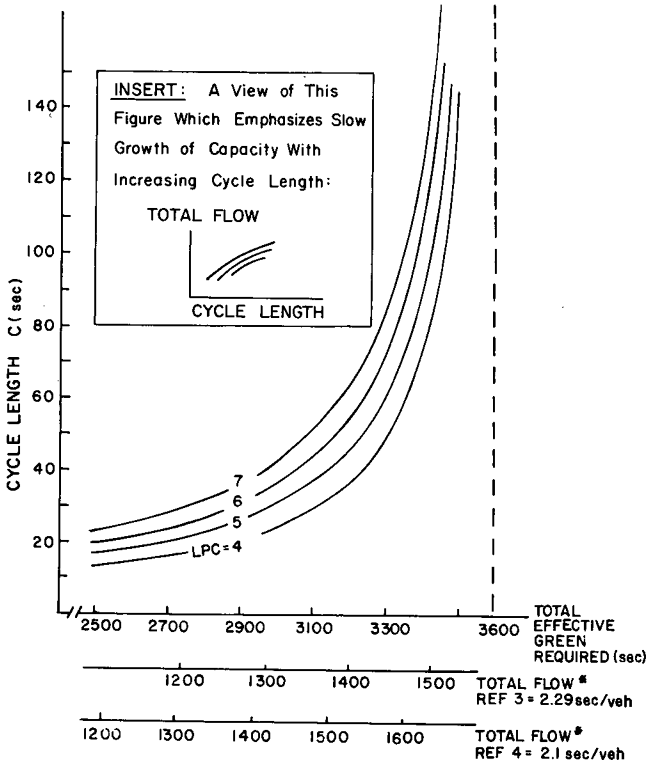
$$C = \frac{\text{LPC}}{1 - \dfrac{1}{3600}\left[\sum_{\text{phases}} \begin{pmatrix}\text{Seconds of Green per Hour}\\ \text{Required for Critical Lane}\end{pmatrix}\right]} \tag{1}$$

where C is the cycle length in seconds, and LPC is the loss time per cycle in seconds. This formula may also be interpreted as capacity as a function of cycle length.

Figure 9 illustrates this equation for a range of values of LPC. For a two-phase operation, LPC $\simeq 6$ sec is reasonable. The horizontal axes show the sum of green required for all critical lanes, and the corresponding sum of equivalent critical lane volumes for the average headways. The headway values were obtained from the literature (*3, 4*), and are consistent with values found in this study.

It should be noted that Figure 9 has two messages: (1) relatively short cycle length suffices for rather high volumes (under 60 sec for up to 1400 passenger cars per hour, pcph, for the sum of critical lanes, with 2.29-sec headways assumed); and (2) substantial increases in cycle length provide little corresponding increase in capacity (at $C = 80$ sec, 1450 pcph; at $C = 120$ sec, 1500 pcph). In the illustration, a 50 percent increase in C above 80 sec provides only a 3.4 percent increase in capacity. One must ask if there is not a better way to gain more increase.

By viewing Figure 9 (see insert in Fig. 9) so that cycle length is the x axis (abscissa) and total flow is the y axis



* SUM OF CRITICAL LANE VOLUMES IN PASS. CAR EQUIVALENTS.

NOTE: $LPC_{REF\ 4} = LPC_{REF\ 3} + N\ (3.7 - 0.75)$ FOR N PHASES

Figure 9. Cycle length required for various flows.

(ordinate), one can emphasize the slow growth of intersection capacity with cycle length.

In the foregoing, it has been shown that cycle length may not be as powerful a capacity-improver as one might think. On the other hand, there has been no evidence that short cycle lengths are of positive value in some cases.

If an approach has a flow of f_1 passenger cars per hour per lane (pcphpl) and a cycle length C, note that there are f_1 $(C/3600)$ vehicles per lane per cycle. A minimum condition is therefore that this platoon does not exceed the available link storage, for, if it did, the potential for spillback would exist. (It is assumed that the entire platoon is stopped, and saturation is approached.) Thus,

$$f_1\left(\frac{C}{3600}\right)\zeta \leq \mathcal{L} \tag{2}$$

where:

ζ = vehicle storage length, 20 ft (20 ft is used in some examples; to use other distances (23 ft is often used), simply substitute the desired value of ζ);
$\mathcal{L}$ = link storage, ft;
C = cycle length, (sec); and
f_1 = flow, pcphpl.

It should be observed that $\mathcal{L}$ need not be the physical length of the link; if a policy decision is made that the stored vehicles should come no closer than within 200 ft of the upstream intersection, $\mathcal{L}$ is 200 ft less than the physical length. Such a policy decision is in accord with the avoidance of the perception of congestion.

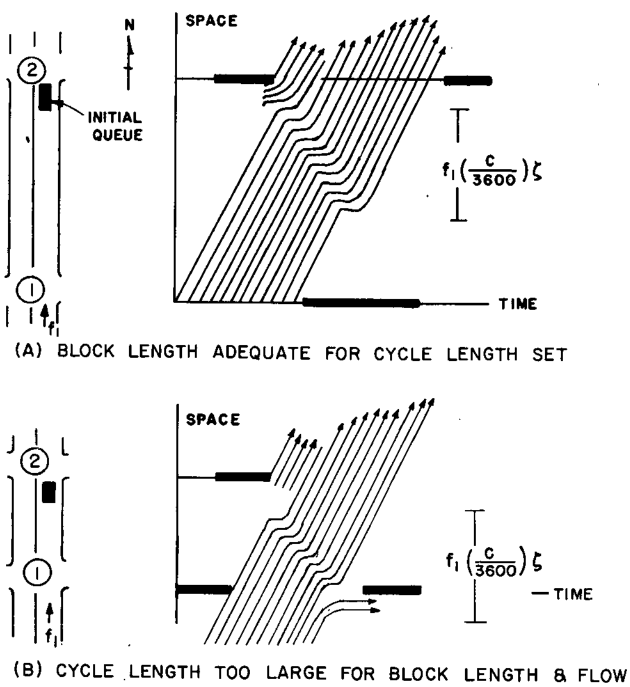
Figure 10 depicts the block length problem. (Note that if it could be assumed that the platoon could be moved through the intersection without stopping, the congestion situation would not, by definition, exist.)

In Figure 10(A), a normal delay to an input platoon is shown. The vehicles stored at the downstream intersection (the critical intersection) may have been turn-ins or may have been internally generated. However, if the upstream intersection were located as in Figure 10(B), the platoon would have been stopped as it tried to cross the intersection. The signal would have turned red before all f_1 $(C/3600)$ vehicles had a chance to enter the link. The input flow would be decreased, the output would be decreased, and the apparent congestion would show up on the next upstream link. Moreover, there is an excellent chance that some drivers will be physically stopped in the intersection proper when the signal changes. This spillback will block the east-west flow, causing delay to it.

In order to avoid this situation, it is necessary to assure that excessively long platoons are avoided. Eq. 2 may be rewritten as a constraint on cycle length:

$$C \leq \left(\frac{3600}{f_1}\right)\frac{\mathcal{L}}{\zeta} \tag{3}$$

Clearly, there are two contrary forces at work: as the total critical lane flow (all approaches) increases, the cycle length increase brings some benefit; at the same time, flow increases on any one approach decrease the maximum cycle



NOTE: In Part B, the stoppage presents some vehicles from discharging Intersection 1. It also causes conditions that could lead to a blockage of Intersection 1 (e.g., If there were any delays in moving through Link 1-2, the cross stream traffic at Introduction 1 would be quickly affected).

Figure 10. Block length problem.

length permitted. An example in the guidelines (Appen. J) illustrates the point.

Regularity of the Traffic Pattern

The regularity of daily traffic patterns is of basic concern when evaluating various candidate control strategies. At one extreme, if little variation exists from day-to-day, a time-of-day (minimal response) controller will undoubtedly suffice. At the other extreme, if no discernible pattern exists, or if it arrives randomly each day, then a highly responsive control is appropriate.

Between these two extremes, there is the possibility of a basic underlying pattern with substantial variation about it. There is also the possibility that on a given day there will be a major deviation from an otherwise extremely regular pattern. To accommodate such possibilities, traffic data may be collected in real time and deviations noted, or a predictor established to estimate future values.

When considering the control of congestion, it is necessary to consider (1) the regularity of daily patterns and (2) the regularity of the times at which various flow levels are reached. This will provide insight into the merits of minimal response control versus highly responsive control in this flow regime.

An extensive data base was acquired through the courtesy of the Metropolitan Toronto Department of Roads and Traffic, and was used to investigate the regularity of the weekday pattern; the regularity of the time at which certain flow levels are reached; the correlation of the flow between lanes on the same approach; and the potential for refining estimates of traffic volumes, based on observation of the deviations from the historic or nominal pattern. These topics are treated in detail in Appendix E.

Table 2 summarizes the amount of data available in 5-min samples at each of the 4 detector pairs. The average pattern was computed for each day by averaging the volume observed in each time slot. Because the weekdays did not differ substantially, a single average pattern was also computed aggregating all weekdays at each detector. Figure 11 shows the average weekday pattern, and the Saturday and Sunday patterns, for Site 2. Also shown is the average pattern for Monday and Friday.

Figure 12 illustrates the average weekday pattern at each of the four sites. The shaded area indicates the region within which one can expect 95 percent of the observations of volume to fall. This is not a confidence bound on the average. The confidence bound is much tighter. It is an estimate of the fluctuations from day-to-day within a specified time slot.

It can be observed in Figure 12 that, although there is substantial variation most of the time, the peaking is quite sharp, with relatively little variation in the time at which certain levels are reached. That is, a specified level, X, is reached at approximately the same time every day; this is variation in the horizontal dimension. It can even be questioned whether the variation in the vertical dimension is as "substantial" as it appears: the 95 percent range is in the order of $\pm$ 120 vphpl, or $\pm$ 10 vpl in a 5-min period. This occurs when the average count is in the order of 40 vpl in a 5-min period. Further, the distribution is approximately normal, so that the deviations tend to be clustered near zero.

Figure 13 shows the distribution of the times at which certain volumes are reached at Site 2. Note that the variation is rather small, leading to the conclusion that minimal response policies could be developed with some assurance that the onset of certain levels could be anticipated with some confidence. At the same time, the variation that does exist precludes extremely rapid preplanned switching of the control settings.

It should be noted that if one takes advantage of the regularity of pattern and does not have surveillance detectors, there is no protection against statistical "rare events"—major deviations from the average pattern. Generally, these can be associated with weather or special community activities.

Equity Offsets

It is unfortunately true that it is not always possible to avoid spillback because there may be too many vehicles attempting to enter a particular link. With extreme traffic congestion, it is not uncommon to see vehicles storing themselves in the intersection, to the great disadvantage of the cross traffic. Figure 14 illustrates such a situation.

One common solution to this spillback problem is to place a skilled traffic control officer at this site to prevent such events. Another approach is an intensive ticketing program for such offenders. The former approach not only is historically more effective, but it also is the one demanded by a public afflicted with spillback.

A possible alternative solution exists by changing the basic concept of what the offset is supposed to accomplish. However, this solution should not be implemented until it is certain that a better offset cannot alleviate the problem. The treatment to be presented now is only for that period after the best possible offset has failed because of the sheer volume demanding access to the link.

The treatment, shown in Figure 15, is as follows: allow the oversaturated direction to have green until the vehicles blocking the intersection just begin to move; switch green to the cross traffic; allow the cross stream to move until it has had an equitable input into the oversaturated link, or at least into the intersection.

The offset that accomplishes this function is defined herein as *equity offset* because of the objective of equitable treatment of the competing flows. It is not determined by the standard relations. The upstream red should begin L/V_{ACC} sec after downstream green initiation, where V_{ACC} is the acceleration wave speed in feet per second. Assuming g_1 as the green time at the upstream intersection (percent of cycle) and g_{CI} as the green time at the critical intersection, thus

$$t_{off} = g_1 C - \frac{L}{V_{ACC}} \tag{4}$$

where C is the cycle length in seconds. Typically, $V_{ACC} \simeq 16$ fps.

There are actually two cases of common interest. In the first case, there are negligible turn-ins from the cross

TABLE 2

DISTRIBUTION OF SAMPLES BY HOUR OF DAY AND DAY OF WEEK

Hour \ Day	1	2	3	4	5	6	7	Total	Percent
1	115.	69.	46.	38.	52.	92.	104.	516.	3.9
2	69.	0.	0.	12.	12.	26.	83.	202.	1.5
3	96.	26.	0.	12.	10.	0.	72.	216.	1.6
4	96.	56.	0.	35.	13.	14.	72.	286.	2.2
5	97.	80.	29.	68.	23.	36.	79.	412.	3.1
6	129.	116.	74.	89.	99.	75.	96.	678.	5.1
7	143.	113.	97.	109.	129.	104.	113.	808.	6.1
8	144.	132.	117.	120.	132.	108.	120.	873.	6.6
9	139.	133.	118.	120.	132.	94.	65.	801.	6.0
10	135.	128.	120.	120.	131.	82.	0.	716.	5.4
11	136.	132.	108.	104.	124.	57.	0.	661.	5.0
12	132.	131.	108.	96.	116.	43.	0.	626.	4.7
13	67.	44.	45.	47.	61.	48.	0.	312.	2.3
14	131.	117.	116.	81.	94.	48.	0.	587.	4.4
15	133.	120.	120.	86.	101.	48.	0.	608.	4.6
16	141.	120.	117.	97.	127.	57.	2.	661.	5.0
17	144.	132.	132.	120.	132.	77.	24.	761.	5.7
18	144.	131.	132.	120.	132.	89.	29.	777.	5.8
19	118.	114.	114.	100.	126.	95.	72.	739.	5.6
20	57.	32.	15.	45.	95.	103.	106.	453.	3.4
21	15.	0.	0.	1.	91.	108.	117.	332.	2.5
22	8.	0.	0.	7.	84.	103.	120.	322.	2.4
23	52.	30.	8.	45.	86.	96.	119.	436.	3.3
24	70.	46.	28.	57.	92.	101.	114.	508.	3.8
TOTAL	2511.	2002.	1644.	1729.	2194.	1704.	1507.	13291.	100.0
PERCENT	19.	15.	12.	13.	17.	13.	11.	100.	0.0

NOTE: Day 1 is Monday; Hour 1 is Midnight to 1 AM

traffic. The split at the upstream intersection can be as commonly determined.

In the second case, there are substantial turn-ins. In general, the cross street should be allowed just enough green to put its "fair share" of vehicles in the oversaturated link. If the turn-ins "steal" all the storage space, the arterial through vehicles cannot enter and the oversaturation will propagate rapidly up the arterial.

It is counterproductive to prepare equations or nomographs to cover all approach width combinations and volume/turn combinations. The following principles should govern:

1. The cross-street green should not be so long as to allow turn-ins to take disproportionate amounts of the storage in the oversaturated link.
2. It should be recognized that if the turning demand accumulates, it will interfere with the through movement on the cross street. If possible, storage should be provided.
3. If the turns are indeed significant, and storage and capacity exists, establishment of turn lanes with separate signalization should be considered. In this way, the through movement could be continued.
4. In allowing such cross-through movements, it should be recalled that the through arterial movement needs only as much green as it can effectively use at the critical intersection. Any additional green is, in fact, wasted, and is only allocated to the arterial to keep it from the cross turning movement (if signalization by movement does not exist).
5. If necessary, turn prohibitions should be considered, so that the cross stream through movement is not severely impacted by queued turning vehicles.
6. If none of the foregoing is feasible, refer to the guidelines in Appendix J, on the regulation of the spread of (unavoidable) congestion from the upstream intersection.

Note that a hierarchy of solutions has been portrayed (see Appen. J, Fig. J-21). More hierarchies of this sort are also presented in the guidelines, where nonsignal options are considered.

Now, recall the assumption that the oversaturation is unavoidable in the one link. And only then is equity offset implemented. The hierarchy applies to other (cross) links. It does not apply to the original (unavoidably oversaturated) link.

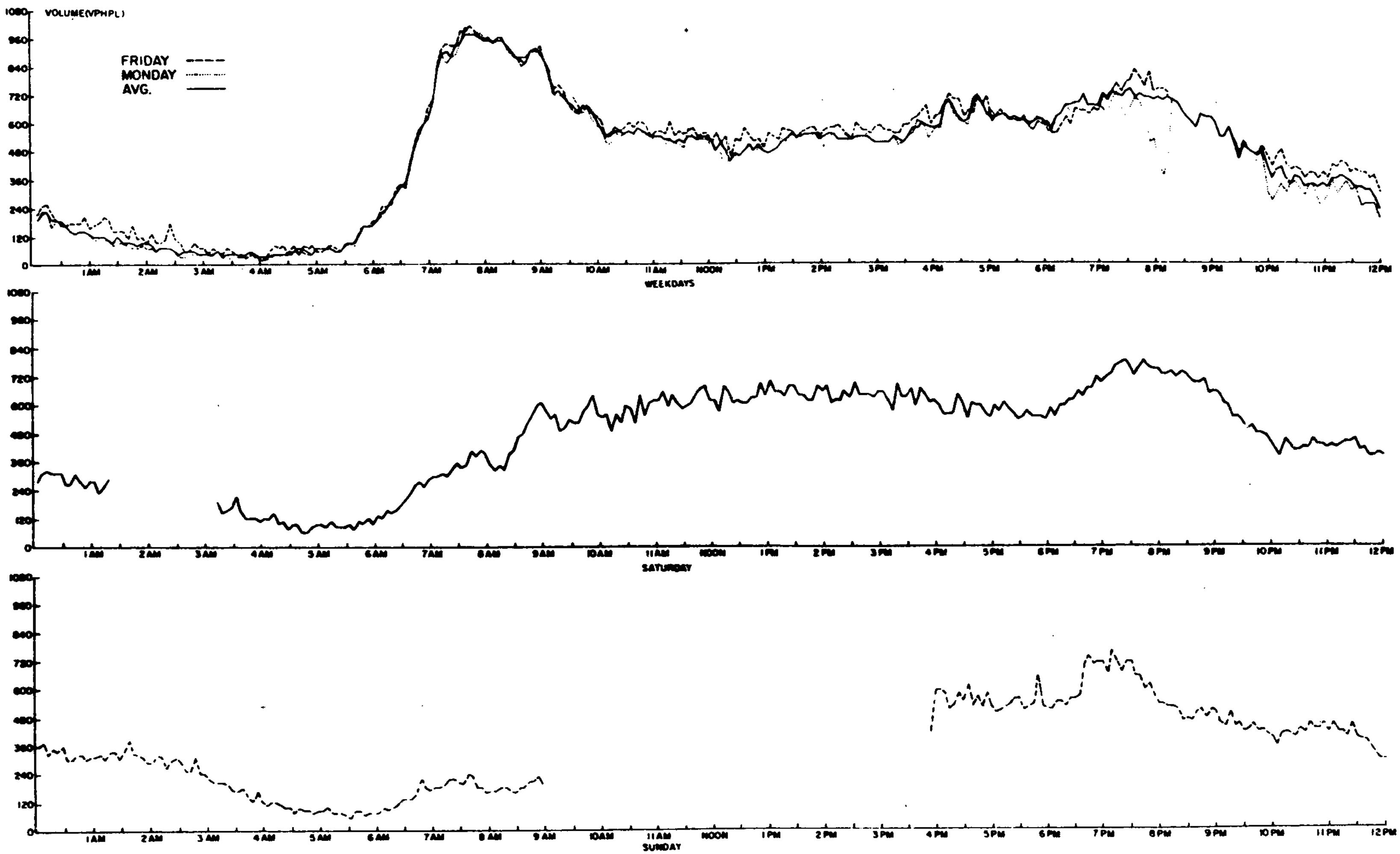
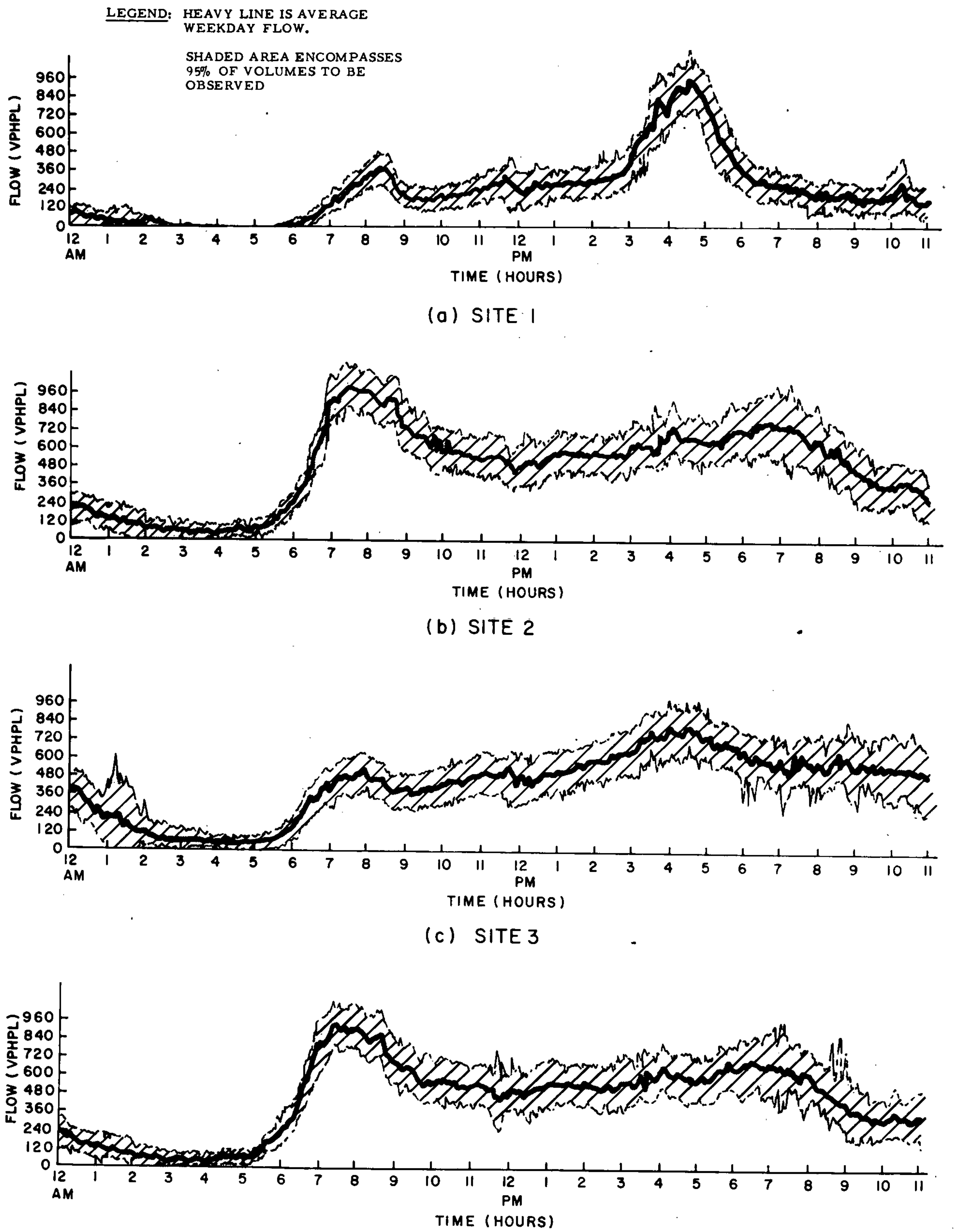


Figure 11. Average patterns at Site 2.



(b) SITE 4

NOTE: (1) SITES 2 AND 4 ARE ONE BLOCK APART ON SAME STREET, IN SAME DIRECTION

(2) ALL SITES ARE TWO MOVING LANES.

Figure 12. Average weekday pattern and variation of individual flows.

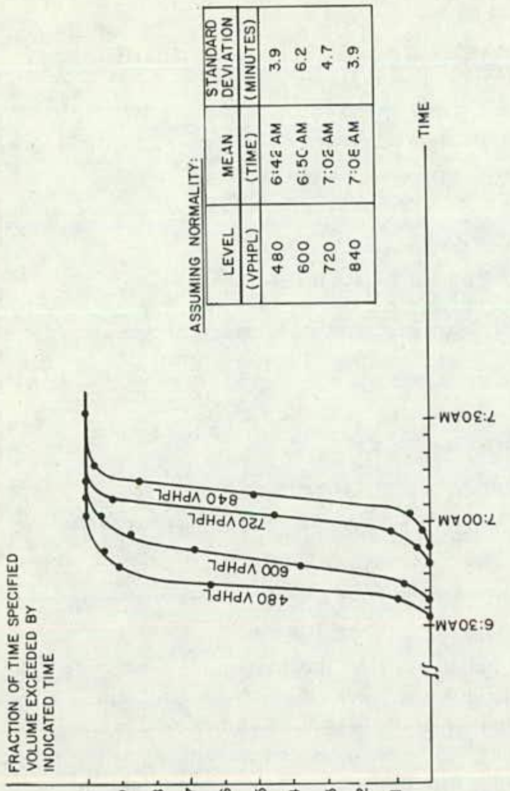


Figure 13. Times at which certain levels reached (Site 2).

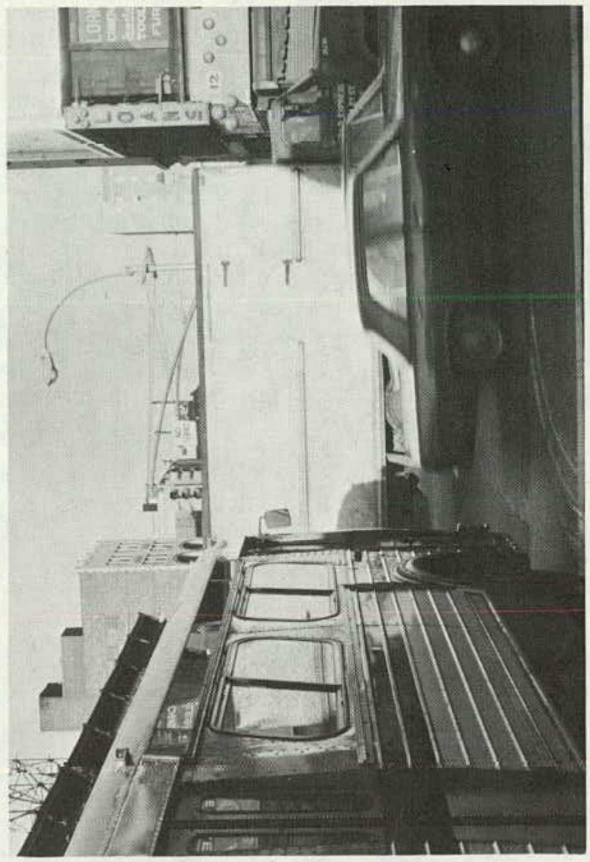


Figure 14. Intersection blockage.

The original link must have unavoidable saturation: neither any signal nor any available nonsignal remedy could have helped it. Only then is this link given up, and the best possible done for other traffic.

Extensive simulation testing indicates the benefits of this offset policy. Illustrations are given in Appendix J.

Highly Responsive Signal Policies

The guidelines give prime emphasis on minimal response signal policies and on nonsignal remedies used in conjunction with such signal policies. Highly responsive signal control policies are not emphasized for three reasons:

1. The regularity in the traffic demand from day-to-day typically supports the utility and viability of minimal response policies.
2. The extensive detectorization and computer-oriented control required with such policies are beyond the resources, both budgetary and staff orientation, of most jurisdictions.
3. There is major work underway, or relatively recently completed, on other projects addressing this topic. The work in Washington, D.C. (three generations of the Urban Traffic Control System (UTCS) and in Toronto, Canada, are prime examples.

It should be noted that standard actuated equipment is quite validly classified as a "highly responsive/intersection" type of control and properly set actuated equipment can, indeed, be used to reduce congestion at a

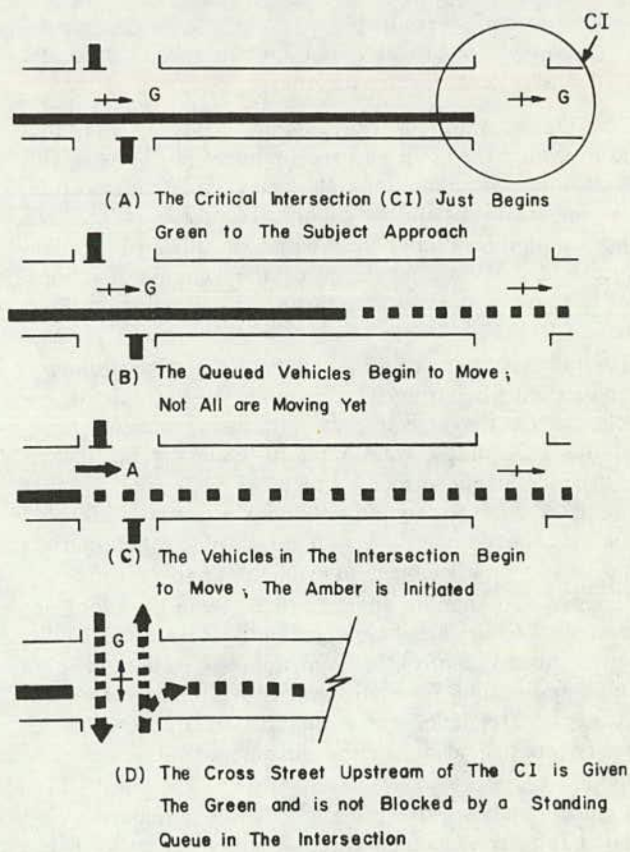


Figure 15. Illustration of equity offsets.

site. The timing of such equipment, including proper detector location, is covered in detail in Ref. (5) and the references cited therein.

However, it must also be realized that, as congestion tends to saturation and oversaturation, cycle length and offset impose a necessary discipline on traffic. Thus, the guidelines—with their minimal response considerations—become more applicable.

The following discussion briefly outlines the relative advantages and disadvantages of some congestion-tailored responsive policies that appear in the literature. It also observes that an additional responsive policy was developed in the research.

Three policies were deemed applicable to the congestion/saturation problem of interest:

1. Smooth flow concept.
2. Minimum delay via split switching (Gazis).
3. Queue proportionality in real time (Longley).

These policies are reported in Refs. (6), (7), and (8), respectively. A fourth policy was proposed by the research agency:

4. Queue-actuated control (Appen. F).

The smooth flow concept has as its primary objective the realtime adjustment of the offsets:

$$t_{off} = \left(\frac{L}{V} - \frac{Q}{R}\right) \tag{5}$$

where:

L = link length, ft;
V = desired speed, fps;
Q = queue length at green initiation, veh/lane; and
R = service rate, veh/sec/lane.

The Gazis' minimum delay policy finds an optimum time to switch the split, and the optimal split values, with the objective of minimum total delay at an intersection. It is not truly an on-line highly responsive policy, but rather an optimization policy based on advanced mathematics and detailed knowledge of the demand functions. The Longley policy is a real-time, rapid adaptive split policy.

The queue-actuated policy switches from phase-to-phase when certain (determined) maximum queue extents are reached, with the objective of minimizing average delay while not exceeding certain limits of queue extent. It does not depend on queue measurement as such (i.e., number of vehicles), but rather on attaining a certain threshold value (i.e., queue of 12—yes or no). Table 3 summarizes some of the disadvantages of each of these policies.

Extensive simulations indicate that of the three policies first cited, Longley's policy is generally the most effective. Queue-actuated control is even more effective at high volume levels and in environments with heavy turn-ins. Because of its deliberate formation of queues, queue-actuated control also has the advantage of greater productivity.

Figure 16 shows the average delay experienced at an isolated intersection under three different control policies. The degree of saturation is as defined by Webster—"the ratio of the actual flow to the maximum flow which can be passed through the intersection from one arm."

TABLE 3

DISADVANTAGES OF SPECIFIC HIGHLY RESPONSIVE POLICIES

SMOOTH FLOW CONCEPT

- Measurement of Queues Required
- On-Line Computations

GAZIS' MINIMUM DELAY

- Demand Function Description Needed
- Queue Extent Can Be Excessive
- Applicable Only to Convex Portion of Demand Curve

LONGLEY'S POLICY

- Measurement of Queues Required
- On-Line Computations

QUEUE-ACTUATED CONTROL

- Detectors Located According to Competing Demand (Historic)
- On-Line Computations

Queue-actuated control can be considered as a viable control policy in situations where minimal response policies would, or do, fail (that is, minimal response policies in conjunction with nonsignal remedies). This does not mean that minimal response policies do not perform better in many situations; indeed, minimal response policies are generally better on a system basis.

Nonsignal Treatments

Enforcement and Prohibition

Two of the most frequently persistent violations that aggravate the congestion/oversaturation problems are intersection blockage and parking regulation violations. Control by equity offset represents an attempt to circum-

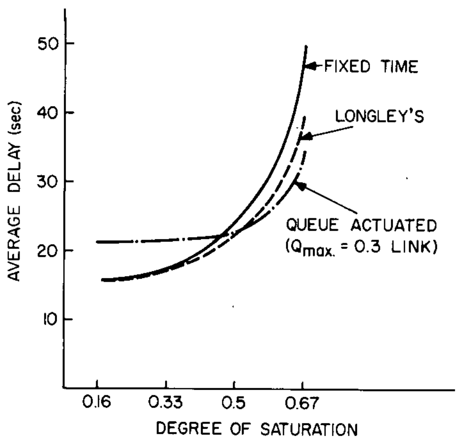


Figure 16. Delay comparison of different control measures: fixed time, Longley's control, and queue-actuated control at an isolated intersection.

vent the first, avoiding or delaying the need for on-scene traffic control officers.

The following discussion illustrates the magnitude of the impact that the second problem (parking violations) can have. Perhaps this will indicate a level of unacceptability to the individual engineer, which will make his decision clear in some circumstances—strict enforcement.

Figure 17 shows the results of a simulation of the impact of double parkers in a 600-ft 3-lane link. The same curve could be used for three moving lanes without an additional (i.e., a fourth) lane for parking. The figure was computed for various V/C ratios (ratio of demand volume to capacity) of the moving traffic, various double-parking positions, and a double-parking duration of 12 min.

Obviously, double parkers have a substantial impact for volumes for which V/C exceeds 0.75, and the impact is least severe when the double parker is at the center of the link. It is most severe when he is at the discharge end of the link. The explanation lies in the opportunity for passing maneuvers, or lack thereof, and in the reduction of discharge capacity or entrance width. Moreover, not only does the impact grow substantially with V/C ratio, but the impact also persists well beyond the double-parking duration at the higher V/C ratios.

Note that this curve can also be used to estimate the impact of adverse uses of a curb lane from which parking was removed so as to increase capacity. For longer durations of double parking, the values in Figure 17 can be proportionately adjusted so that it can serve as an approximation (the values in Figure 17 can be multiplied by (15/12) 100% = 125% if, for instance, a 15-min duration is being considered).

Conversely, local realities may dictate that strict enforcement is not practical. These realities may include the fact that many of the vehicles are goods vehicles essential to local vitality, or that the vehicles would only circulate, adding to the apparent traffic flow. Note that one vehicle circling a city block once every 4 min can add 15 vph to the count on 4 streets in 1 hr.

Thus, the estimated impacts set levels that must be remedied by other treatments contained in the guidelines (Appen. J).

If neither decision is clear cut (i.e., strict enforcement or compensation via other treatments), the engineer may wish to formulate a cost-benefit or cost-effectiveness analysis to aid him. Delays, such as those indicated in Figure 17, can be translated into costs. However, development of such a methodology is outside the scope of this project.

Right-Turn Bays

A detailed analytical discussion of turn bays is contained in Appendix G. As a result of that work, it is apparent that the creation of a right-turn bay allows:

1. An increase in productivity or—if it is desired—a decrease in the effective green allocated to the phase.
2. A decrease in the local delay, with the right turners realizing most of the delay savings.

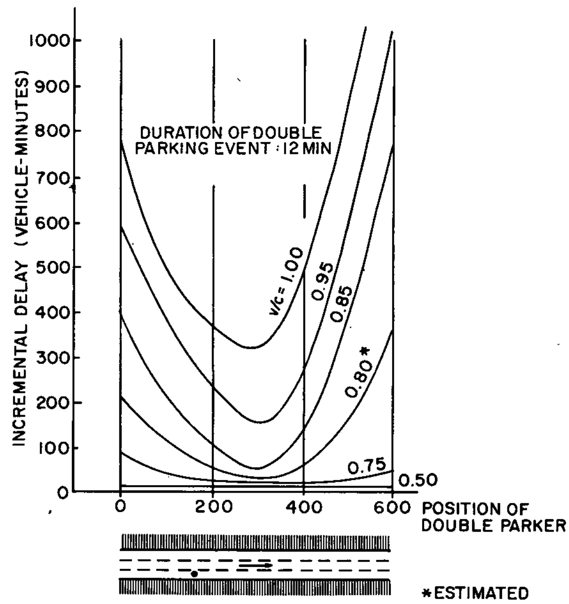


Figure 17. Illustration of delay caused by double parker.

The increase in productivity, expressed as a percentage, can be comparable to the right-turn percentage. The length of the turn bay should be approximately the same length as or slightly longer than the queue that typically forms. In this way, maximum presorting can occur. Thus, in practical terms, short cycle lengths and this objective complement each other, because released platoons are smaller and the necessary length is easier to achieve.

An example in the guidelines illustrates, however, that much of the benefit is achieved by the existence of a bay of even moderate size. Nonetheless, short cycle lengths aid presorting, and should be used as a companion measure.

Consider a 2-lane arterial with input flows capable of oversaturating the system because of the 50/50 split at the critical intersection. Figure 18 presents the spillback history (seconds of spillback in a 30 min simulation) along the arterial upstream of the critical intersection. It is interesting that the turn bays, which serve some benefit to the system, do not necessarily provide a remedy. The volumes do assure that there will be oversaturation eventually. In Link 2, where the critical intersection implies longer queues, it is only the longest bay length that has any impact. Link 3 has more discharge capacity than Link 2, and does not benefit as much from the turn bays. (Recall from the very definition of a critical intersection that Link 2 could not previously handle the demand being put on it. That is, its discharge point—the intersection with the 50/50 split—could not.) Link 4, being farther away from the critical intersection, had shorter queues and thus benefited more than Link 3.

Figure 19 presents the spillback history with the backward progression in force. The backward progression approximates an equity offset for the case at hand. Two results emerge:

1. The signal treatment is more effective than the nonsignal remedy (comparison shown in Figure 18 by the dashed curve—offset helped more than turn bays).

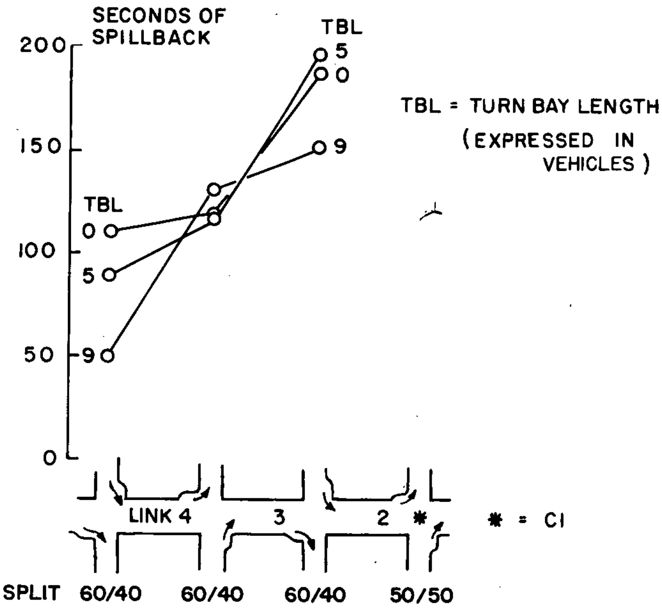


Figure 18. Spillback in a right turn illustration.

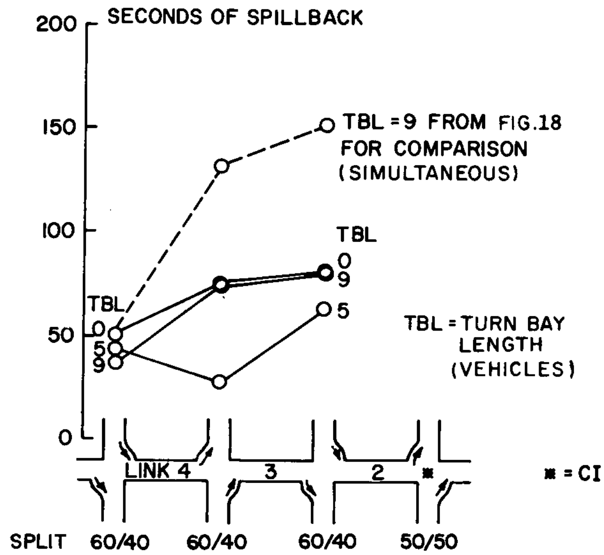


Figure 19. Spillback in a right turn illustration using backward progression.

2. Longer bays may actually do harm. In this particular case, this is so because turners are being given preferential access to the arterial, contrary to the intent of the oversaturation settings represented by the (approximately) equity offsets.

The engineer must take care not to confuse his objectives: in this case, he may select cross-traffic enhancement via equity offset, but must be careful about the turners' enhancement via long turn bays to make sure that the preferential access does not create yet another problem.

Other Nonsignal Remedies

The guidelines in Appendix J incorporate consideration of all remedies (possible treatments; see also Table 7), including left-turn bays.

The results of the research indicate that right-turn-on-red (RTOR), which is considered as a possible remedy, can do significant harm because turning vehicles have unrestricted access to congested arterials, "stealing" space and transferring the delay to the arterial inequitably. In congested/signalized environments, RTOR must be used judiciously.

The magnitude of pedestrian-induced traffic disruptions is shown to be significant only at the most extreme combinations of vehicular V/C ratio and of pedestrian intensity, for normal turning percentages.

Other results are incorporated into the guidelines, including rules-of-thumb on when to use simultaneous and other progressions, rules-of-thumb for productivity increases due to left-turn bays, circumstances under which multiple phases may actually not increase the cycle length, illustrations of the relative impact of alternative treatments, and flow charts or "decision charts" for considering some sets of treatments. By referring to the guidelines, the engineer can better sort out the issues in his own application and consider more implications.

UTCS-1 SIMULATOR

The simulations were run using the UTCS-1 model simulator, which had been developed and validated as part of the Federal Highway Administration's Urban Traffic Control System (UTCS) program. (The UTCS-1 simulator is the predecessor of the NETSIM simulator.) To independently ensure the validity of the simulator for the type of simulation under consideration, a calibration was performed using a section of Flatbush Avenue Extension. Traffic volume distributions and signal timing of the street are contained in Appendix H.

Since the type of simulation under consideration is directly related to queue length and link occupancy, the link occupancy by simulation for the critical link is shown in Figure 20 with the actual link occupancy obtained from the films taken for data collection (see Appen. D). The results are shown to be in good agreement. Analysis indicates that the coefficient of correlation between the two lines ($R = 0.86$) is statistically significant.

It was found useful to also write some programs that manipulate the optional UTCS-1 outputs. Three additional manipulations were of interest:

1. A side-by-side comparison of queue versus time for any two cases, for a specified link.
2. A side-by-side comparison of occupancy versus time for any two cases, for a specified link.
3. An aggregation of sets of links (actually, of movements) so as to present some relevant statistics—output, delay, etc.—by these aggregations for successive intervals.

The programs written to accomplish these manipulations are contained in Appendix I.

CONCEPT OF THE GUIDELINES

This section outlines the concept and structure of the guidelines contained in Appendix J, and highlights some of the findings incorporated into the guidelines.

The nature of the problem—traffic congestion—virtually assures that one is dealing with signalized intersections. However, it does not follow that one has only signal remedies at hand. Indeed, the possible treatments may be broadly classified as signal (timing and coordination) and nonsignal.

Within the signal classification, there are two major subclassifications: minimal response (i.e., preplanned) and responsive signal control. It is not at all well established that highly responsive control is justified over preplanned signal control, particularly for the heavier flow range. This is an indication that is being reinforced by trends in major computer-based research projects. Not only is traffic flow rather regular from day-to-day, but the practical limit on detectorization (due to economics) does not allow sufficient information to be extracted so that a highly responsive plan can truly be implemented. Even if it were, it is not certain that it would be cost effective.

Within the nonsignal classification, there are also two major subclassifications: regulatory and operations. Regulatory consists of enforcement and of prohibitions. Operations, as used herein, consists of all other traffic measures.

Within this section, the following topics are addressed:

1. What improvement is to be sought?
2. What solutions are available?
3. What should be done initially?

Knowledge on these topics, combined with the "plan of attack" presented later, represents the recommended framework for dealing with the problem of congestion and saturation.

Improvements Desired

Before enumerating possible treatments, it is appropriate to dwell on the ends to be achieved: what improvements is one seeking?

Table 4 lists a set of the most common improvements that a traffic engineer may wish to make, faced with a traffic congestion problem. These improvements are stated in the broad terms usually encountered as goals or objectives. A closer inspection of the items on this list, however, establishes that, depending on the traffic levels (relative to capacity), some of these are truly relevant, and others are merely consequences of these primary items. This list is not operative; that is, it does not say how to do various things. Depending on the observed appearance of the problem, there is a logical predisposition to how much of the traffic stream can be helped. Of those that can be helped, there are two groups: the principals involved in the traffic situation and contributing to it, and those affected by the traffic situation caused by the principals.

Table 5 summarizes the traffic stream flow components and their nature in terms of an identified critical intersection (an intersection whose capacity is the limiting value of a segment of roadway or of an entire system). The engineer's task is twofold:

1. Alleviate, or prevent, the traffic congestion problem to the maximum extent possible.

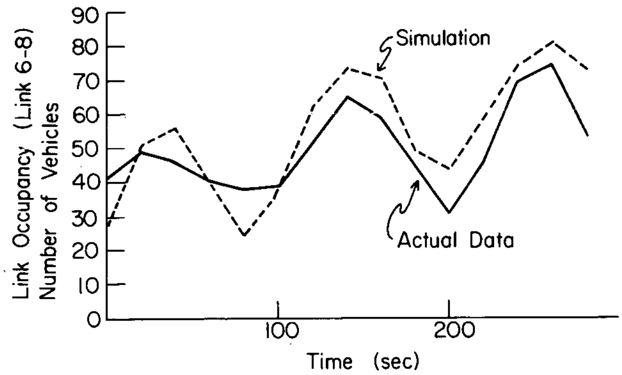


Figure 20. Validation of simulation.

2. While doing this, provide as much benefits as possible to groups adversely affected, but without harming the over-all remedy.

As can be seen in Table 5, the turning movements at the critical intersection itself can be classified in both categories. The treatment of these movements depends on the particular problem manifested and on the relative magnitude of the movement in contributing to the criticality. To illustrate: a right-turn bay may provide some increase in throughput capability (thus providing some alleviation of the criticality), but its principal impact may be significant savings in delay that would otherwise be encountered by right turners. The bay may be justified strictly on the basis of aiding this one flow component.

It was noted earlier that some of the items in Table 4 are actually secondary to other items in that table for given flow levels. Table 6 summarizes the primary objectives that should be sought by the engineer—these are dependent on the traffic level; first and foremost, the engineer must realize that, at the more extreme flow levels, it is most important to "avoid spillback." This is the explicit objective. The mathematical niceties of minimum stops or minimum delay collapse in the face of intersections blocked by vehicles.

Some comments on Table 6 are in order. First, it should be recognized that the primary objective to which the engineer should address himself does change depending on the flow level. Secondly, the major index of performance (or measure of effectiveness—MOE) also changes. However, both sets are well correlated to queue extent measures, so that queue or occupancy patterns—particularly during red and at the onset of green—are good indicators across the entire range of conditions.

It is appropriate to also call attention to the last item in Table 4—"improve the regularity of service." During simple congestion, the variance as well as the mean of the delay per cycle increases as the demand approaches capacity (many readers are probably acquainted with the exploding nature of the average queue as the demand rate approaches the service rate; the same sort of queuing analysis also reveals that the variance—the spread of values—also explodes as the demand rate approaches the service rate). Thus, the individual driver will be exposed

TABLE 4

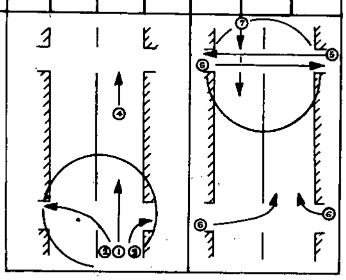
LIST OF IMPROVEMENTS DESIRED

- Reduce Geographic Spread of Congestion
- Reduce Rate of Spread of Congestion
- Increase Throughput
- Reduce Delay
- Reduce Stops
- Improve Regularity of Service

TABLE 5

FLOW COMPONENTS AND THEIR CLASSIFICATIONS

	MAIN STREAM(S)				CROSSING STREAM(S)		
PRIMARY NATURE OF FLOW	① THRU	② LEFT	③ RIGHT	④ DOWN STREAM	⑤ CROSSING	⑥ TURN-INS	⑦ OPPOSING DIRECTION
CONTRIBUTING TO THE CRITICAL SITUATION	*	*	*	–	–	*	–
IMPACTED BY THE CRITICAL SITUATION	–	*	*	*	*	–	*



NOTE: CRITICAL INTERSECTION CIRCLED FOR EMPHASIS

to greater variability in his individual experience from day-to-day. Improvements that minimize the mean delay will also enhance the regularity of the delay suffered.

Likewise, on the higher flow levels, the actual blockage of intersections occurs probabilistically. If the potential for blockage is increased (by filling the downstream link), a variable phenomenon is introduced as to whether there will actually be blockage at a fixed time after the potential occurs. Avoiding the potential thus contributes to enhancing the regularity that the driver perceives.

There are times when a basic solution has been achieved, and some possible benefits remain to be realized by additional improvements addressed to specific subgroups. Right-turn bays are such a case because they frequently have little impact on such a measure as average delay of all vehicles, but thus have truly substantial benefits for a smaller segment of the traffic stream—the right turners.

Solutions Available

Table 7 summarizes the range of solutions available. There is no simple statement of an ordered list of recommended solutions, in decreasing order of preference. There are, however, indications of how much of one solution must be accomplished to have an equivalent impact of so much of another solution. The engineer must use this knowledge in conjunction with local conditions and practices to reach a decision. There are also indications of how best to use two solutions in conjunction with each other. An initial classification and elimination checklist that can be followed is contained in Table 8.

In brief, the engineer must reach a preliminary judgment of the underlying cause of the problem. At the same time, he must be aware that the solution is not clear-cut. For example, undesirable or extensive queues may lead one into the depths of these guidelines too quickly; such problems can arise because of poor offsets, outdated splits, and excessive cycle lengths. Therefore, having prepared a preliminary opinion on the underlying cause, the engineer must choose between a number of possible solutions. Much of the guidelines is addressed to the candidate solutions, the considerations involved, and the relative merits. The following summarizes the recommended "plan of attack" in the treating the problem of congestion and saturation.

Plan of Attack

It cannot be overemphasized that it has not been possible to develop unequivocal statements of when particular techniques or combinations of techniques are better than others. However, certain categorical statements can be made, and a logical analysis framework specified. The steps that can be taken constitute an over-all "plan of attack"; briefly, these are:

1. Address the root causes of congestion.
2. Update and, if necessary, improve the signalization.
3. Provide more space by use of turn bays and parking restrictions.

TABLE 6

OBJECTIVES TO BE SOUGHT DEPEND ON TRAFFIC CONDITION

TRAFFIC CONDITION	Congestion →	Saturation/Oversaturati
OBJECTIVE TO BE DESIRED	{Minimize Delay, Minimize Stops} →	{Avoid Spillback, Provide Equitable Service}
IMPACT OF USING OR TRYING TO ACHIEVE ABOVE OBJECTIVE (i. e., RESULTS)	⇓ • Intersection-specific congestion, not area-wide • Throughput Adequate • Service Regularity Enhanced • Enhancements to Special Groups (such as right turners) Possible	⇓ • Reduce Spatial Extent and Rate of Spread of Congestion • Potential for occurrence of area-wide Congestion reduced, and service regularity thus enhanced • Throughput not Impeded • Special Treatment of User Groups Possible
MAJOR INDEX OF PERFORMANCE	Delay, Stops	Queue Extent and Occupancy

4. Consider both prohibitions and enforcement realistically—Is an effort futile? Will it only transfer the problem?

5. Take other available steps, such as right-turn-on-red, recognizing that the benefits will generally not be as significant as either signalization or more space.

6. Develop site-specific evaluations where there are conflicting goals, such as providing local parking versus moving traffic, when the decision is ambiguous.

The following provides an exposition of these key elements. The framework by which a problem should be considered has two components: (1) a focus on the identification of the problem in terms of probable cause and (2) a focus on the categorization of the possible solutions, as previously discussed, so that they may be readily found within the guidelines. The recommended plan of attack summarized in the following discussion and the candidate solutions defined earlier are covered in detail in Appendix J.

Address Root Causes

First, foremost, and continually, it is vital that the engineer attack the problem of congestion and oversaturation by preparing a preliminary opinion on the underlying cause.

Improve Signalization

Contacts with traffic engineers have revealed that a systematic consideration of signalization for congestion and saturation is rarely done. Yet, poor signalization is often the basic problem. Once the signalization is improved through reasonably short cycle length, proper offsets (including queue clearance), and proper splits, many problems disappear. Sometimes, of course, there is just too much traffic. At such times, if other options cannot be called on, equity offsets to aid cross flows and different splits to manage the spread of congestion are appropriate.

A study of representative traffic patterns lends strong credibility to the result that minimal response (i.e., preplanned) signal policies generally suffice in combatting problems of congestion and oversaturation.

Provide More Space

If a problem cannot be remedied by signalization, the next major set of actions is summarized in two words: more space. Left-turn bays and, where appropriate, right-turn bays can aid individual movements as well as remove impediments to the through flows. Without question, additional lanes are a benefit. However, this tends to be an arterial-long solution.

Two-way left-turn lanes offer special advantages, particularly along strip development sites.

One-way systems, arterials with unbalanced lanes, and reversible lanes offer advantages, but also represent either major implementation problems or site-specific treatments. One-way systems require studies quite beyond congestion, although that may be the prime motivator for such a study. Unbalanced lanes require certain volume patterns.

Consider Prohibition and Enforcement

Before instituting any prohibition or enforcement program, the traffic engineer must decide:

1. Can it be enforced strictly enough to realize most or all of the projected benefit? (Curb parking prohibition to provide a moving lane is an example.)

TABLE 7

RANGE OF SOLUTIONS AVAILABLE

CONTROL MEASURE	TYPE	WHERE APPLIED	CHANGE POSSIBLE
SIGNAL	MINIMAL RESPONSE	INTERSECTION	CYCLE LENGTH
			BLOCK LENGTH
			SPLIT
			EXTRA PHASES
		SYSTEM	PREPROCESSING TRAFFIC
			PHASE ARRANGEMENT TO AID PROGRESSION
			EQUITY OFFSETS
			SPLITS TO APPORTION STORAGE AVAILABLE
			DOWNSTREAM EFFECTS
	HIGHLY RESPONSIVE	INTERSECTION	EXISTING POLICIES
			A MAXIMUM-QUEUE POLICY
			EXISTING HARDWARE
		SYSTEM	SPREAD FROM INTERSECTION
NON-SIGNAL	REGULATORY	ENFORCEMENT	AREA-WIDE
			DOUBLE-PARKING
			OTHER
		PROHIBITIONS	PARKING
			TURNING
			OTHER
	OPERATIONS	TURNING	LEFT BAYS
			RIGHT BAYS
			DUAL TURN LANES
			TWO-WAY LEFT TURN LANES
			RTOR
		LANE ARRANGEMENT (MAJOR)	ONE-WAY STREETS
			UNBALANCED
			REVERSIBLE
		DISRUPTION	PEDESTRIAM IMPACT
			BUS STOPS
			PARKING/UNPARKING
			DOUBLE-PARKING
			MID-BLOCK

2. Will it simply transfer or even accentuate the over-all problem? (The impact on other streets from circulation of vehicles that would otherwise be double parking is a prime example.)

Only then can he consider that there is a potential benefit.

Some solutions that are available can have either a net benefit or a net disbenefit, depending on the site and the situation. Right-turn-on-red (RTOR) is such a case. If it allows vehicles to "escape" from a congested arterial, it is quite suitable. If, however, it allows vehicles to "steal" available space on such an arterial, it is inappropriate.

The question of such prohibitions as turning prohibitions arises. Those can only be used if alternate routes exist.

Consider Solutions In Terms Of Economics

Very often, application of these guidelines will clarify the issue and identify a solution. In some cases, the final decision will rest on conflicting desires that might be usefully viewed in economic terms. Is the removal of 5 parking spaces worth the delay savings to the traffic stream? Are off-street loading zones justified economically? Are pedestrian phases justified in terms of total person-minutes saved? What is a proper allocation of curb space?

If it is necessary, the engineer can develop such an analysis for his individual case. More general treatment of these situations is recommended for future research. Some work on curb space management for goods facilities has been done in other research efforts (*9*).

TABLE 8

AN INITIAL CLASSIFICATION AND ELIMINATION CHECKLIST

Apparent Problem	Initial Steps
Area-Wide Congestion	• Identify the Critical Intersection (CI). • It is unlikely that there are two CI's. If there are, it is likely that each can be considered with its own area of influence. • Do not erroneously identify the intersection downstream of the CI as the CI. • Classify the type of oversaturation (Types I through V, as defined)
Spillback in a Link	• Determine whether a simple split adjustment is sufficient. • Determine whether the cycle length is too long for the block length and flow. • Determine whether the offset is poor for the primary and secondary flow mix. • Identify any special blockages in the link (double parking, queues for garages, car washes, etc.).
Single Intersection	• Isolate primary symptom.
→ Single Approach	• Check split and offset as above. • Check burden caused by turns.
→ Two Approaches, same Right-of-Way	• Check same as one approach, but with emphasis on interference with each other.
→ Two or More Approaches More than One Right-of-Way	• Consider methods for increasing net capacity. • Consider methods for minimizing spatial extent of possible oversaturation and area-wide congestion.

CHAPTER THREE

APPLICATIONS

THE GUIDELINES

The research conducted as part of this project has culminated in the development of the guidelines contained in Appendix J. These guidelines are recommended for use as a tutorial on congestion/saturation and as a working reference on how to approach specific applications. Because of the extreme variety of actual cases, these guidelines emphasize systematic methodology and representative illustrations rather than comprehensive, unequivocal, all-encompassing decisions and hierarchies of treatment. The latter approach was deemed both infeasible and impractical.

The guidelines are recommended as a tool for traffic engineers in all localities suffering traffic congestion and oversaturation.

DEFINITIONS

The definitions proposed in this report for congestion, saturation, and oversaturation are recommended for use by researchers and practicitioners. In this way, the future literature can be cross referenced with some confidence and without ambiguity.

REGULARITY OF TRAFFIC PATTERNS

The results stemming from the work performed in connection with the regularity of traffic patterns are available for comparative studies as recommended in Chapter Four. The regularity of the daily traffic pattern in (at least) the one city considered (Toronto) is impressive, and must be considered in evaluations of the need for control policies that are highly responsive to traffic variations.

QUEUE-ACTUATED TACTICAL CONTROL POLICY

This "working paper" on queue-actuated tactical control policy is contained in Appendix F and is recommended for consideration by those involved in highly responsive control policies.

UTCS-1 SIMULATOR

The UTCS-1 simulator has been used extensively and found quite suitable. Some validation of it has been done. As one of the early users of the UTCS-1 (during the first stages of the project particularly), difficulties with it and limitations of the output structure were discussed with the developers for consideration in later revisions.

Some support programs were written and are included in Appendix I.

CHAPTER FOUR

CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

CONCLUSIONS

The conclusions of this research can be summarized as follows.

The problem of congestion and saturation is widespread, are a broad variety of techniques is being applied in an effort to combat the problem. However, there is no specific direction of attack. On the basis of the project results, the researchers conclude that definite measures can be taken, but preventive action addressing the root causes must be given a high priority. Among the measures that can be taken, those relating to signalization generally can have the greatest impact. There are distinct signal plans for avoiding spillback and for "living with" spillback. The nonsignal remedies are in no way to be minimized, particularly those that provide space either for direct productivity increases or for removing impedances to the principal flow. The guidelines developed in this work provide both a tutorial and an illustrated reference in what techniques to consider and how to consider them systematically.

The following addresses specific recommendations.

RECOMMENDATIONS

The Guidelines

The guidelines in Appendix J should be published in manual form for the benefit of traffic engineers confronted with the problems of congestion and saturation. An effort should be made at some future time also to ascertain whether it is used as a one-time tutorial or a continual reference and, if the latter, what sections have proven most useful.

A Cost-Effectiveness Methodology

The trade-off between alternative treatments is sometimes not so clear that a simple decision can be made. Moreover, some treatments involve primarily economic impacts not necessarily related to the moving traffic stream; for example:

- The impact of removing N parking spaces in order to provide a right-turn bay that will save some delay to the moving stream, the delay saved being a function of volume.
- The timing of an intersection to minimize total person-minutes of delay, including pedestrians.
- The proper allocation of curb space.
- The provision of space to a transit lane.

Such evaluations have been carried out, particularly in regard to curb space management and goods movement (*9*). Further studies are recommended that emphasize the implications of such evaluations in terms of their cost effectiveness. From these studies would emerge not only a methodology, by which traffic operations decisions could be traded off one against the other, but also an insight into what the involved numbers and costs are. How much is an on-street parking location "worth" to the community? How much does it now cost in alternative delay? What is the cost to transit operations of prohibiting a turn? Of making a street one-way?

It is judged that much benefit could be realized just by encouraging the engineer to quantify such decisions. Among other things, subtle costs could be identified early—increased costs to the transit operator, for instance. More than that, such an approach could serve as a rational basis for both questioning and motivating improvements that affect the most visible component—the moving traffic stream.

Regularity of Traffic Patterns

The research results of this study, using the data provided by the metropolitan Toronto Department of Roads and Traffic, indicate an extreme regularity of daily traffic patterns and of the time at which certain flow levels are first attained. These results have implications that make it necessary to evaluate the cost effectiveness of highly responsive signal control, as discussed in Appendix E.

In addition, it should be noted that virtually nothing appears in the literature from which it may be ascertained as to whether these results are typical. Therefore, those with data concerning other localities should make this information available in publications on this subject, or data should be acquired for such purposes.

REFERENCES

1. MOSKOWITZ, K., "Long Cycles, Short Cycles, and "Lost Time" at Signalized Intersections." Paper presented at the Workshop on Intersection Capacity, Highway Research Board (Jan. 1974).
2. WEBB, G. M., and MOSKOWITZ, K., "Intersection Capacity." *Traffic Engineering* (Jan. 1956) pp. 147-152.
3. CARSTENS, R. L., "Some Traffic Parameters at Signalized Intersections." *Traffic Engineering* (Aug. 1971) pp. 33-36.
4. GREENSHIELDS, B. D., ET AL., "Traffic Performance at Urban Street Intersections." *Technical Report No. 1*, Yale Bureau of Highway Traffic, New Haven, Conn. (1947).
5. PARSONSON, P. S., "Small Area Detection at Intersection Approaches." *Traffic Engineering* (Feb. 1974) pp. 8-17.
6. YAGODA, H. N., "The Control of Arterial Street Traffic." Paper presented at the 1967 TFAC Control Conference, Haifa, Israel.
7. GAZIS, D. C., "Optimum Control of a System of Oversaturated Intersections." *Operations Research*, V. 12, No. 6, pp. 815-831 (Nov-Dec. 1964).
8. LONGLEY, D., "A Control Strategy for a Congested Computer-Controlled Traffic Network." *Transportation Research*, Vol. 2, pp. 391-408 (1968).
9. HABIB, P. A., "Accommodating Goods-Movement Vehicles in the City Center." Ph.D. dissertation, Polytechnic Institute of New York (June 1975).

APPENDIX A

NATIONWIDE SURVEY ON EXTENT OF CONGESTION/SATURATION

THE QUESTIONNAIRE

A copy of the questionnaire as distributed in November of 1971 is included as Figure A-16.

Congestion was defined, for purposes of the questionnaire, as a condition, at an intersection, in which all waiting vehicles cannot pass through the intersection on one signal cycle.

Saturation was defined as an extension of congestion, as that condition when vehicles leaving an intersection are prevented from moving freely because of the presence of vehicles either in the intersection itself or backed up in any of the exit legs of the intersection.

DISTRIBUTION OF RESPONSE

A total of 569 questionnaires was sent to cities, states, and counties in the United States and Canada. Responses totaled 205, or 36.4 percent.

Questionnaires were sent to all cities of over 50,000 population, and to those cities with less than 50,000 population which had a city traffic engineer. Counties and states were included with the view of learning whether such entities encountered this typically urban problem, but most responses indicated a lack of direct involvement with the congestion-saturation syndrome.

Distribution of the city sizes queried is shown in Figure A-1. Generally, responses were from 25 to 82 percent, with the over-all response at about 36 percent. It can be seen from Table A-1 that the greatest percentage return (81.5 percent) was from the largest cities (over 500,000 population).

CHARACTERIZATION OF RESULTS

Responses were tabulated by groupings—city size, county, state, and Canada. Totals and percentages were taken to give a representation by such groupings and some indication of the scope of problems described.

Table A-2 and Figures A-2 through A-5 summarize the results. These summaries include all viable portions of the available responses. Any comments or interpretation judged appropriate is made in the text and in the guidelines (Appen. J).

A review of some of the locality maps provided with the questionnaire responses shows the following:

1. For Kansas City, Mo. (pop. 459,405) saturation is indicated for the CBD core, plus intersections along major routes into the CBD; congestion is indicated with queues both less than, and greater than, halfway to the upstream intersection on intersections surrounding the CBD and approaches to intersections on major routes.
2. For Hamilton, Ontario (pop. 298,121), saturation is shown at intersections along major routes into the CBD.
3. For Cincinnati, Ohio (pop. 448,444), a situation similar to that of Kansas City exists.

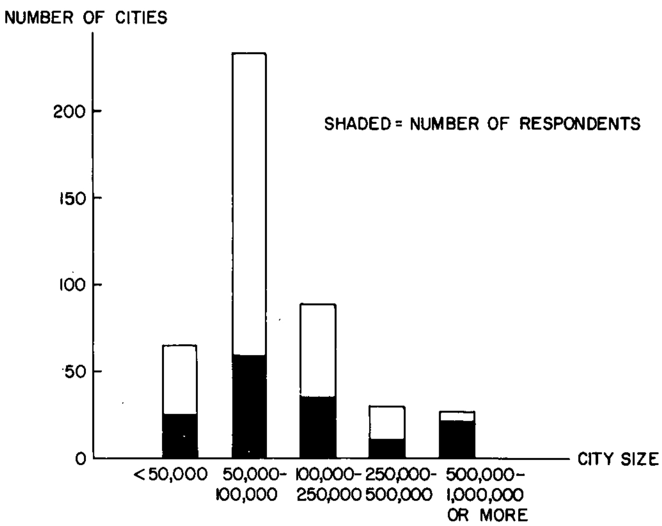


Figure A-1. Distribution of city sizes queried.

TABLE A-1
QUESTIONNAIRE RETURNS

	CITIES (Population)									
	50,000	50 to 100,000	100 to 250,000	250 to 500,000	500 to 1,000,000 +	Total Cities	Counties	States	Canada	Total
Sample Size	66	233	88	31	27	445	56	50	18	569
Returns	26	59	34	11	22	152	19	25	11	207
% Return	39.4	25.3	38.6	35.5	81.5	34.2	33.9	50.0	61.1	36.4

4. For Wichita, Ka. (pop. 273,507), saturation is indicated along First Street (an arterial) and also at two major intersections on US 54; congestion is indicated with queues over half the upstream block long on other feeder routes to the CBD.

5. For St. Paul, Minn. (pop. 308,606), saturation occurs at 5 CBD intersections and 3 main route intersections; congestion occurs at intersections peripheral to the CBD.

PATTERNS OF CONGESTION AND SATURATION

The maps provided by many respondents were studied for patterns of occurrence of congested and saturated intersections. General impressions were the following:

1. *Cities of Less than 50,000 Population*—Saturation occurs typically at a few downtown intersections, not always contiguous to one another. It may also occur at isolated intersections or nodal points of major routes into the city. Congestion occurs in relation to saturated points in the CBD, and in relation to saturated points on major approach routes.

2. *Cities Between 50,000 and 100,000 Population*—A pattern similar to the foregoing appears to be common, with greater repetition of chains of contiguous intersections along major routes subject to saturation or congestion at various levels. Saturation/congestion at approaches to tunnels or bridges appears frequently, as it does in the vicinity of freeway interchanges.

3. *Cities Between 100,000 and 250,000 Population*—Depending on the conformation of the city in this size grouping, saturation/congestion will occur in concentrated form in a downtown street pattern, or will appear at isolated intersections along major routes approaching downstream. Where contiguity occurs, 3, 4, 5, or 6 contiguous intersections may be affected. Major routes with no alternatives are particularly prone to saturation/congestion, especially as the CBD is approached. Generally, however, proportionately fewer strictly CBD intersections are affected in this size grouping, with the phenomenon affecting intersections beyond the limits of the CBD.

4. *Cities Between 250,000 and 500,000 Population*—Saturation again afflicts major portions of the CBD traffic system or substantial portions of feeder arterials in this size group.

5. *Cities of Greater than 500,000 Population*—The largest cities show little CBD saturation, but much saturation occurs on major routes into the city, with CBD intersections relatively less affected by varying degrees of congestion.

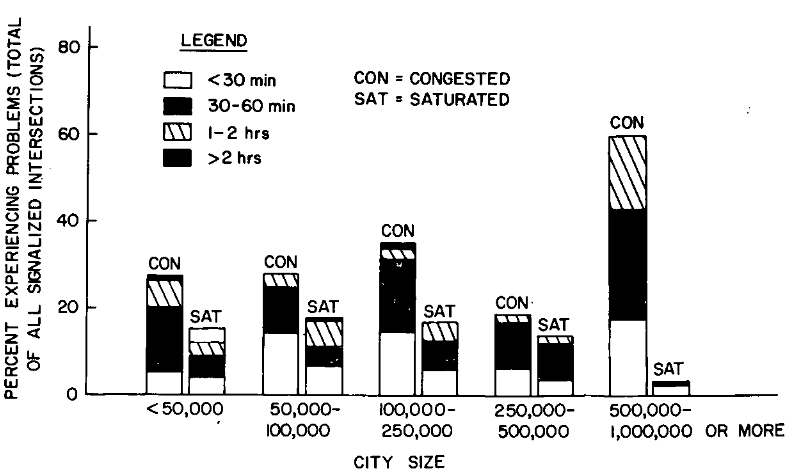


Figure A-2. Extent and duration of traffic problems, by city size.

TABLE A-2

NUMBER OF SIGNALS IN JURISDICTION (AVERAGE)

	in CBD		on Arterials		in Other Locations	
City Size	No.	%	No.	%	No.	%
50,000 <	13	35	21	58	7	7
50-100,000	22	27	41	58	23	15
100-250,000	51	28	109	59	32	13
250-500,000	94	24	247	63	49	13
500,000 +	156	12	426	63	429	25
County	55	22	77	67	18	11
State	380	26	82	42	735	32
Canada	53	23	160	70	43	7

ASSESSING CONGESTION AND TIMING SIGNALS

Figure A-6 shows the methods used by various size cities to assess congestion; Table A-3 summarizes the methods used to determine signal timing plans. Note the extreme emphasis on "hand" solution.

PHYSICAL AND OPERATIONAL IMPROVEMENTS RECOMMENDED

The following items were cited in the various responses as improvements to be used or considered:

PHYSICAL IMPROVEMENTS	OPERATIONAL IMPROVEMENTS
Lane widening	Signal modernization
Channelization	• optics
Grade separations	• multidial
Sight distance	• traffic actuated

Skid resistance of pavements
Location of signals
Bus lanes (exclusive)
Truck routes
Interconnection of CBD signals (minimum)
Progression timing for signals
Capacity analyses
Parking prohibitions
Turn prohibitions
Separate signal phases
Pedestrian constraints
Signing

The responses showed a heavy emphasis on applying the principles of the *Highway Capacity Manual* (1965), followed by an appropriate cost/benefit analysis to determine priority of improvement and choice of a particular type of improvement at a given location. Between the two broad categories mentioned, neither one predominated in any given city. Rather, a close relationship between major physical improvements and the application of modern operational controls predominated in the majority of recommendations under the TOPICS program.

CONTROLS UTILIZED

Figures A-7 through A-15 summarize responses to inquiries on the use of control techniques to handle congestion/ saturation.

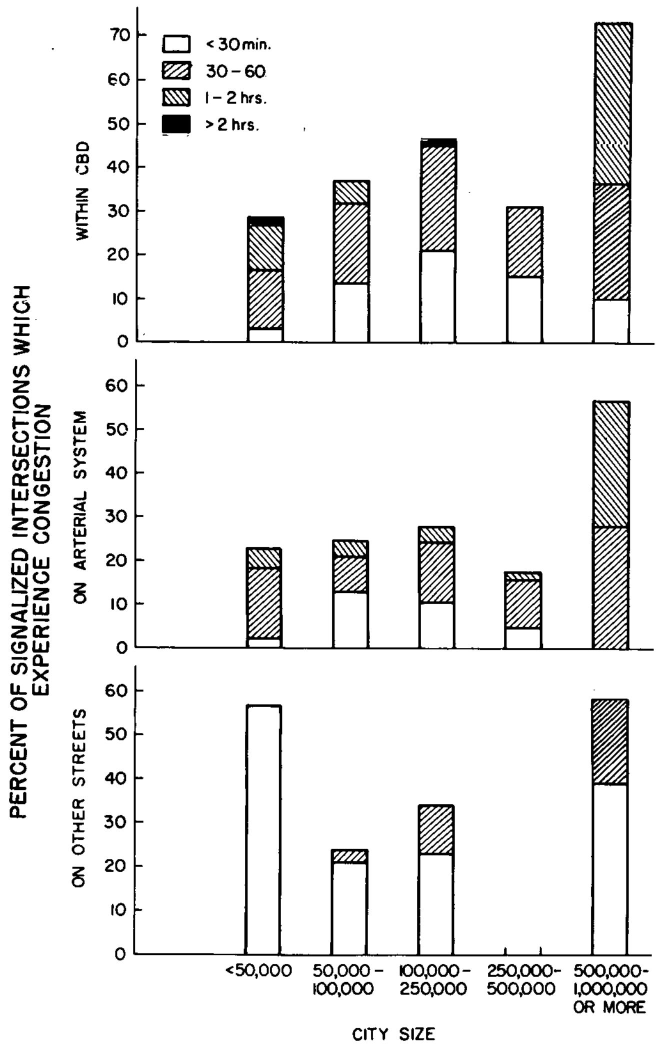


Figure A-3. Extent and duration of traffic congestion as function of locale and city size.

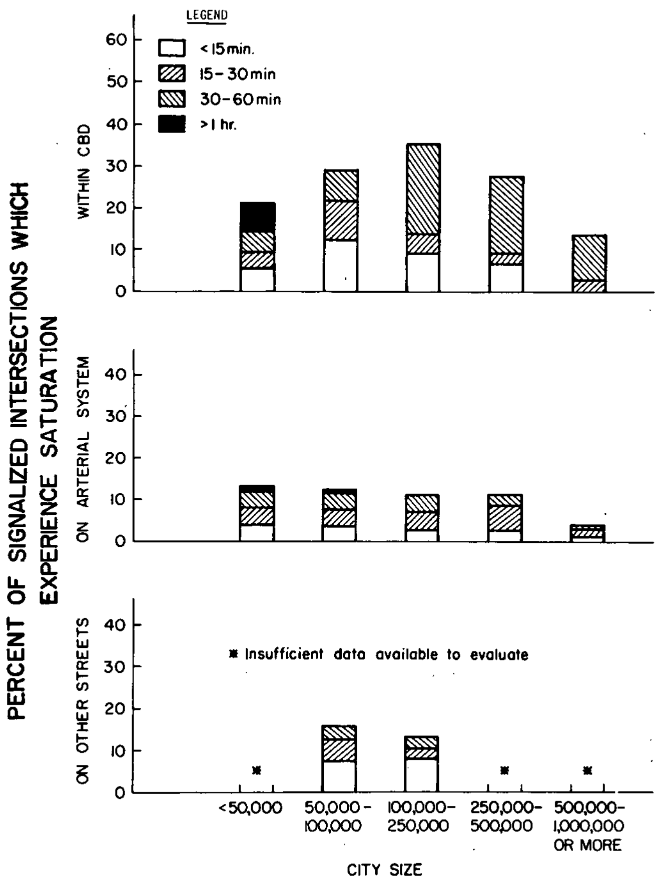


Figure A-4. Extent and duration of traffic saturation as function of locale and city size.

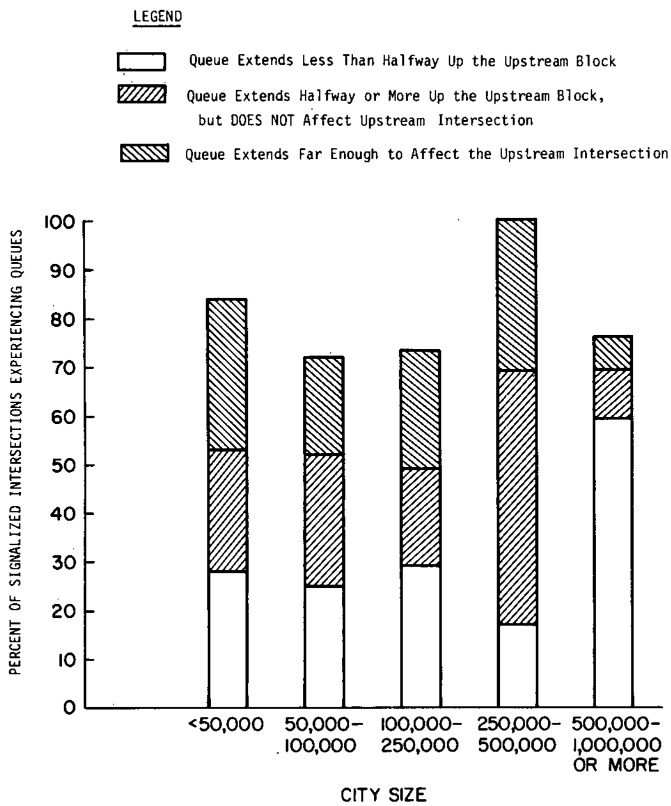


Figure A-5. Queue length at congested intersections as a percent of all signalized intersections experiencing congestion, by city size.

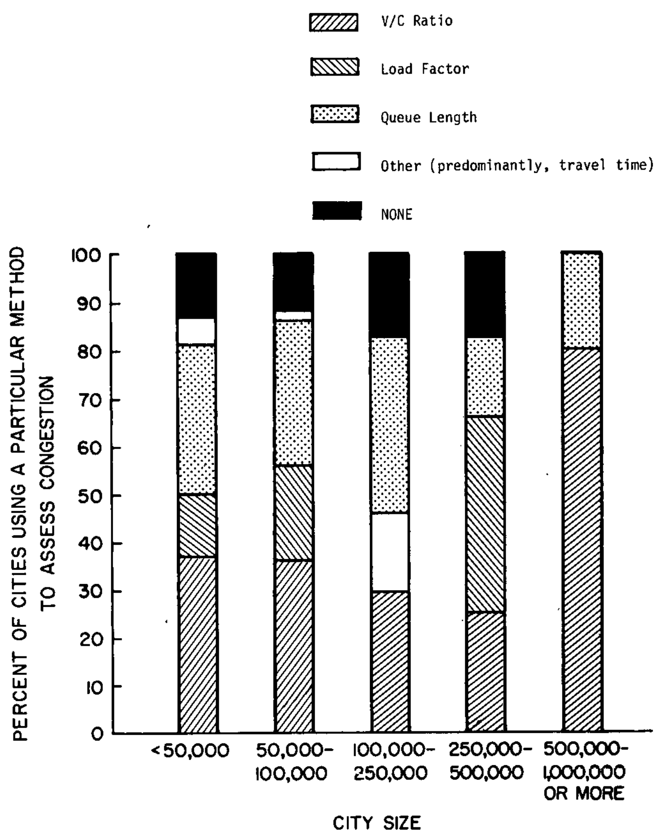


Figure A-6. Methods used to assess congestion in various city sizes.

TABLE A-3

REPORTED METHODS OF TIMING SIGNALS

CITY SIZE (Population)	METHOD OF PREPARING SIGNAL TIMING (NUMBER OF RESPONSES): Computer Calculation of Time/Space Diagram	SIGOP*	SIGRID**	HAND***	Average Number of Signals in Jurisdiction
50,000 <	1	-	-	21	35
50,000-100,000	4	2	-	39	70
100,000-250,000	2	1	1	25	183
250,000-500,000	-	3	1	5	391
500,000 +	1	1	-	5	1,377
All	8	7	2	95	

* Traffic Signal Optimization Program (a network optimizer)

** A Predecessor of SIGOP - Developed for Toronto

*** Manual computation of Signal Timing, usually by Development of a Time/Space Diagram

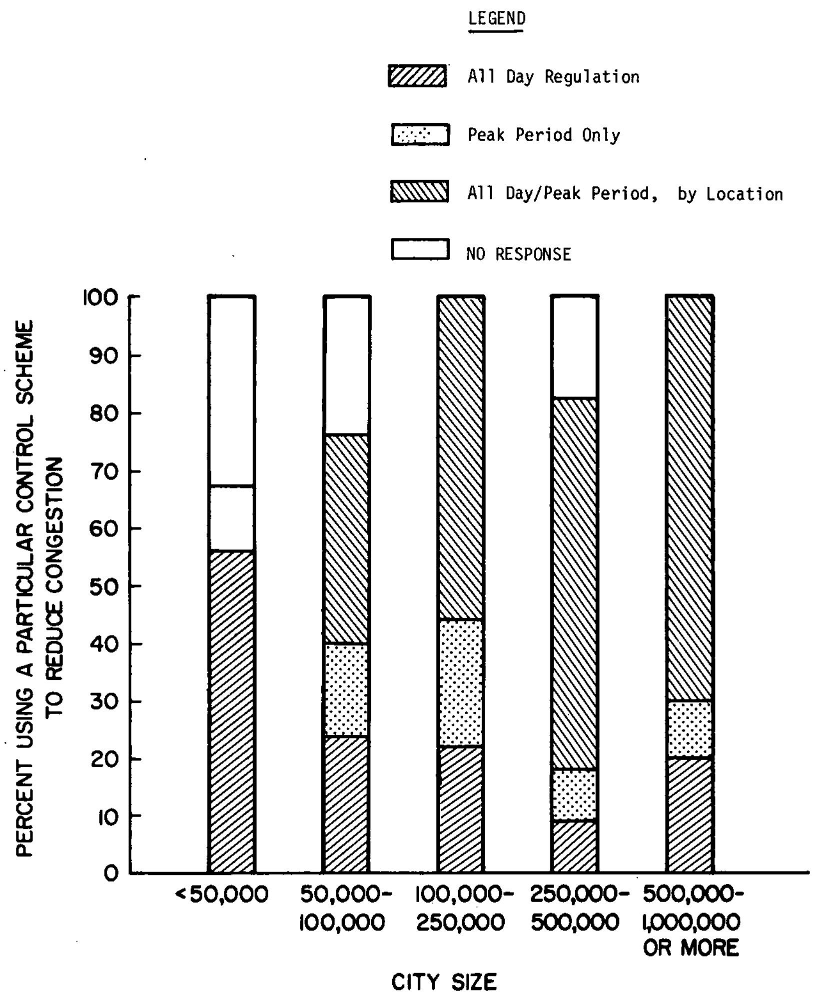


Figure A-7. Parking regulations used for handling congestion and/or saturation, by city size.

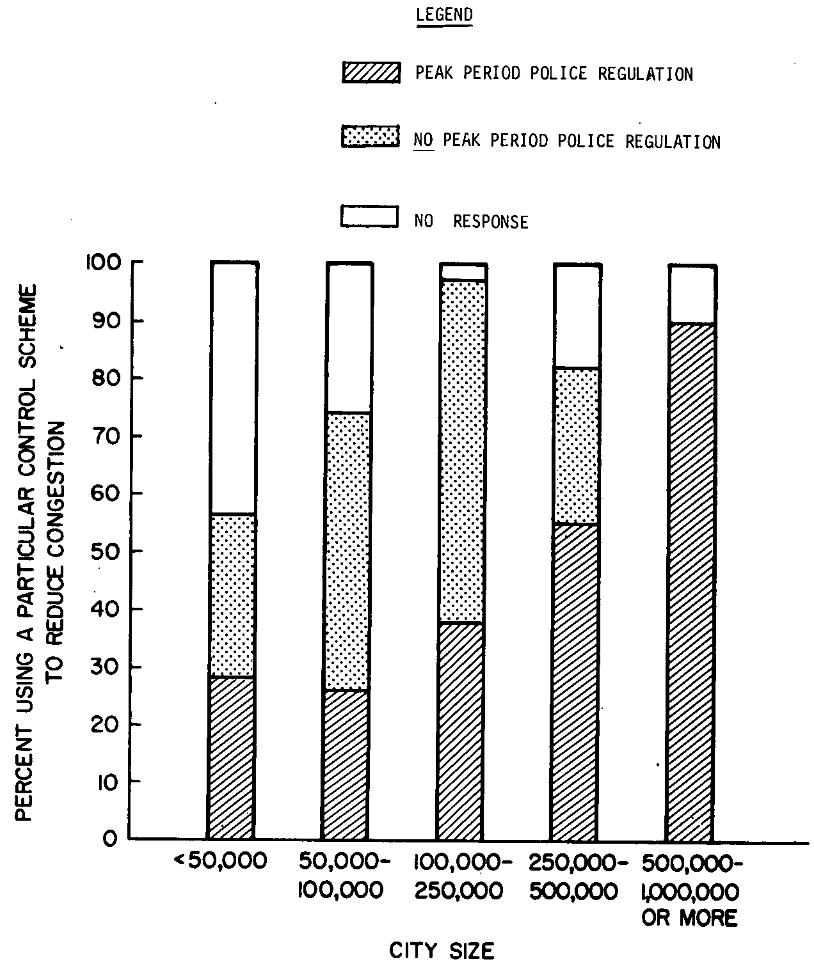


Figure A-8. Use of police regulation for handling congestion and/or saturation, by city size.

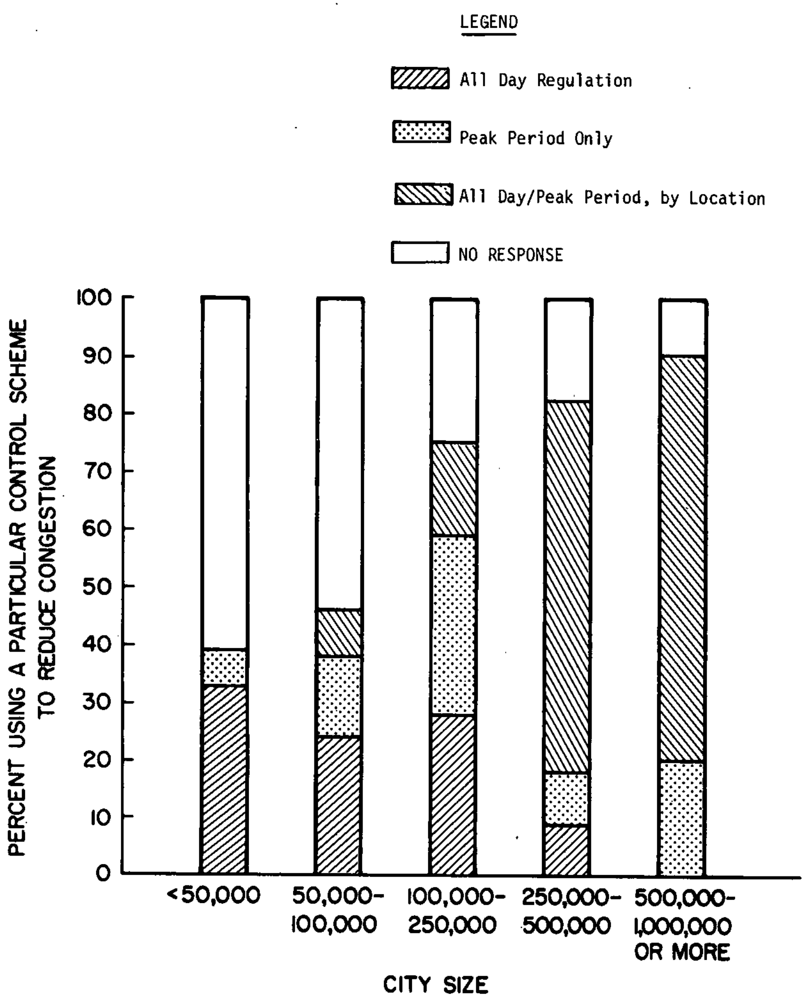


Figure A-9. Left-turn regulations used for handling congestion and/or saturation, by city size.

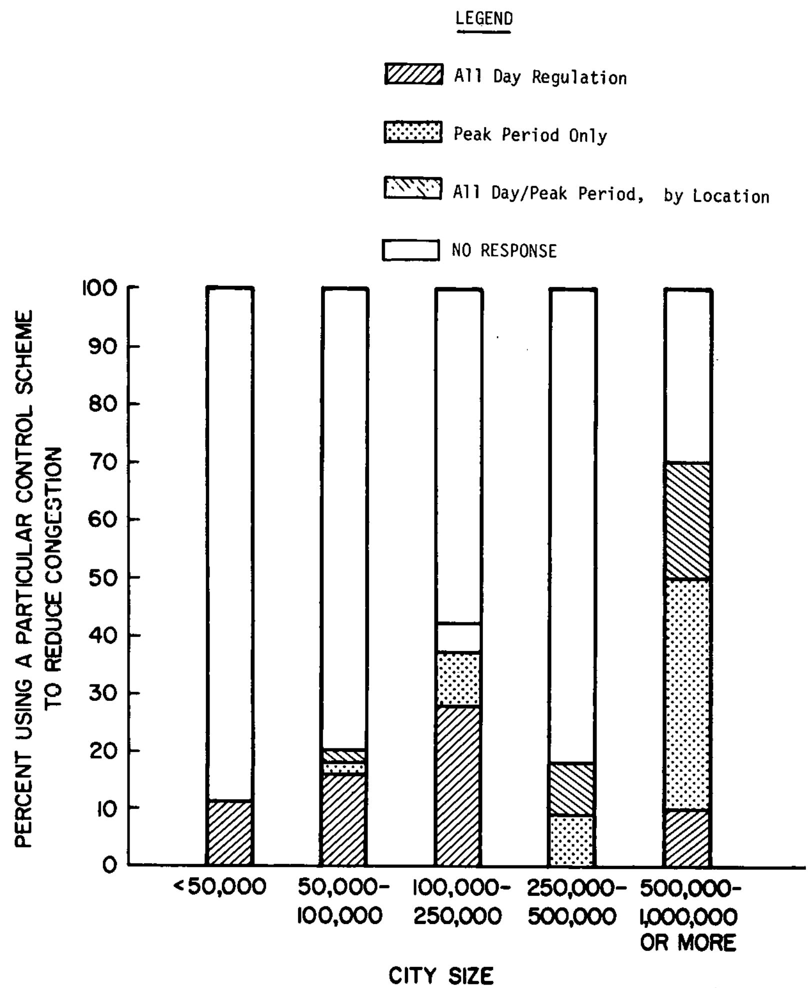


Figure A-10. Right-turn regulations used for handling congestion and/or saturation, by city size.

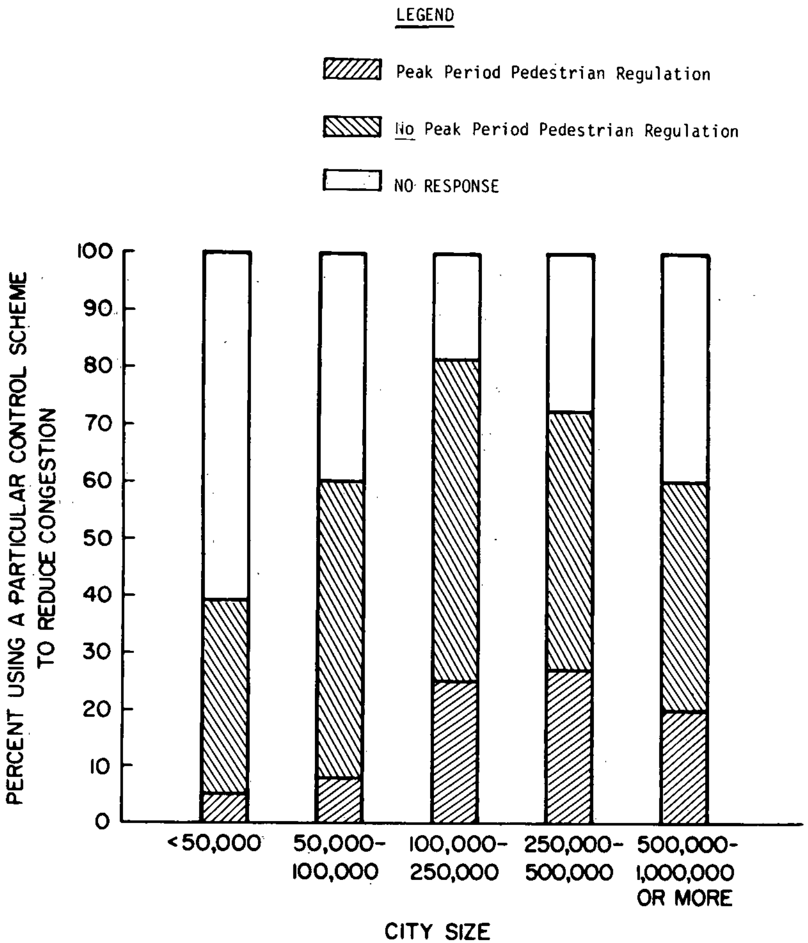


Figure A-11. Use of pedestrian regulation for handling congestion and/or saturation, by city size.

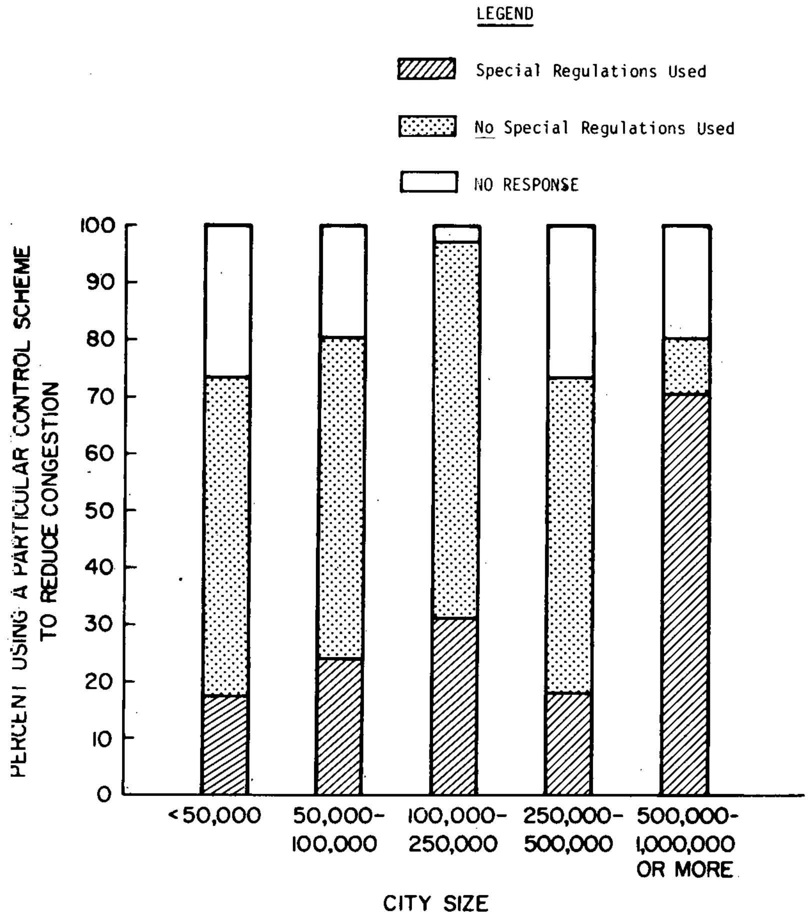


Figure A-12. Use of special commercial vehicle regulations for handling congestion and/or saturation, by city size.

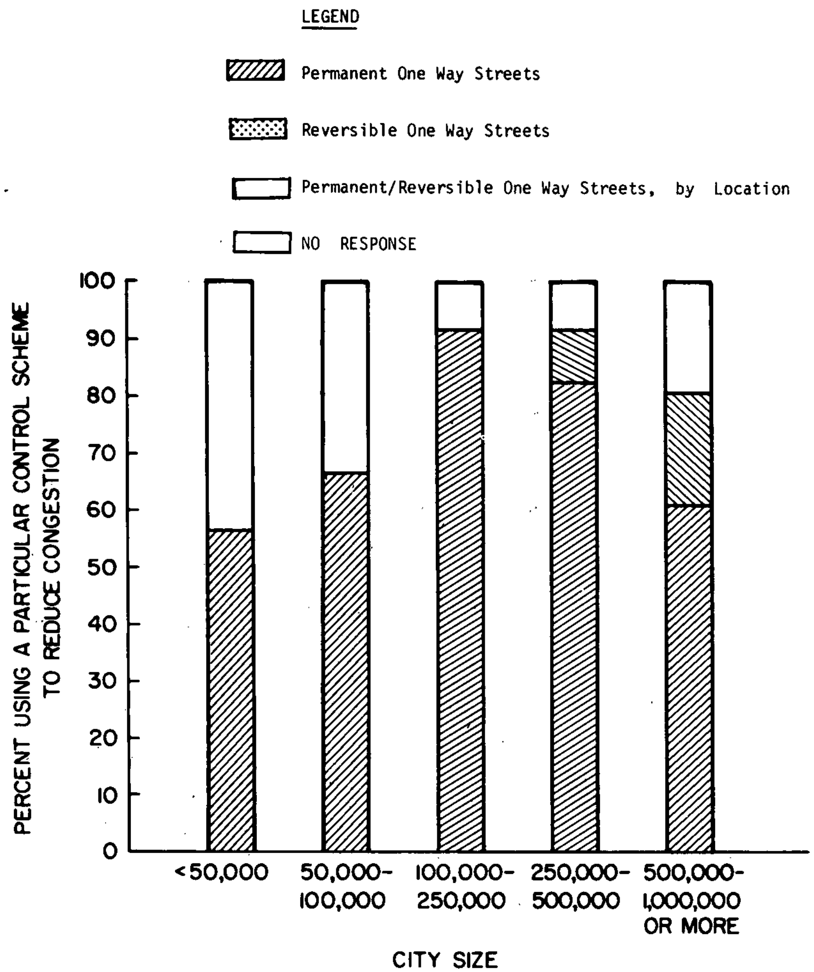


Figure A-13. Use of one-way street for handling congestion and/or saturation, by city size.

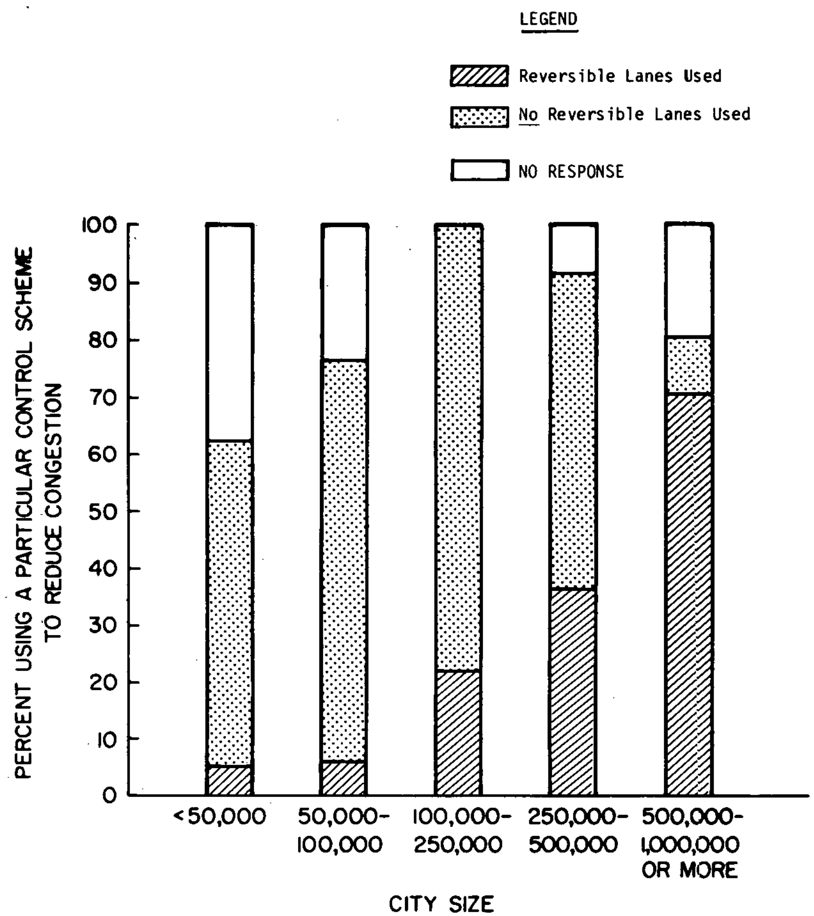


Figure A-14. Use of reversible lanes for handling congestion and/or saturation, by city size.

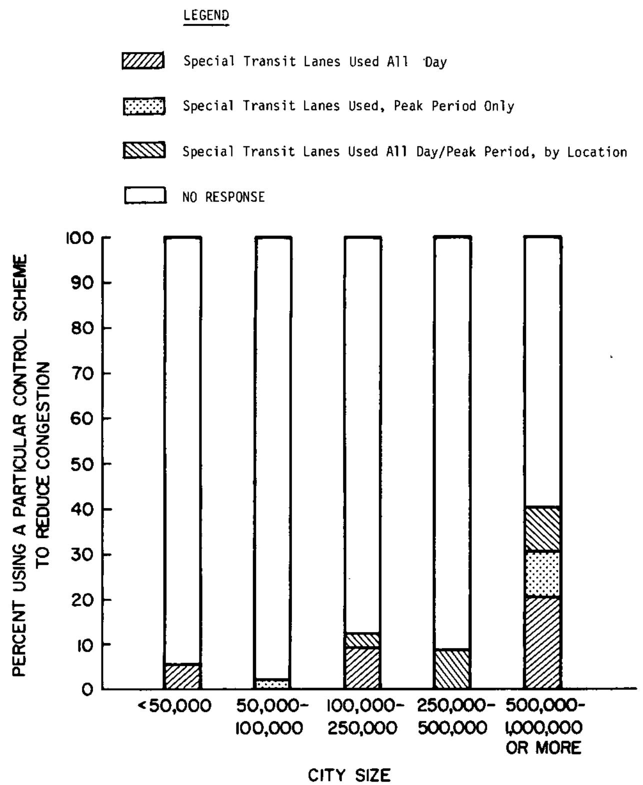


Figure A-15. Use of special transit lanes for handling congestion and/or saturation, by city size.

POLYTECHNIC INSTITUTE OF BROOKLYN
333 Jay Street, Brooklyn, New York 11201

Department of
TRANSPORTATION
Planning & Engineering

24 November 1971

NCHRP Project 3-18(2)

TRAFFIC CONTROL IN OVERSATURATED STREET NETWORKS

As mentioned in my previous communication, we are investigating problems of oversaturated street networks and problems of traffic congestion generally, in selected cities of the United States and Canada. The major objectives of the project are:

- To develop a better understanding of the extent and seriousness of the problem;
- To understand how current traffic operations techniques are being applied to combat the problem; and
- To investigate advanced concepts as to their potential utility.

A questionnaire is enclosed, and I would very much appreciate it if you would give it your consideration and complete and return it by 1 January 1972. The knowledge gained through your cooperation will be of service to all, and will add to the meaningful success of this valuable project.

Kenneth W. Crowley
Project Manager

Figure A-16. Questionnaire form used in survey.

The Department of Transportation Planning and Engineering
Polytechnic Institute of Brooklyn
333 Jay Street, Brooklyn, N.Y. 11201

under contract to

National Cooperative Highway Research Program
Highway Research Board
National Academy of Science - National Research Council

NCHRP Project 3-18(2)

QUESTIONNAIRE

on

Problems and Practices of Traffic Practitioners in the Area of

Congestion and Saturation of Street Networks

PLEASE CORRECT OR COMPLETE YOUR ADDRESS LABEL.

[Please return this cover sheet with your completed questionnaire.]

PLEASE RETURN TO: Kenneth W. Crowley
Department of Transportation Planning & Engineering
Polytechnic Institute of Brooklyn
333 Jay Street
Brooklyn, N.Y. 11201

PART A - Please Complete

1. How many signals do you have under your jurisdiction? ________

2. Of these, about how many are
   a) in the CBD network? ____________
   b) outside the CBD on the arterial system? ________
   c) in other locations? ___________

3. By what means do you asses congestion at an intersection?

☐ V/C ratio ☐ load factor ☐ queue length ☐ other (specify below) ☐ none

____

If we define CONGESTION at an intersection as that condition when all waiting vehicles cannot pass through the intersection on one signal cycle,

4. Do you regularly experience peak period congestion at locations in
   a) the CBD network? ☐ yes ☐ no If yes, about how many? ______
   b) the arterial system? ☐ yes ☐ no If yes, about how many? _____
   c) other locations? ☐ yes ☐ no If yes, about how many? _____ What locations are urban (non-CBD) ☐
   suburban ☐

5. How long would you estimate typical peak-period congestion lasts in
   a) the CBD network?
   ☐ less than 30 minutes ☐ 30 min. to 1 hour ☐ 1 hour to 2 hrs. ☐ longer than 2 hours
   b) the arterial system?
   ☐ less than 30 minutes ☐ 30 min. to 1 hour ☐ 1 hour to 2 hrs. ☐ longer than 2 hours
   c) other locations?
   ☐ less than 30 minutes ☐ 30 min. to 1 hour ☐ 1 hour to 2 hrs. ☐ longer than 2 hours

If we defined SATURATION as that condition when vehicles leaving an intersection are prevented from moving freely because of the presence of vehicles either in the intersection itself or backed up in any of the exit legs of the intersection,

6. Do you regularly experience peak period saturation at locations in

a) the CBD network? ☐ yes ☐ no If yes, how many? ______

b) the arterial system? ☐ yes ☐ no If yes, how many? ______

c) other locations? ☐ yes ☐ no If yes, how many? ______ What locations are they? urban (non-CBD)
suburban

7. How long would you estimate typical peak-period saturation lasts in

a) the CBD network?

☐ less than 15 minutes ☐ 15 min. to $\frac{1}{2}$ hour ☐ $\frac{1}{2}$ hour to 1 hr. ☐ longer than 1 hour

b) the arterial system?

☐ less than 15 minutes ☐ 15 min. to $\frac{1}{2}$ hour ☐ $\frac{1}{2}$ hour to 1 hr. ☐ longer than 1 hour

c) other locations?

☐ less than 15 minutes ☐ 15 min. to $\frac{1}{2}$ hour ☐ $\frac{1}{2}$ hour to 1 hr. ☐ longer than 1 hour

8. Do you consider saturation to be a significant problem in your jurisdiction? ☐ yes ☐ no

To give us a clearer visual perspective of the scope of this problem, please provide us with a map of the area under your jurisdiction.

9. Approximately what percentage of your intersections which have congestion have typical maximum queues which extend less than halfway up the upstream block? ______

Mark them on your map in BLUE.

10. Approximately what percentage have typical maximum queues which extend halfway or more up the upstream block (without affecting the next intersection)? ______

Mark them on your map in GREEN.

11. Approximately what percentage have typical maximum queues which affect the next (upstream) intersection? ______

Mark them in RED.

12. Do you have any special control procedures for handling congestion and/or saturation? ☐ yes ☐ no

If yes, do they involve

Regulations

Parking: ☐ All day ☐ Peak Periods Only

Turning

Left Turn: ☐ All day ☐ Peak Periods Only

Right Turn: ☐ All day ☐ Peak Periods Only

Peak Period Police Control: ☐ yes ☐ no

Pedestrian: ☐ yes ☐ no Please specify:

______

______

One Way Streets ☐ Permanent ☐ Reversible

Reversible Lanes ☐ yes ☐ no

Special Transit Lanes ☐ All Day ☐ Peak Periods Only

Special Commercial Vehicle Regulations ☐ yes ☐ no Please specify:

Other Regulations or Controls Please describe:

Please do not hesitate to further describe these methods, or to attach appropriate reports, if you feel it would be helpful.

PART B - We would appreciate this too.

1. Would you describe the special control scheme(s) you use to alleviate congestion?

2. By what method was the signal timing prepared? (by hand, SIGOP, etc.)

3. To what extent are commercial vehicles a cause of the problem of congestion?

4. To what extent are pedestrians a cause of the problem of congestion?

5. To what extent are bus operations a cause of the problem of congestion?

6. To what extent are land use policies a cause of the problem of congestion?

APPENDIX B

DEFINING MEASURES OF OVERSATURATION

This appendix concerns itself with the definition of measures of effectiveness that are appropriate for use in the characterization or control of oversaturated street systems.

Two major subareas of interest are presented. The first includes a detailed examination of candidate measures of effectiveness for use in the oversaturated street traffic problem. The second presents the results of a limited data collection and analysis effort designed to check the reasonableness of the selected measures of effectiveness.

TYPES OF MEASURES

A broad range of measures of effectiveness is in current use today to characterize some aspect of traffic performance. Measures include those that relate to stops, delays, travel time, and production. These several measures, in general, fall into one of two over-all categories:

1. Those measures that describe conditions at specific intersections or points within the system.
2. Those measures that describe over-all conditions in the system or in subsections of a larger system.

In addition, measures of effectiveness may be characterized by use. Here, too, there are two major categories:

1. Those measures that are useful in characterizing the state of the point or system under consideration.
2. Those measures that have utility in aiding the control of the point or system under consideration.

Some measures may have utility both as descriptors and for control purposes, just as some measures may allow for characterization of local points or systemwide effects. There may also be measures that have both point and system utility.

In this appendix a number of measures of effectiveness will be reviewed and evaluated with respect to their utility in the current problem. As such, consideration will be given to how well they allow for characterizations consistent with the segmentation of congestion into its three subregions of interest, as described in Chapter Two. Thus, it may be that a measure will be satisfactory for defining the boundary between uncongested and congested operations, but might not, for example, permit the description of onset of oversaturation.

The literature was reviewed to determine which, if any, measures of effectiveness were available for use in the oversaturation problem. The highlights of the literature review are discussed in the following text. Appendix C is an annotated bibliography of relevant publications.

TABLE B-1

EXISTING AND NEW PARAMETERS OF POSSIBLE UTILITY

TYPE OF MEASURE	UTILITY AS:	
	DESCRIPTOR ONLY	DESCRIPTORS AND/OR CONTROL PARAMETER
INTERSECTION	Load Factor Saturation Factor Maximum Individual Delay/Vehicle Number of Cycles to Clear Intersection Number of Stops/Starts to Clear Intersection	Queue Length Total Delay Average Delay Link Length Minus Queue Length Ratio of Queue Length to Link Length Input-Output Trapped Vehicles Volume/Capacity (V/C) Ratio
SYSTEM	Total Delay Average Number of Stops/Vehicles Density Mean Velocity Gradient Occupancy Total Travel Time	Density Occupancy Input-Output

LITERATURE REVIEW OF MEASURES OF EFFECTIVENESS

A number of possible measures of effectiveness were revealed in the literature review, or proposed by the research agency. These measures are given in Table B-1, stratified by range of utility (i.e., by point or intersection, or by system) and by use (i.e., description only or description and/or control parameters). As can be seen, there are significantly more potential measures of local or point interest than of system utility. This reflects not only the state of the art, but also the research agency's belief that the problem of oversaturation, regardless of how wide it eventually spreads, may be characterized at the start by point or local measures.

Each of the candidate measures (intersection and system) is described in the following. The measures are then evaluated to determine the one or ones most suitable for describing or characterizing saturation/oversaturation, as well as for utility as control parameters. It should be noted that the choice of a measure or set of measures cannot avoid a certain amount of subjectiveness.

DESCRIPTION OF POINT MEASURES OF SATURATION

Load factor is the percentage of fully loaded cycles in the peak hour. A load factor of 1.00 indicates at least saturation throughout the peak hour. Load factors less than unity indicate that congestion did not exist throughout the peak hour, although it may have existed during some shorter period during the hour. Load factor is an intersection measure, and is useful primarily as a descriptor of conditions. Load factor is essentially a technique of field measurement of intersection performance, and does not correlate well with other operating characteristics. Identification of loaded cycles requires some judgment on the part of the field observer. Simulation studies (*17, 18*—note that these references cite items in annotated bibliography of Appendix C) revealed that load factors related quite well to average individual delay at load factor values less than 0.6. Beyond this there was evidence that the relationship became unstable.

Saturation factor is the number of saturated cycles at a given approach to an intersection divided by the total number of cycles during a specified time period. Saturated is defined as a signal cycle for which, at a given approach to a signalized intersection, the number of vehicles stopped at the end of the red interval is greater than the number of vehicles moving through on the following green interval (*12*). It is quite similar to the load factor but is considered easier to develop. A loaded cycle need not be saturated. Saturation factors were utilized in *NCHRP Report 113* for evaluation of intersection levels of service.

Maximum individual delay/vehicle is the maximum delay experienced by a single vehicle in the time duration under consideration. This value may be used as an indicator of the magnitude of oversaturation. It is an intersection measure useful primarily as a descriptor of conditions. Maximum individual delay generally rises with increased input rate (*9*). Maximum individual delay is also found to be well correlated with mean system delay, maximum individual stopped delay, or maximum queue length (*7*). However, maximum individual delay is harder to measure because individual vehicles have to be traced.

Average delay is the average delay per vehicle on an intersection approach during some time period. It is a good indicator of the magnitude of saturation. It has been widely used as a control parameter and also as an optimization criterion of traffic flow through an intersection or through networks (*10, 12, 21*). In the past 25 years, several delay models have been developed and validated. Extensive study of delay at signalized intersections also has been undertaken using simulation (*12, 14*). Delay is also widely used as a measurement of the effectiveness of an intersection or network (*7, 9, 13*). Average delay is well correlated to other characteristics of intersections such as volume, queue length, and characteristics of signal operations (*7, 11, 12, 21*). Another important aspect of the delay parameter is its usefulness as an input to economic studies of an intersection or network operation.

Aggregate delay is the *total delay* to all vehicles on an intersection approach during some time period. It is an indicator of the magnitude of oversaturation, and is generally useful as a control parameter. A form of aggregate delay measure is utilized as a control parameter on volume-density controllers and some full-actuated controllers. Although total delay is correlated to other parameters such as queue length, average delay, and input volume (*7, 14*), average delay is used more frequently than total delay as a control parameter or optimized criterion, because average delay is a better indicator of levels of service to the individual driver. Total delay is more suitable to economic studies of intersection performance, and in this sense it is redundant with travel time.

V/C ratio is the ratio of demand volume to capacity. This is an intersection measure indicative of impending saturation. The measure has little value once saturation has occurred, because it is widely variable under conditions of forced flow. It may be used as a control parameter as well as to characterize existing conditions before saturation sets in. However, V/C ratio in conjunction with other measures, such as queue length, might have some utility as a measure of saturation. (V/C + queue) represents the degree of saturation.

Queue length is the number of vehicles in, or the length in feet of, a given queue. This is an intersection measure of utility in both characterizing conditions and on-line control. It is indicative of the magnitude of saturation, especially when compared to the "no delay" queue size (i.e., that queue which may be cleared during one green phase). Queue length is well correlated with other intersection measures, such as delay and input, and functions of the geometrics and operational characteristics of an intersection. It is the most frequently used control parameter with delay. Queue length may be the intersection measure that most directly affects drivers' behavior under saturated conditions because it is most readily observable. Various detection systems and techniques for estimating queue length have been developed in recent years (*6, 11, 14, 15, 29, 33*). Queue length is also an important parameter for the Urban Traffic Control System (UTCS) for Washington, D.C. (*33*), and the TRW South Bay System (*34*) for a segment of the Los Angeles area. In the latter system, in particular, queue length is used as a congestion indicator.

Number of stops/starts to clear intersection is related generally to the number of cycles to clear intersection, and is of similar utility. *Number of cycles to clear queue* is the number of cycles necessary to clear a given queue through an intersection, generally taken as a representative queue size for the period under consideration. Both measures are intersection measures indicative of the magnitude of saturation. Changes in the number of cycles to clear average queue or average number of stops and starts to clear an intersection may be used to detect the effect either of queues extending from downstream intersections or of temporary blockages (accidents, illegal parking, stalled vehicles) of moving lanes at the approach.

Input-output is related to continuous monitoring of demand into an intersection (input = volume through + queue build-up), and of volume leaving the intersection, and yields a continuous accounting of the number of vehicles accumulating on a given link. The rate and amount of vehicle accumulation in a given link not only indicates the degree of saturation, but it also is useful in predicting the onset of oversaturation in the link; however, its use is considered redundant with queue measurements.

Trapped vehicles refers to the existence, and numbers, of vehicles trapped in an intersection, and may be used as a clear indication that contamination from adjacent links has already occurred or that turning movements from the intersection in question are oversaturated. This measure, however, describes only the condition of an intersection after contamination has occurred and is not useful in describing the transition condition or the degree of oversaturation until contamination occurs.

Queue length/link length ratio is the ratio of the queue length to the length of the street link, where the link length is specified as the distance from the intersection under study to the next upstream intersection of note (minor side-street intersections might be reasonably excluded). *Link length minus queue length* is similar to queue length/link length ratio except that the absolute length or distance between the adjacent upstream intersection and the tail of the queue on the link is used as a parameter, rather than the relative percent occupancy of a link by a queue. The apparent advantage of these two measures over a simple queue measure is that the measures are a function of queue and link length, which is one of the most important geometric factors in traffic performance through urban networks. Link length affects vehicle speed, queue storage capacity, number of stops and delays per unit length of a network, interactions between two adjacent intersections, and signal operation.

EVALUATION OF CANDIDATE POINT MEASURES

A number of measures have been presented as being potentially useful as descriptors or as control parameters in the area of traffic control in the saturated/oversaturated environment. For the purposes of such description and/or control, it is important that any measure or measures finally selected should be able to perform two functions: describe and predict.

In its descriptive function, the measure should adequately indicate in which of the defined traffic performance states in the congested region of operation (i.e., stable saturation, unstable saturation, or oversaturation) the system of interest is operating.

In its predictive function, the measure should adequately indicate the probable onset of each of the defined traffic performance states in the congested region of operation.

Ideally, there should be one measure that can be used to both describe and predict over the entire region of interest. This may not be an achievable goal, and some set of measures individually or in combination may be required.

In addition to these prime requirements, the following factors must also be taken into consideration:

1. Does the measure have general applicability? That is, the measure should not be so dependent on specific physical and operating conditions that it requires extensive recalibration each time it is applied.
2. Is the measure easily observable under actual conditions?
3. Are one or more of the measures under consideration functionally equivalent? That is, can one candidate measure be expressed in terms of one or more other measures?
4. The data requirements for any measure as well as its utility as a control parameter also must be considered in the final selection of a measure or set of measures.

These requirements are summarized in Table B-2, and serve as the basis for evaluation of the available candidate measures. The following discussion centers on the application of these evaluation criteria to the candidate set of point measures. On the basis of these evaluations, a recommended set of point measures is selected. It should be noted that some candidate measures may be eliminated on the basis of a single criterion, whereas others may be eliminated through applications of a combination of criteria.

Are the Measures Capable of Adequately Describing the Defined Traffic Performance States? Are They Capable of Adequately Predicting the Probable Onset of the Various Traffic Performance States?

The majority of the candidate measures describe or model saturated traffic operations to varying degrees of adequacy. Some, however, do not measure saturation in direct terms, and some describe conditions only in combination with other measures.

Measures such as load factor or saturation factor do not differentiate saturation, either stable or unstable, from oversaturation. As these factors approach 1.0, they do indicate the onset of congestion and, as such, its first subcategory, stable saturation. They are unable, however, to describe any more of the congested range, in particular those traffic characteristics that presage the onset of oversaturation. Therefore, load factor and saturation factor are eliminated from further consideration.

Trapped vehicles (i.e., vehicles caught within the confines of an intersection) only occur when oversaturation has been reached. Thus, this measure would be useful to characterize or describe the existence of oversaturation. It is not useful, however, as a descriptor of any of the other traffic performance states or even as a predictor of the onset of oversaturation and is, therefore, eliminated from further consideration.

Are the Measures Functionally Equivalent?

Having eliminated a few candidate measures that obviously are inadequate, one is still faced with a considerable number that appear potentially useful and, therefore, merit additional consideration and evaluation. At this point, the several remaining measures will be investigated as to their functional equivalence. The measures in question are the following:

- V/C ratio.
- Maximum individual delay per vehicle.
- Number of cycles to clear intersection.
- Number of stops and starts to clear intersection.
- Total delay.
- Average delay.
- Queue length.
- Input-output count.
- Link-length to queue length relationship.

In reviewing these candidate measures it is apparent that they are not all independent of each other. For example, number of cycles to clear the intersection is dependent on the position of a given vehicle in a queue and, therefore, directly relates to queue length. Similarly, the delay measures all relate in some manner not only to each other but also to queue length. In terms of a direct relationship, the difference between input and output is queue length. Similarly, for a V/C ratio greater than 1.0, demand exceeds capacity and vehicles will not be able to pass through the signal in a single cycle. Thus, V/C ratio relates to queue length to input-output count, as well as to all the delay measures. On reviewing the remaining candidate measures one sees that they all relate to each other. The subject of functional equivalency was extensively analyzed and reported in (*15*). When two or more measures are available, the simplest, most precise, and easiest

TABLE B-2

EVALUATION CRITERIA

- Capable of Adequately Describing Traffic Performance States?
- Capable of Adequately Predicting Onset of Traffic Performance States?
- Easily observable in Field/Reasonable Data Requirements?
- Best of functionally Equivalent Subset of Measures?
- Utility as a Control Parameter?

to use should be selected. In fact, all of the measures can be related directly to some form of queue length measure. This does not mean that queue length is necessarily the measure to use, because some other, less direct measure might be significantly easier to observe and might have far lower data requirements.

Although all of the candidate measures relate to each other in one fashion or another, they are not all as easy to use. Consider, for example, the volume/capacity (V/C) ratio. A V/C ratio of less than 1.0 indicates less than saturated operation. At a V/C ratio equal to 1, demand exactly equals capacity, and the first boundary between congested and stable saturated operations has been reached. As V/C exceeds 1.0, demand exceeds capacity, and one is somewhere in the saturated range. Thus, V/C ratio is a potentially useful measure. How useful is it? Suppose one is able to determine that the demand volume/capacity ratio is, say, 1.5. Does this impart any information as to what is the current traffic state? Not really. For a measure such as V/C to be useful, it would be necessary to have a continuous record of V/C with time of the sort shown in Figure B-1.

Having such a continuous record of V/C, it would be possible to determine in which traffic performance state the system under consideration is operating. The whole process would, however, require considerable data to be useful, and is, in fact, less direct than other measures that are also volume dependent, as is the V/C ratio. For example, queue length is nothing more than the difference between demand, V, and capacity, C. Furthermore, a measure such as queue length does not require the same continuous history of data to be of immediate utility. Thus, although the V/C ratio is potentially useful it is not practically useful—particularly when compared to other, more direct measures.

Does the Measure Have General Applicability?

What is the general applicability of the candidate measures of interest? By general applicability is meant the capacity of a measure to be used directly, without qualification, regardless of the physical and operational characteristics of the particular intersection of interest. That is, a specific value of the measure means the same thing regardless of the specific physical and operational conditions encountered. This is an ideal situation, not usually attainable regardless of the measure used; nor is it the purpose for which it is used. Even the rather universal measure of flow vehicles per hour usually requires such qualifiers as number of lanes and mix of traffic. In any case, the "best" measures will approximate, as closely as possible, the ideal of complete generality of applicability.

In terms of general applicability, the set of queue-based measures appears to offer the best path for finding a general form of measure. In particular, those queue measures that explicitly incorporate the spacing between adjacent intersections offer the best path. These measures have already built into them one major physical variable, block length. The other candidate measures do not explicitly take geometry into account, and would, therefore, have to be calibrated for a range of different physical situations.

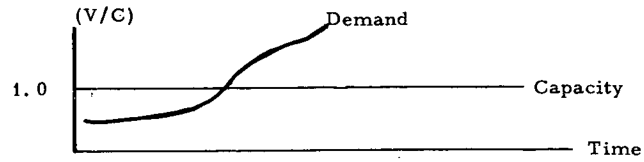


Figure B-1. Volume/capacity vs. time.

Does the Measure Have Utility as a Control Parameter? Does It Have Reasonable Data Requirements?

One control parameter for both existing and advanced control schemes is most desirable, if at all possible. An easily observable and measurable candidate is also desirable for existing control schemes without any detection systems. Several control schemes with the queue as a control parameter have been developed (*6, 7, 8*), and some of them are applied. Delay is currently the most widely used parameter in existing control schemes. However, once saturation occurs, optimization of flow, based on a delay parameter, is not valid because delay is no longer a primary problem. The primary task of the control scheme would then be prevention of contamination of other intersections. Queue parameters are also more easily observed and measured than delay parameters. Thus, a candidate measure expressed in terms of queue appears to be the most promising.

Selected Measures

The selected measures are queue length, queue length to link length ratio, and link length minus queue length. Each of the three point measures of saturation has the capability of describing the defined traffic performance states, as well as predicting the onset of each state. The latter two measures, those which explicitly incorporate link length, are considered to be desirable, because they are more general forms.

A limited field study has been conducted to aid in the determination of whether queue-related measures do adequately meet evaluation criteria in reality, and which of these three candidate point measures best describes the various levels of saturation. The field study and its results are discussed later in this appendix. In addition, simulation will be undertaken to further aid in determining which one of these measures best describes all conditions.

SYSTEM MEASURES OF SATURATION

Travel time is the most commonly used measure of system performance. Travel time through a system can be defined as the difference between the time a vehicle enters and exits the system. If an ideal travel time is defined as free-flow or desired travel time, the difference between the actual and the ideal travel times is delay, and can be used as an indication of the existence of congestion or a measure of the degree of congestion in a system.

Service ratio is usually expressed as vehicle-miles per a given time period, and is frequently used as a measure of system performance in conjunction with other measures, such as travel time or over-all travel speed. It has, however, more economic implications (*15*).

Delay ratio is defined as the ratio of average delay time to average travel time, and describes the magnitude of delay in a system (*12*).

Density is the number of vehicles per unit length of roadway or lane.

Occupancy is the percentage of time a detector is occupied or activated by vehicles. It is a function of density, velocity, and vehicle length. A measure of this type has been used in the Toronto, Canada, traffic control system in which the density and the queue length, estimated by vehicle speed over a detector, are used as control parameters (*14*). Density plus the speed profile of a link may be used to indicate the degree of saturation in the link.

Mean velocity gradient is acceleration noise divided by the mean velocity, where acceleration noise is the standard deviation of the change in a vehicle's velocity over time, and is a measure of the difference between a vehicle's actual speed and some uniform speed. It is a system measure of traffic irregularity and network performance. Studies (*12, 23, 24*) have found that it has a high degree of correlation with travel time on urban networks.

Average number of stops per vehicle is simply the average number of stops incurred in traversing a given segment of roadway, generally of appreciable length. It is widely used in travel time and delay studies to evaluate system performance, and is also used as an input to economic studies.

EVALUATION OF SYSTEM MEASURES

As in the evaluation of candidate point measures, the prime evaluation criteria are those that ask how adequate the measure is in (1) describing congestion states and (2) predicting the probable onset of each traffic performance state of interest. Each of the listed measures is reviewed in the following in terms of these two items.

The measure, travel time, suffers from ambiguity and insensitivity, when considered in terms of the problem at hand. Consider that a given value of travel time may represent two completely different system-wide congestion states (e.g., stable saturation throughout large sections of the network or oversaturation at a few locations with the rest of the network undersaturated). Since travel time does not enable differentiation between the several traffic performance states, it is rejected as a system measure.

Service ratio and delay ratio also suffer from the same inability to differentiate between saturation and oversaturation as does travel time.

Density is not sufficient in itself, because the presence of high density does not necessarily guarantee the presence of queues sufficiently long as to cause oversaturation. This measure may be useful in conjunction with queues or velocity measures. Such combined measures must be calibrated for individual links to permit their use as descriptors and predictors. If possible, a more general type of measure is desirable.

Occupancy, like density, does not provide a direct easy-to-use measure that will allow description of the various congestion states.

Mean velocity gradient has potential utility as a system measure, in that it reflects the degree of congestion that is desirable. It suffers, however, from being difficult to measure.

Average number of stops per vehicle, as with the travel time measure, suffers from being insensitive to degree of congestion. A given average number of stops per vehicle may occur from a system that is almost uniformly in a stable saturated condition, or from one that has a few localized points of oversaturation and is uncongested elsewhere.

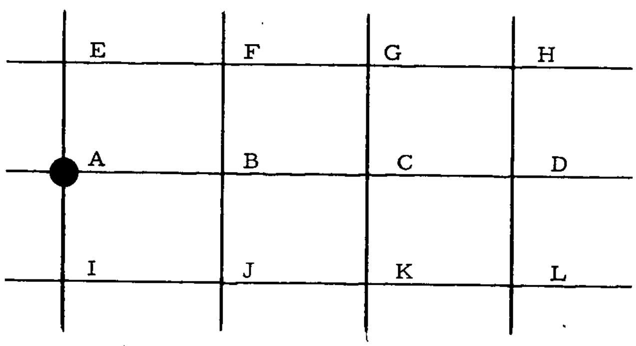
CONCLUSIONS ON SYSTEM MEASURES

It can be seen that all the existing candidate measures of system performance are inadequate when considered in terms of their ability to describe or predict the onset of the various defined levels of congestion.

The research agency judges that an appropriate point measure would be adequate for any foreseen purposes. This is based on the view that the problem itself, while frequently appearing to be system or subsystemwide initially, generates from a small number of local points of oversaturation.

The recommended approach is to use local measures in a system orientation, as shown in Figure B-2. It is structured to be consistent with the definition of oversaturation developed in this research. No attempt has been made to give it a special name.

If intersection A is oversaturated, and affects only the performance of one or more of intersections B, E, and I, but does not affect the performance of intersections beyond the adjacent ones, the oversaturation is considered to be local. However, if intersection B becomes oversaturated because of oversaturation at intersection A, and affects one or more of intersections C, F, or J, oversaturation is considered to be systemwide. In other words, when oversaturation at one intersection propagates to intersections beyond the adjacent ones, it is defined as system oversaturation. If two or more intersections oversaturate independently, this is independent local saturation and not system oversaturation.



Note: "A" is the Critical Intersection

Figure B-2. Local vs. system oversaturation.

APPENDIX C

LITERATURE REVIEW AND ANNOTATED BIBLIOGRAPHY

The literature was surveyed for all articles and papers concerning urban congestion, delay and saturation, and their root causes. HRIS was used, as well as independent reviews of major publication sources, including HRB records, bulletins and special reports, NCHRP reports, and the journals *Traffic Engineering, Traffic Engineering and Control, Transportation Science,* and *Transportation Research.*

Those articles of greatest applicability and relevance to the current research are listed, as follows, in the annotated bibliography. Additional references of secondary interest are also included.

This appendix is a companion to Appendix B, which cites references by the numbers used herein.

1. "System Analysis Methodology in Urban Traffic Control Systems—Phase II." *TRW Report 17283-4-11-RO-00.*
   Study of an oversaturated intersection in the midst of a coordinated progressive system. Mathematical optimization techniques are used to quantitatively measure the effect of the "bottleneck" intersection on the performance of the system. Various control strategies are examined and their effectiveness is quantitatively measured by simulation techniques.
2. ANCKER, C. J., and GAFARIAN, A. V., "A Simple Renewal Model of Throughput at an Oversaturated Signalized Intersection." *Transportation Research,* Vol. 1, pp. 57-65 (1967).
   Concerned with the effect of headway variability on the flow through an oversaturated signalized intersection (infinite queues). A probabilistic model is constructed to simulate the relationship between through outputs and other parameters, such as cycle length, average headway, and start-up time.
3. ANCKER, A. J., GAFARIAN, A. V., and GRAY, R. K., "The Oversaturated Signalized Intersection—Some Statistics." *Transportation Science,* Vol. 4 (Nov. 1968).
   Describes the examination of an oversaturated signalized intersection with no turning movements (where the queue before the intersection is only partly served during a single green phase). During the green phase, data on the time between arrivals of successive vehicles have been collected and analyzed statistically.
4. GAZIS, D. C., and POTTS, R. B., "The Oversaturated Intersection." *Proceedings of the Second International Symposium on the Theory of Traffic Flow,* London (1963).
   Study of the problem of minimizing delay at traffic signals that are overloaded during the peak period. In the first instance, it is assumed that the cycle is fixed and that each green period is restricted in range between a minimum and a maximum value. It is shown that, in the particular case in which the saturation flows in the two critical directions are equal, the minimum delay is given by the fixed-time settings that cause both queues to disappear at the same time.
5. GAZIS, D. C., "Optimum Control of a System of Oversaturated Intersection." *Operations Research,* No. 12 (1964).
   Describes the problem of optimizing the control of two oversaturated intersections (arrivals exceed the capacity-queue build-up), solved by using a semi-graphical method (see Ref. 4). A three-stage operation of the traffic lights is found to be the optimum control for a system of two oversaturated intersections. At every stage the service rate of any stream is either maximum or minimum.
6. LONGLEY, D., "A Control Strategy for a Congested Computer-Controlled Traffic Network." *Transportation Research,* Vol. 2, pp. 391-408 (1968).
   A new performance criterion for a saturated network is proposed. The basic control philosophy suggested in this paper is based on the fact that traffic lights cannot clear queues in conditions of primary congestion, and their function is thus to maintain the queues in a given ratio in order to delay the onset of secondary congestion arising from the blockage of other junctions by primary congested traffic. The implementation of this control scheme requires a central computer control and queue detectors.
7. GERLOUGH, D. L., and WAGNER, F. A., "Improved Criteria for Traffic Signals at Individual Intersection." *NCHRP Report 32* (1967) 134 pp.
   The effort includes development of multipurpose simulation model and applications to various traffic control techniques. Five control schemes were analyzed by simulation: (1) fixed-time single dial, (2) fixed-time optimized, (3) basic queue control, (4) queue length arrival rate control, and (5) modified space-presence control. Field tests were made of three control schemes—fixed-time control, volume density control, and experimental basic queue control—to determine their effectiveness.
   Simulation studies were also made of various measures —total delay, average delay, and mean system delay— to determine the adequacy of these measures in describing intersection performance.
8. WAGNER, F. A., GERLOUGH, D. L., and BARNES, F. C., "Improved Criteria for Traffic Signal Systems on Urban Arterials." *NCHRP Report 73* (1969) 55 pp.
   Involves investigation of various traffic signal control

concepts for arterial systems and testing of the following on exising urban arterials and by network simulation model: (1) Webster optimization of cycle and splits, (2) Yardeni's time-space design model, (3) Little's maximal bandwidth model, (4) delay/difference of offset method, (5) cycle and offset selection mode, (6) basic queue control mode, and (7) mixed cycle mode. The effectiveness of different concepts is determined by field tests and simulation and the results are compared. The "TRANS" network simulation model was refined and validated in the field.

9. WAGNER, F. A., BARNES, F. C., and GERLOUGH, D. L., "Improved Criteria for Traffic Signal Systems in Urban Networks." *NCHRP Report 124* (1971) 86 pp.
Involves investigation by computer simulation and field testing of various advanced methods of traffic signal control in urban networks with emphasis on the simple modification of fixed-time signal settings that commonly exist at the present time. The following 10 different signal timing schemes were investigated in various combinations to determine their effectiveness in improving traffic flow through urban networks: (1) Webster's cycle and split optimization, (2) delay-difference method, (3) volume-priority method, (4) preferential street method, (5) mixed-cycle method, (6) Little's maximal bandwidth method, (7) SIGOP, (8) British combination method, (9) Allsop's graphic theory method, and (10) basic queue control method.

10. WEBSTER, F. V., "Traffic Signal Settings." *Road Research Laboratory Technical Paper No. 39* (1961).
A delay model of fixed-time signalized intersections is developed and simulated. The model is also verified through field measurement. The model is a function of cycle length and the proportion of effective green to the cycle length as well as flow and the degree of saturation. Optimum settings of fixed-time signals, split and cycle length, which minimize the average delay, are derived from the delay model. Other characteristics of signalized intersection are also studied, including queues, among others. Average and maximum queue lengths are studied theoretically and in the field.

11. KELL, J., "Intersection Delay Obtained by Simulating Traffic on a Computer." *Highway Research Record 15,* (1963) pp. 73-97.
Time simulation of an intersection of two, 2-lane, two-way streets, controlled by STOP signs. Delay is studied under varying volumes and turn volumes to determine the relationships between intersection delay and approach volumes, as well as turning movements. Intersection delays are also compared for intersections controlled by STOP signs and signals to determine the effect of signals on the intersection delay.

12. PONTIER, W. E., MILLER, P. W., and KRAFT, W. H., "Optimizing Flow on Existing Street Networks." *NCHRP Report 113* (1971) 414 pp.
Comprehensive investigation of traffic flow improvements and resulting benefits in downtown areas that can be achieved by application of traffic engineering techniques. Field experiments were made to quantify the effect of traffic improvements. The experimental improvements can be grouped into the following six major areas: (1) directional control and land use, (2) curb-lane control, (3) channelization, (4) signal control, (5) inclement weather effects, and (6) bus operations. Various relationships were also developed to describe the quality of traffic flow, attaining a level-of-service definition for downtown streets. Methods were also developed for application of the results to other areas.

13. WEINBERG, M. I., GOLDSTEIN, H. MCDADE, T. J., and WAHLEN, R. H., "Digital-Computer Traffic Signal System for a Small City." *NCHRP Report 29* (1966) 82 pp.
Describes the manner in which a control doctrine was developed, and a digital-computer-controlled traffic signal system is synthesized for a small city. A delay model that minimizes aggregate delay is developed. It also describes requirements for instrumentation and interpretation of sensing by detectors. The City of White Plains, N. Y., was selected as a traffic model to study the system requirements as well as the system cost.

14. CHRISTENSEN, A., "Use of a Computer and Vehicle Loop Detectors to Measure Queues and Delays at Signalized Intersections." *Highway Research Record 211* (1967) pp. 34-53.
Vehicle detection system used in the Toronto computer-controlled signal system. A computer program obtains, from the pulses of loop detectors, the traffic parameters of volume, speed, space and time headways, and density. A computer program finds queue length and delay at signalized intersections by finding the relations between time headway and queue length and time headway and delay. Estimated queue lengths are verified by field data for accuracy.

15. "System Analysis Methodology in Urban Traffic Control Systems—Phase I." *TRW Report FH-11-6883* (June 1969).
The primary purpose of the study was to develop a second-generation surveillance methodology for automatically collecting traffic data in an urban network, transmitting it to a central computer, and processing it for system evaluation purposes. An extensive review of an existing or planned computerized surveillance and control system is made. The objective of a traffic control system is defined as measures of effectiveness for determining how well these objectives are met. The traffic parameters required for such measures of effectiveness are also determined. Studies are also made through simulation on sensor location, communications concepts, and data processing methodology to determine accuracy and instrumentation requirements. The state of the art of flow/control interaction theories and models is reviewed and recommendations for future studies are made.

16. GREENSHIELDS, B., *The Quality of Traffic Flow.* Bureau of Highway Traffic, Yale Universiy (1961). The report presents the development of a measure called a quality index of traffic flow for determining the level of quality of traffic movements, principally in urban areas. The quality index is a function of average speed, absolute sum of speed changes per mile, and number of speed changes per mile. The index was field tested and found to be well correlated with other intersection performances in urban networks.
17. MAY, A., JR., and PRATT, D., "A Simulation Study of Load Factor at Signalized Intersection." *Traffic Engineering* (Feb. 1968).
18. MAY, A., JR., and GYAMFI, P., "Extension and Preliminary Validation of a Simulation of Load Factor at Signalized Intersections." *Traffic Engineering* (Oct. 1969).
19. GEORGE, H. P., *Measurement and Evaluation of Traffic Congestion.* Bureau of Highway Traffic, Yale University (1961).
20. KELL, J. H., "Analyzing Vehicular Delay at Intersections Through Simulation." *HRB Bulletin 356* (1962) pp. 28-39.
21. NEWELL, G. F., "Approximation Method for Queue with Application to the Fixed Cycle Traffic Light." *SIAM Review,* Vol. 7, No. 2, (April 1965).
22. LITTLE, J. D., MARTEN, B., MORGAN, J., "Synchronizing Traffic Signals for Maximal Bandwidth." *Highway Research Record 118* (1966) pp. 21-47.
23. HELLY, W., and BAKER, P. G., "Acceleration Noise in a Congested Environment." *Vehicular Traffic Science, Proceedings of the Third International Symposium on the Theory of Traffic Flow* (1967).
24. UNDERWOOD, R. T., "Acceleration Noise and Traffic Congestion." *Traffic Engineering and Control* (July 1968).
25. MILLER, A. J., "A Computer Control System for Traffic Networks." *Proceedings of the Second International Symposium on the Theory of Traffic Flow* (1963).
26. HOLROYD, J., and HILLER, J., "Area Traffic Control in Glasgow," *Traffic Engineering and Control* (Sept. 1969).
27. HEWTON, J. T., "The Metropolitan Toronto Traffic Control Signal System." *Proceedings of the IEEE* (April 1968).
28. "Improved Street Utilization Through Traffic Engineering." *HRB Special Report 93* (1967) 234 pp.
29. BOTTGER, R., "A Simulation Study of a Detector-Arrangement for Determination of Queue Length and Delay." *Simulation,* Ier Symposium International Sur La Regulation Du Traffic.
30. "Detector Location." An ITE Information Report, *Traffic Engineering* (Feb. 1969).
31. SAGI, G., and CAMPBELL, L., "Vehicle Delay at Signalized Intersection." *Traffic Engineering* (Feb. 1969).
32. HARRISON, A., and FINDLAY, C. M., "Some Characteristics of Saturation Flow." *Traffic Engineering and Control* (Aug. 1969).
33. *Advanced Control Technology in Urban Traffi Control Systems.* U.S. DOT FHWA Contract No. FH-11-6932, Sperry Rand Corporation (Oct. 1969).
34. STANFORD, M. R., "Los Angeles County to Install Nation's First Computerized Traffic Control System." *Better Roads* (Aug. 1971).
35. *Network Flow Simulation for Urban Traffic Control System.* U.S. DOT FHWA Contract No. F4-11-7462-2, Peat, Marwick, Mitchell and Co.
36. MARRUS, B. and MAIN, M., "New Method Improves Signal Timing." *Traffic Engineering* (June 1964).
37. GAZIS, B., "Traffic Control Time Space Diagram, and Networks." *Traffic Control-Theory and Instrumentation,* Horton, T. (Ed.), Plenum Press, N.Y. (1965).
38. YAGODA, H., "The Control of Arterial Street Traffic." Paper presented at the 1967 IFAC Control Conference, Haifa, Israel.
39. CHANG, A., "Synchronization of Traffic Signals in Grid Networks." *IBM Journal of Research & Development* (July 1967).
40. MCSHANE, W., "The Control of Vehicular Traffic on Urban Arterials and Networks." Ph.D. dissertation, Polytechnic Institute of Brooklyn (June 1968).
41. YAGODA, H., STEVENS, E., and SABALA, E., "A New Technique for the Optimal Timing of Two-way Arterials." *ASME Publication 70-Tran-39* (1970).
42. WHITEHEAD, D. W., "The Toronto System—Intersection Evaluation and Control." *Traffic Engineering* (Sept. 1969).
43. GREENSHIELDS, B., SHAPIRO, D., and ERICKSEN, E., "Traffic Performance at Urban Street Intersection." *Technical Report No. 1,* Yale Bureau of Highway Traffic (1947).
44. SIEGEL, R. L., "Buses on Urban Networks." Master Project, Polytechnic Institute of Brooklyn (June 1972).

APPENDIX D

DATA ANALYSIS—CHARACTERISTICS AND STUDY OF SELECTED MEASURES

The primary purpose of a data check of the selected measures (see Appen. B) is to determine their adequacy in describing or modeling the properties of oversaturation, as well as the onset of oversaturation in reality. The need for a data check is even more acute because the measures selected may not necessarily be the best under actual conditions. A data check is also necessary to aid in determining which of the selected measures best meets the evaluation criteria. For these reasons, and because a very limited amount of work has been done in regard to oversaturation, a set of measures (queue length, link length minus queue length, and ratio of queue to link length) was selected for study on the basis of a literature search, an extensive review of numerous candidate measures, and a program of field work. The following information was sought at each study site:

1. Speed of individual vehicles through the intersection, departing from the upstream intersection, and queue position (moving and standing queues) at the critical intersection to determine the effect of the queues on traffic leaving the upstream intersection.
2. Cycle-by-cycle measurements of input and output, at both the critical and the upstream intersections, and queue length at the critical intersection to evaluate the stability of queues with respect to the input and output flows.
3. Signal offset between a pair of intersections to evaluate the effect of offset on the vehicles leaving the upstream intersection.
4. Signal timing, intersection movements, traffic composition, pedestrian flow, and geometric data to determine the capacity of each approach, as well as the effect of pedestrians on the capacity.

To properly evaluate the properties of oversaturation and the onset of oversaturation, it was also necessary to examine traffic behavior under uncongested conditions. Thus, the study covers the range of performance states from uncongested to oversaturated operations.

STUDY SITES

Extensive field trips were made by the research staff throughout the New York Metropolitan Area to locate suitable sites. Several meetings were held with local traffic engineering units for the same purposes. Upon careful evaluation of various candidate sites, three pairs of intersections in Brooklyn, N.Y., were selected.

The reasons for selecting the three sites, as well as a brief description of the three, are presented in the following. Figure D-1 locates the sites, and Figures D-2 through D-4 provide geometric and timing data.

There are many intersections in Manhattan that become oversaturated during peak hours. However, no sites in Manhattan were selected because of short block length (200 to 300 ft) and unusually high percentages of taxicabs.

Traffic Characteristics

Sites 1 and 2

Sites 1 and 2 were selected for study because oversaturation is observed at these sites very frequently, almost regularly during peak hours, and the duration of oversaturation is sufficiently long to make a meaningful study. The degree of oversaturation and duration is of a magnitude that normally is observed on arterials in large metropolitan areas where problems of oversaturation frequently exist. There are relatively small turning movements at the approaches of interest, thus minimizing the effects of turning vehicles and simplifying the basic analysis.

Site 3

Traffic volumes are lower at this location than at Sites 1 and 2. However, queues extend to the upstream intersection during afternoon peak hours from time-to-time. The condition persists for two or three cycles whenever it occurs. Unlike Sites 1 and 2, queue build-up is caused by heavy right-turning movements into the link of interest, thus presenting an opportunity for study of intersection oversaturation due to cross-street traffic turning into the arterials.

Geometrics

Geometric features of the test sites are generally adequate. Sites have block lengths of 500 ft; the third site has a 900-ft long block. Lane widths are 11 ft in all cases. Grades are sufficiently low as to not affect traffic performance.

Sites 1 and 2 are on a 6-lane arterial street, and Site 3 is a two-way street with one moving lane and one parking lane in each direction. Thus, these three sites represent a typical range of urban streets.

Traffic Control and Traffic Pattern

Signs and markings at these sites conform to standards. The intersections are controlled by fixed-time signals that are most common in urban areas. The signals are not physically interconnected. The signal timing offset at Site 2 is progressive, and offsets at Site 1 and 3 vary because cycle lengths at the upstream and the critical intersections are different. This varying offset presents an excellent opportunity to study the effects of offset on vehicles leaving the upstream intersection.

Site 1

This site is a section of Flatbush Avenue Extension northbound, near downtown Brooklyn. Flatbush Avenue Extension is one of the major arterials in Brooklyn. At its northerly end it is connected to the Manhattan Bridge, which is one of four East River crossings. This site includes intersections at Myrtle Avenue and Tillary Street, which have 100-sec and 120-sec cycles, respectively. During morning peak hours, congestion occurs at Tillary Street, and queues frequently extend to Myrtle Avenue.

The morning peak normally lasts about 2 hr, between 7 and 9 a.m., and oversaturated conditions occur at about 7:15 a.m. and last for approximately 30 min. The site has three through lanes plus one parking lane in each direction. Parking is prohibited during peak hours and a good compliance is observed.

Oversaturation at Site 1 could be attributed to the following:

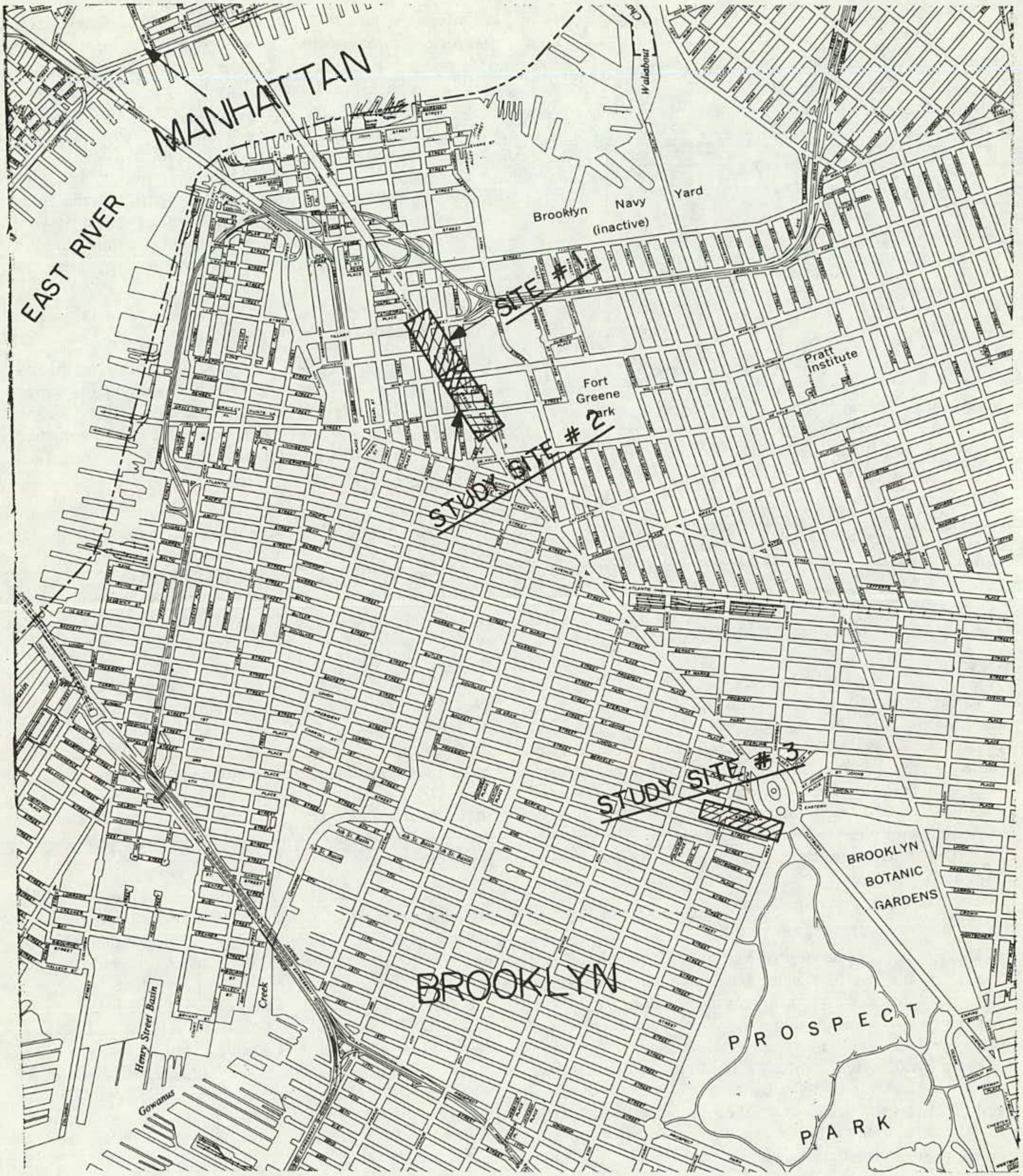


Figure D-1. Study sites location map.

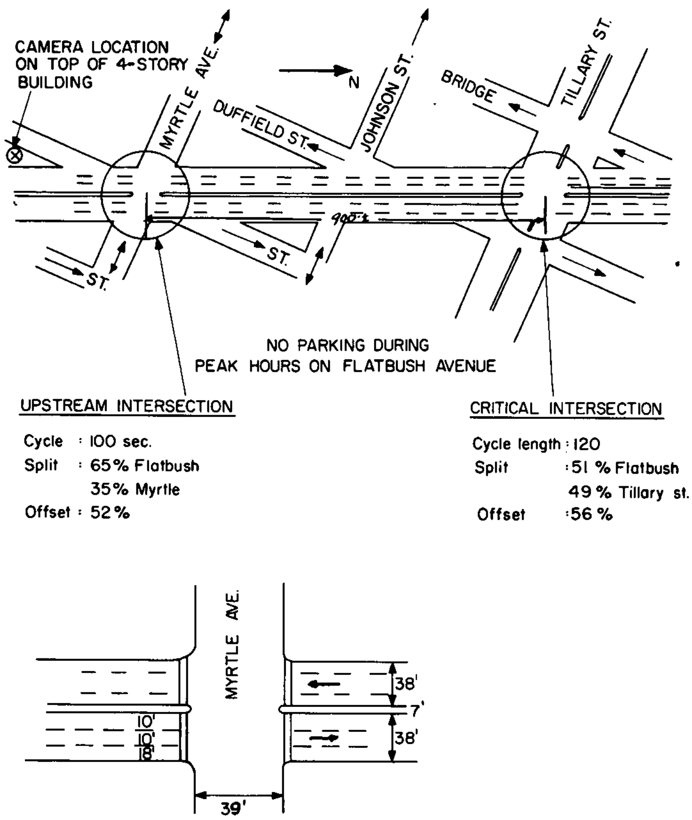


Figure D-2. Geometrics and timing at Site 1.

1. Use of a 120-sec four-phase signal at Tillary Street, resulting in a smaller green/cycle ratio and thus less capacity along Flatbush Avenue Extension as compared to the two-phase 100-sec cycle at the upstream intersections.
2. Heavy left-turning volumes from Flatbush Avenue Extension to Tillary Street (29 percent). The present 12-sec left-turn phase is found inadequate.
3. Two different cycle lengths creating offsets ranging from zero to 100 sec.

Site 2

This site is also a section of Flatbush Avenue, adjoining Site 1 on the south, which includes the intersection of Willoughby Street as the upstream intersection and Myrtle Avenue as the critical intersection (an intersection whose capacity is the limiting value of a segment of roadway or of an entire system). Most of the traffic and operational characteristics here are similar to those at Site 1. Congestion at this site is caused by the shorter green at Myrtle Avenue as well as by the queues that extend from Site 1. Analysis indicates that the latter is the major cause of congestion.

Capacity analysis shows that only the Tillary Street intersection has a real capacity shortage.

Site 3

The third site is a section of Union Street between 8th Avenue and Flatbush Avenue in Brooklyn, located about 2 mi south on Site 1. Union Street is a two-way street with one through lane in each direction, with parking allowed on both sides. Congestion occurs at the intersection with Grand Army Plaza during early afternoon peak hours, and queues frequently extend to the 8th Avenue intersection. A large number of vehicles making turns into Union Street from 8th Avenue, a short green period at the Grand Army Plaza intersection, and varying offsets due to two different cycle lengths are the main causes of congestion. Peak volumes were observed between 2:30 and 4:30 p.m., and oversaturated conditions that last for a few cycles at a time occurred frequently during this period.

DATA COLLECTION

Two possible methods of data collection were investigated: manual collection and filming techniques. Upon reviewing the various aspects of these alternatives, including their advantages and disadvantages, the filming method was selected for the following reasons:

1. Because of the complexity and microscopic characteristics of the data required, a manual method not only requires a great deal of manpower and training effort, but also involves a greater amount of human error.
2. Filming provides an opportunity for visual studies as well as continuous coverage of the transition behavior of traffic from undersaturated to oversaturated conditions (onset of oversaturation), which are difficult to evaluate by manual study methods.
3. The research agency had in operation a complete system of ground-based cameras, which were developed for NCHRP Project 3-15, "Weaving Area Operations Study." The project also provided a well-trained crew, thus minimizing the cost for system development and crew training.
4. There are many high-rise office and apartment buildings near the sites, which provided potential vantage points. Good vantage points are essential for ground-based camera data collection.

Two cameras were used to film Site 1, each camera covering one intersection. Site 2 was filmed by one camera because of the shorter block length and an excellent vantage point. Only one camera was used at Site 3, which did not give complete coverage because of problems associated with the vantage point. Thus, only a limited data check was feasible at Site 3. Signal indications, as well as intersection movements, are clearly visible on the films except for those from Site 3. The films also show a built-in clock, providing a reference time on each frame. Filming was done by 2 frames per sec to 4 frames per sec to provide accuracy. Approximately 30 to 60 min of data were collected at each site.

DATA REDUCTION

Data reduction was achieved by "pseudo-field" observation of films, by a team of two technicians. The actual field operation of the study intersections was projected on a screen and viewed at a different rate from real time, with stop-go, back-up, and restart capability, as necessary, when

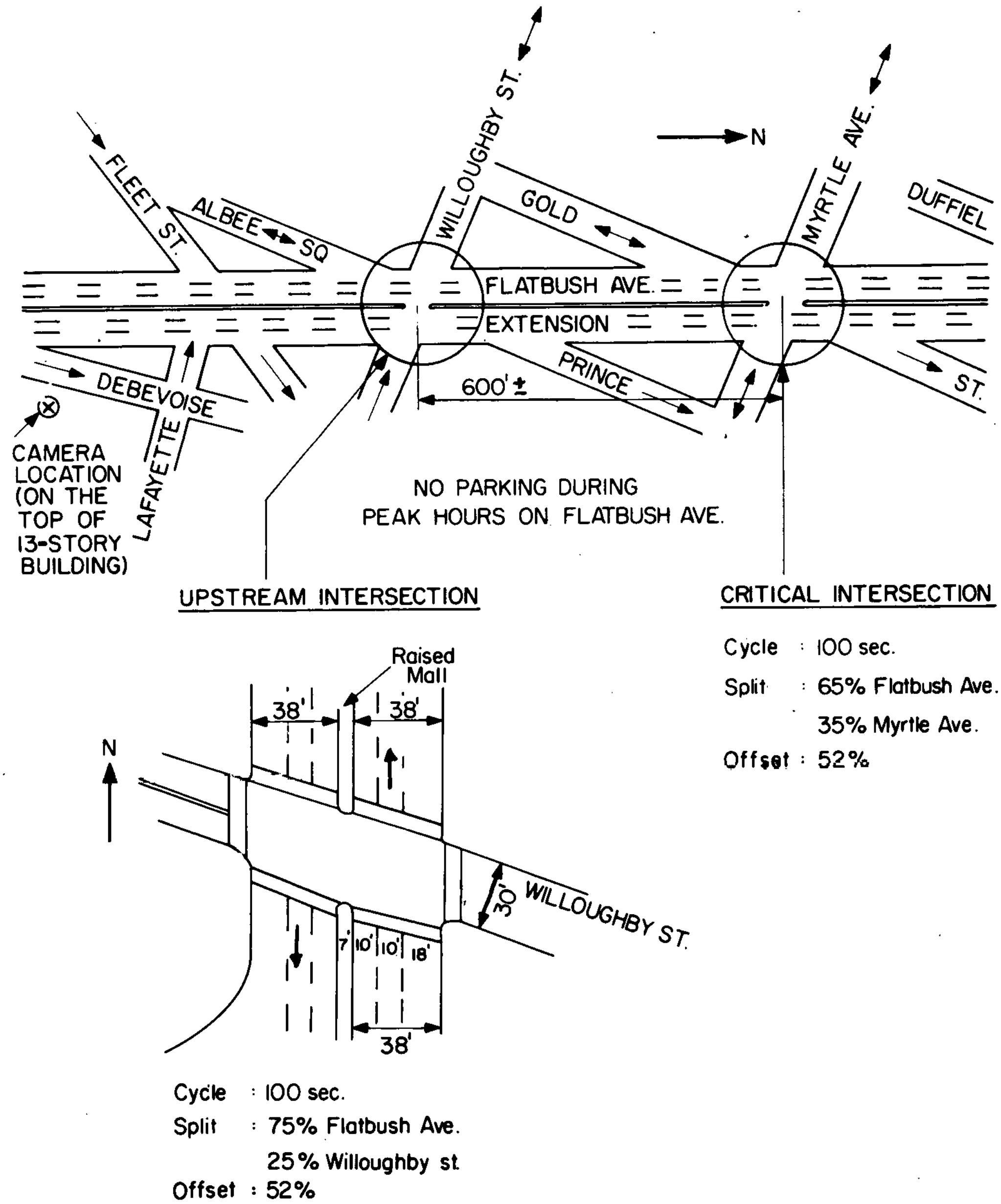


Figure D-3. Geometrics and timing at Site 2.

vehicles were being traced through intersections. Relevant information was reduced and recorded onto reduction forms. Because of the limited amount of data involved, data analysis was performed manually.

To measure queue lengths at critical intersections, the link between the critical and the upstream intersections was divided into a number of segments (each segment is approximately 50 to 100 ft long), using reference markers available at the site (see Fig. D-5). Thus, the ends of queues can be identified by {n + 1} different positions (n segments of the link plus a no-queue position). The ends of queues are identified by lane each time an observed vehicle crosses into the upstream intersection. This is referred to as the instantaneous queue length.

Data from all three sites were reduced. However, data from Sites 1 and 2 were mainly used to check the adequacy of the selected measures of effectiveness (MOE). Use of the data from Site 3 was limited because the end of the queue was out of the film more frequently than expected. Refilming of Site 3 was ruled out because of the difficulty of reaccessing the vantage point. Furthermore, Sites 1 and 2 were considered adequate to achieve the intended objectives of the field data check of the MOE.

Reduced data include the following information recorded on the reduction forms:

1. *Cycle-by-Cycle at Upstream Intersection*—The following data are classified by lead vehicles, vehicles that were stopped at beginning of "green," and vehicles that did not stop.

   a. Time of the beginning and end of "green" at both upstream and critical intersections.
   b. Time vehicles cross near-side and far-side STOP lines. For turning vehicles, only the time that

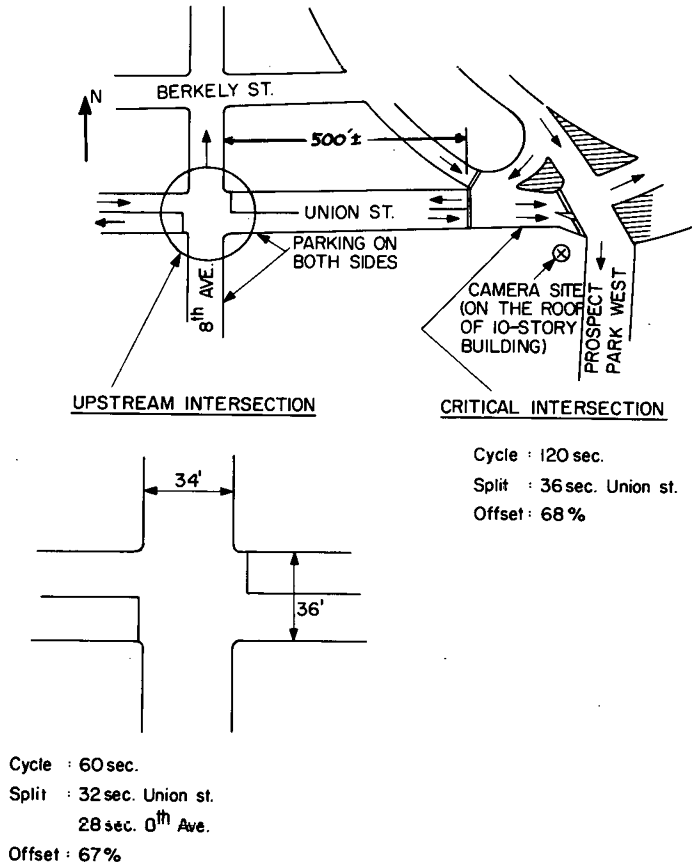


Figure D-4. Geometrics and timing at Site 3.

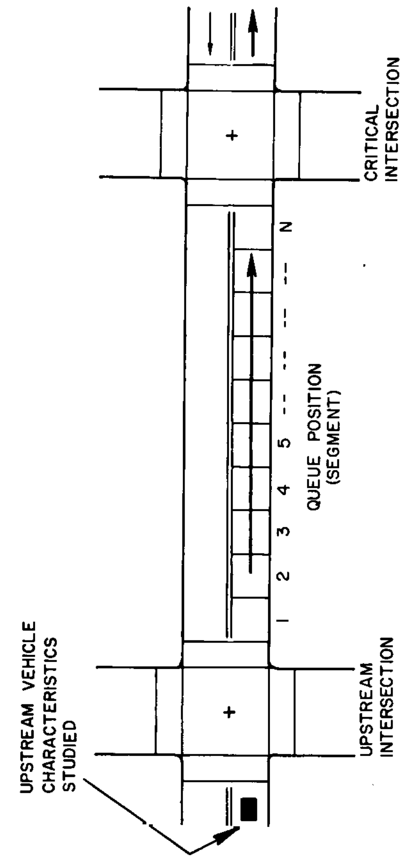


Figure D-5. Definition of segments for queue study.

vehicles cross the near-side STOP line is recorded.

c. Positions of queues by lane at the critical intersection at the time an observed vehicle enters into the upstream intersection.

2. *Cycle-by-Cycle at Both Intersections*

a. Intersection movements.
b. Movements by traffic composition.
c. Pedestrian flow.

The data analysis was mainly directed toward determining whether queues at the critical intersection significantly affect the performance of traffic leaving the next upstream intersection. In particular, start-up time, speed of vehicles leaving the upstream intersection, and headway distribution were examined.

DATA ANALYSIS

Start-Up Time

Start-up time of lead vehicles leaving the upstream intersection (i.e., time lag between the beginning of "green" and the time the lead vehicles start motion) was analyzed with respect to stopped-queue positions at the critical intersection to determine the effects of queue position on start-up times. Because of the limited number of data points, the analysis was carried out for the entire approach instead of by each lane.

Figure D-6 shows that the least start-up time is required when the queue is between positions 4 and 6. However, this result is considered inconclusive because of the very limited number of data points at both extreme ends of the queue positions (queue positions 1, 2, 7, and 8 in Table D-1).

Statistical tests indicate that the difference in the average start-up time between the group of queue positions 1, 2, and 3 and the group of queue positions 4, 5, and 6 is significant, and so is the difference between the group of queue positions 4, 5, and 6 and the groups 7 and 8. Additional data, however, are required for more detailed results. No analysis of start-up time was made for Site 1 because of the far smaller number of data points available.

Crossing Time

Crossing time is needed to determine, in terms of a selected MOE, the time when one intersection affects the performance of the upstream intersection. The data are analyzed by lane as well as for the entire approach. Analysis of crossing time is also made with respect to moving queues, but no consistent relationships are found—possibly, partly because of the difficulties inherent in the definition of moving queues. (*Moving queues* are defined as vehicles approaching standing queues and decelerating—differentiated by brake light actuation).

In this project, the time required for vehicles to cross the intersection (STOP line to STOP line) when leaving the upstream intersection was analyzed with respect to both vehicles stopped at the critical intersection and vehicles not stopped. The objective of these analyses was to determine the effect of queue position on the traffic leaving the upstream intersection, as well as to identify the point at which the traffic behaves differently.

The performance of an intersection can be expressed in two terms: qualitative and quantitative. Quality is usually measured in terms of speed, travel time, or delay. Analysis of the crossing time (speed) involves measurement of the quality of intersection operation.

Quantitative intersection operation can be measured in terms of the productivity, such as the output, of an intersection (see discussion under "Headway Distribution," treated later).

Vehicles Stopped at Beginning of "Green"

As shown in Figure D-7(a), the average crossing times at Site 2 for stopped vehicles plotted against queue positions for each lane indicate that average crossing times decrease at a high rate as the queue length decreases from position 1. However, when the queue end is at position 4 (this is about the ⅓ point of the link), or shorter, crossing times decrease at a much lower rate. The average approach crossing time shows a much clearer relationship between the average crossing times and queue position than do those on a per lane basis. For Site 1 (Fig. D-8(a)), no such conclusions are possible because no data points are available. The average crossing time, standard deviation, and sample numbers available for Sites 1 and 2 are given in Table D-2.

Vehicles Not Stopped

An analysis similar to that made for vehicles stopped at the beginning of "green" was made for vehicles that did not stop. Figures D-7(b) and D-8(b) and Table D-3 show the average crossing time, standard deviations, and sample numbers available for Sites 1 and 2, by lane as well as for the approach. Average crossing times generally decrease as queue length decreases, at a faster rate between positions 1 and 3 and at a much slower rate thereafter. Unlike the case of vehicles stopped at the beginning of the green, the decrease in mean crossing time is still significant even after the queue length decreases to beyond position 3.

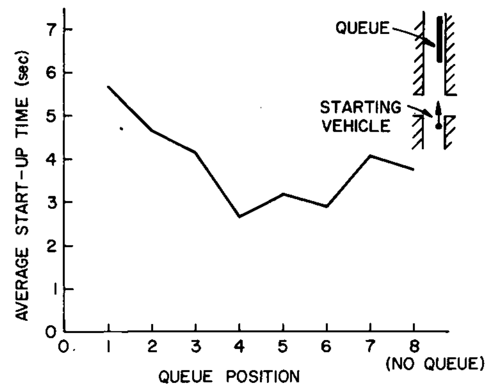


Figure D-6. Start-up time (Site 2).

This may be an indication that speeds of vehicles that did not stop are much higher and, thus, more sensitive to queue positions than speeds of vehicles that stopped.

There appears to be more inconsistency and scatter of data for Site 1 than for Site 2. This may be due to the more dynamic nature of operation, at Site 1, resulting from the heavy left turn (29 percent) at the critical intersection and varying offsets.

Statistical analysis indicates that, for both stopped vehicles and vehicles that did not stop, there are significant decreases in the average crossing times when queue length is equal to or smaller than queue position 4.

Left-Turn Effects on Crossing Time

For both "stopped vehicles" and "nonstopped vehicles," a further analysis was made to determine the effects of turning vehicles on the crossing time. Though turning vehicles are not included in the computation of average crossing times, they certainly would have some effect on the vehicles following them.

The average crossing times with and without including the vehicles immediately following turning vehicles are compared for both Sites 1 and 2. The differences are not statistically significant. There are no right turns at either site, and left turns are 3 percent and 6 percent at Sites 1 and 2, respectively.

Thus, it appears that turns of this magnitude do not have any effect on the speed of vehicles leaving upstream

TABLE D-1

START-UP TIME OF UPSTREAM VEHICLE BY QUEUE POSITION, SITE 2

Queue Position in Downstream Link	Average Start-Up Time (Sec.)	Standard Deviation	Number of Data Points Available
1	5.67	4.72	6
2	4.63	1.93	4
3	4.14	2.32	11
4	2.65	0.82	10
5	3.17	1.64	27
6	2.90	1.23	39
7	4.06	0.99	4
8 (no queue)	3.75	1.22	4

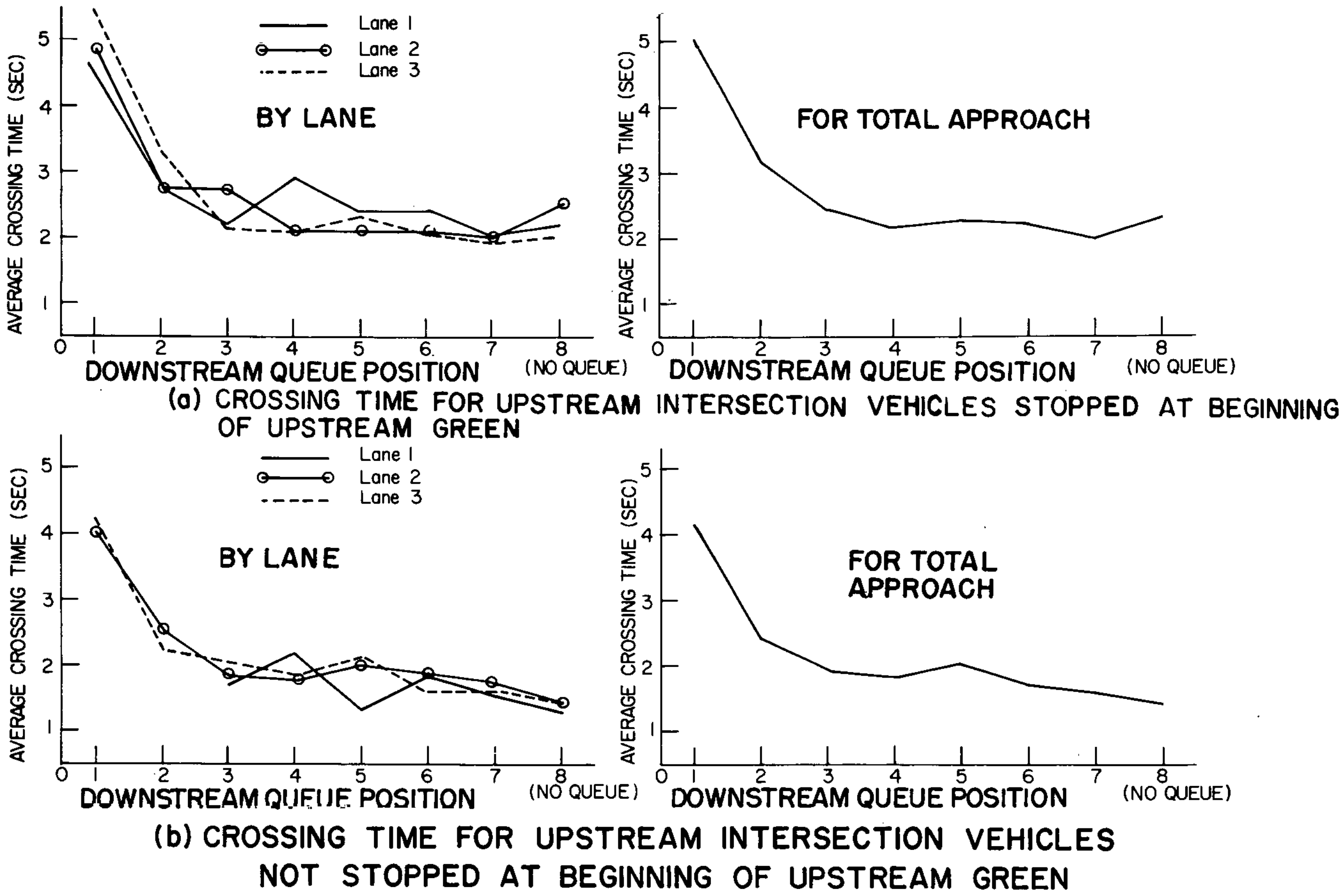


Figure D-7. Average crossing time (Site 2).

intersections with respect to queue length at the downstream intersection.

Lane-by-Lane Variations in Crossing Time

The average crossing times are also compared by lane for both sites. Although it appears that there are some differences, they are not statistically significant. Left-turn vehicles may be responsible for any difference observed between lanes.

Signal Offset and Crossing Time

The effect of signal offset (actually signal state) between the critical intersection and the upstream intersection on crossing time was also investigated. The crossing times of vehicles leaving the upstream intersection when the signal was red at the critical intersection were compared with those vehicles leaving the upstream intersections when the signal was green at the critical intersection. The results of this comparison are shown in Figure D-9 as a function of stopped-queue position. The difference between the two crossing times is statistically significant. The two curves, however, are very similar in characteristics to each other as well as to the average crossing-time plot (developed without differentiation of the state of the critical intersection signal state). Table D-4 gives the average crossing time, standard deviation, and sample sizes differentiated by the critical intersection signal state.

One notes from Figure D-9 that crossing times of vehicles faced with a red signal indication at the critical intersection are longer than those of vehicles faced with a green signal indication at the critical intersection. It would appear as if those seeing a red signal at the downstream intersection realize that it is likely that they will have to stop and join the queue that has developed at the critical intersection. Thus, they appear to be less motivated to exit from the upstream intersection than are those drivers who see a green and, perhaps optimistically, expect the queue to be moving when they reach it.

Average Crossing Time vs. Average Queue Length

As mentioned earlier, a relationship has been demonstrated to exist between the average crossing time of a vehicle and the instantaneous downstream queue length. Should a similar relationship be found to exist between average crossing time and average queue length, one would have a more easily measurable MOE and its utility, as a measure for traffic engineering practitioners, would be increased.

Figure D-10 and Table D-5 show the average crossing times compared against average queue length over a cycle. For vehicles stopped at the beginning of "green," the

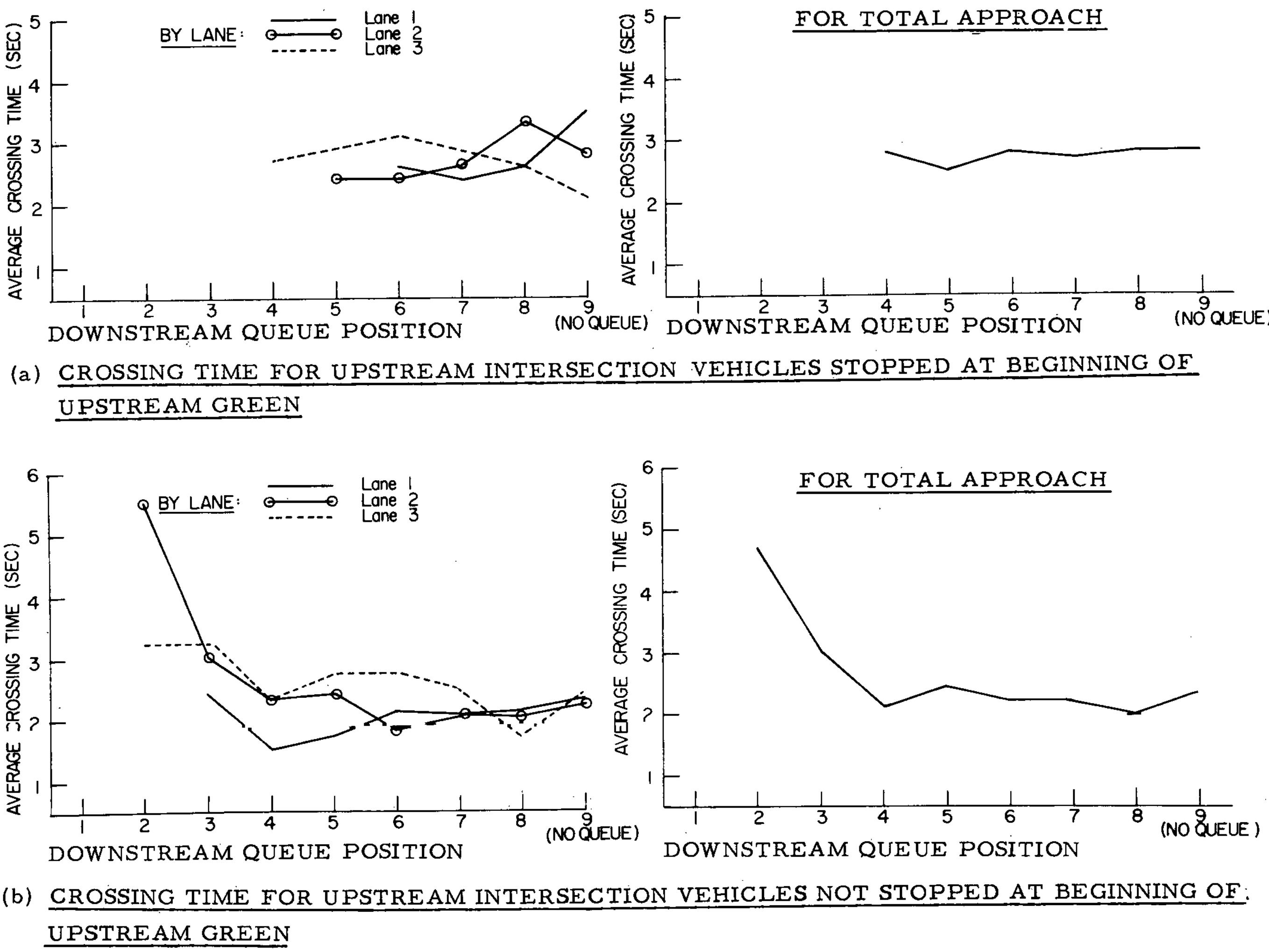


Figure D-8. Average crossing time (Site 1).

relationship is quite similar to that for the instantaneous queue. However, the relationship for vehicles "not stopped" is quite different from that for instantaneous queues, especially when the average queue position is between 1 and 3. This is probably because of the higher speed of vehicles "not stopped," which is likely to be much more sensitive to the queue length. Thus, it appears that the average queue over a cycle is too insensitive to be an effective MOE in the desired range of traffic operations.

HEADWAY DISTRIBUTION

As noted initially, the performance of an intersection can be measured in qualitative and quantitative terms. The quality is herein measured by crossing time (speed). To quantify the performance of an intersection, the headway distribution of vehicles leaving an upstream intersection is analyzed with respect to the queue length at the downstream intersection.

As shown in Figure D-11, there is no consistent relationship between the average headway and queue length for either "stopped" vehicles or "not stopped" vehicles. There is also no statistically significant difference between the average headways.

Because FLOW = 1/HEADWAY, the implication is that the output of an intersection remains essentially the same regardless of the queue length at the downstream intersection as long as the queue does not attain such length as to actually cause physical blockage of the upstream intersection.

PEDESTRIAN EFFECTS

Pedestrian activity during the study period was minimal, and their effect was not analyzed. However, the research agency recognizes that oversaturated networks in or near a CBD may be greatly affected by pedestrian conflicts.

The research agency did not specifically test the exclusive pedestrian phase "Barnes dance" form of control. The effectiveness of this form of control at a saturated intersection is doubtful in that it requires relatively long cycles with a significant amount of time not available to vehicle movement. This tends to generate long queues that hasten the speed of congestion. Local experience in the vicinity of

TABLE D-2

AVERAGE CROSSING TIMES FOR STOPPED VEHICLES BY LANES

SITE #2

Queue Position	LANE #1			LANE #2			LANE #3			APPROACH		
	$\bar{X}$	S D	Sample Number	$\bar{X}$	S D	Sample Number	$\bar{X}$	S D	Sample Number	$\bar{X}$	S D	Sample Number
1	4.50	-	1	4.82	1.83	11	5.42	2.44	6	5.00	1.96	18
2	2.75	-	1	2.75	0.71	9	3.33	2.83	18	3.13	2.26	28
3	2.20	0.27	5	2.72	0.63	33	2.13	0.57	32	2.41	0.65	70
4	2.90	0.43	5	2.10	0.50	20	2.09	0.72	37	2.16	0.68	62
5	2.38	0.67	32	2.11	0.62	49	2.28	0.63	32	2.23	0.64	113
6	2.38	0.57	31	2.08	0.56	15	2.05	0.36	25	2.20	0.52	71
7	2.06	0.72	4	2.00	-	1	1.90	0.28	5	1.98	0.46	10
8(no queue)	2.19	0.58	22	2.50	0.80	16	1.98	0.60	12	2.24	0.68	50

SITE #1

Queue Position	LANE #1			LANE #2			LANE #3			APPROACH		
	$\bar{X}$	S D	Sample Number	$\bar{X}$	S D	Sample Number	$\bar{X}$	S D	Sample Number	$\bar{X}$	S D	Sample Number
1												
2												
3												
4							2.7	0.57	4	2.8	0.51	5
5				2.4	0.64	13				2.5	0.63	14
6	2.6	0.40	3	2.4	0.50	6	3.1	0.79	11	2.8	0.72	20
7	2.4	0.41	6							2.7	0.68	9
8	2.6	0.63	23	3.3	0.91	18	2.6	1.46	21	2.8	1.08	62
9(no queue)	3.5	1.38	7	2.8	0.90	9	2.1	0.36	9	2.8	1.06	25

$\bar{X}$ = estimate of mean   S D = estimate of standard deviation

TABLE D-3

AVERAGE CROSSING TIMES FOR VEHICLES THAT DID NOT STOP

SITE #2												
	LANE #1			LANE #2			LANE #3			APPROACH		
Queue Position	$\bar{X}$	S D	Sample Number	X	S D	Sample Number	X	S D	Sample Number	X	S D	Sample Number
1				4.04	2.33	7	4.19	2.30	17	4.15	2.26	24
2				2.52	0.84	32	2.22	0.76	23	2.39	0.82	55
3	1.70	0.27	5	1.86	0.53	32	2.03	0.51	39	1.94	0.51	76
4	2.20	0.76	5	1.78	0.80	56	1.85	0.84	37	1.83	0.81	98
5	1.31	0.22	8	1.99	1.47	45	2.14	1.42	44	2.00	1.40	97
6	1.83	0.81	10	1.85	1.01	55	1.59	0.54	66	1.72	0.80	131
7	1.51	0.51	38	1.73	0.83	64	1.60	0.40	71	1.63	0.62	173
8(no queue)	1.28	0.36	81	1.45	0.62	124	1.43	0.50	176	1.41	0.52	381
SITE #1												
	LANE #1			LANE #2			LANE #3			APPROACH		
Queue Position	X	S D	Sample Number	X	S D	Sample Number	X	S D	Sample Number	X	S D	Sample Number
1				19.7	20.29	12	12.7	17.57	13	16.1	18.90	25
2				5.5	6.16	7	3.2	0.98	4	4.7	4.94	11
3	2.4		2	3.0	0.60	8	3.2	0.91	5	3.0	0.74	15
4	1.5	0.39	6	2.3	0.72	9	2.3	0.89	12	2.1	0.79	27
5	1.7	0.65	7	2.4	0.55	7	2.7	1.52	13	2.4	1.18	27
6	2.1	0.52	16	1.8	0.38	8	2.7	1.02	11	2.2	0.84	35
7	2.0	0.76	16	2.0	0.51	11	2.5	0.97	19	2.2	0.83	46
8	2.1	0.63	38	2.0	0.64	31	1.7	0.47	21	2.0	0.62	90
9(no queue)	2.3	0.51	7	2.2	0.64	8	2.4	0.69	9	2.3	0.60	24

$\bar{X}$ = estimate of mean   S D = estimate of standard deviation

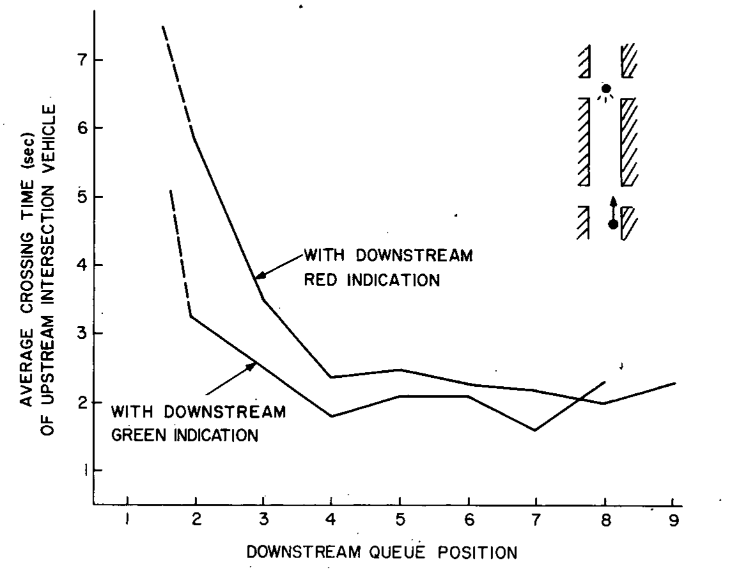
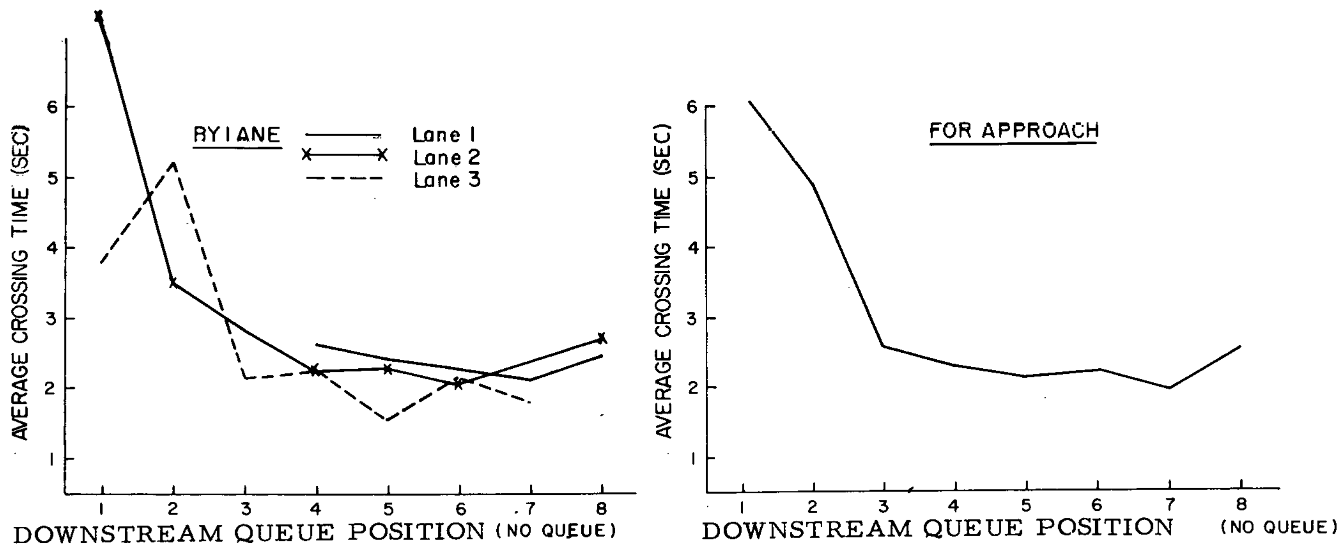


Figure D-9. Average crossing time by downstream signal indication.

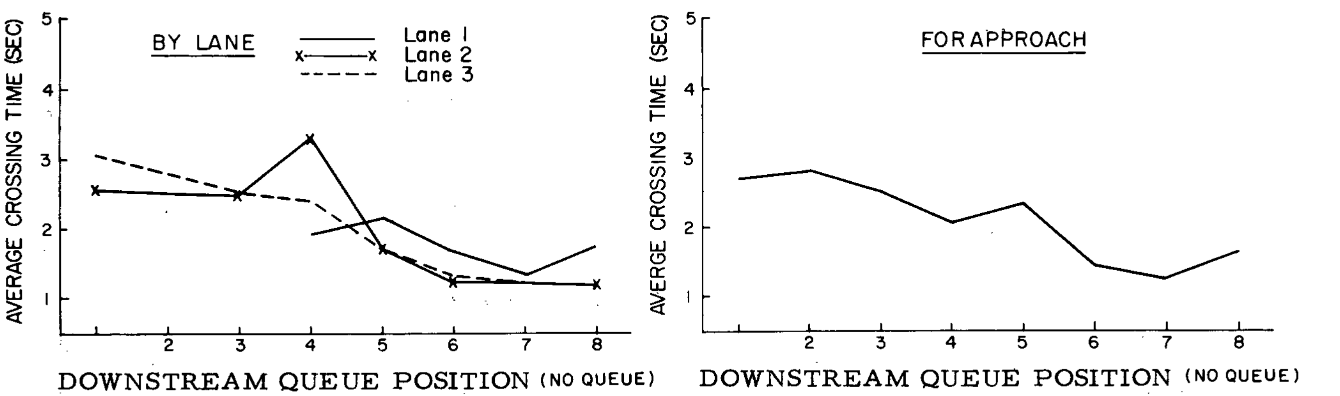
TABLE D-4

AVERAGE CROSSING TIMES BY DOWNSTREAM SIGNAL INDICATION

Queue Position	Lane #1	Lane #2	Lane #3	Approach	
	$\bar{X}$	$\bar{X}$	$\bar{X}$	$\bar{X}$	SD
RED INDICATION					
1					
2		6.9	4.0	5.9	6.6
3		3.2	4.2	3.5	0.7
4	1.1	2.9	2.5	2.4	1.0
5	1.4	2.6	2.8	2.5	1.3
6	2.0	2.0	2.7	2.3	0.9
7	2.1	2.0	2.6	2.2	0.8
8	2.2	2.0	1.6	2.0	0.6
9 (no queue)	2.4	2.2	2.4	2.3	0.6
GREEN INDICATION					
1					
2		3.7	2.5	3.2	1.5
3		2.7	2.6	2.5	0.5
4	1.7	1.9	1.9	1.8	0.4
5	2.1		2.4	2.1	0.6
6	2.3	1.7		2.1	0.4
7	1.5		1.6	1.6	0.8
8	2.0			2.3	0.7
9 (no queue)					



(a) CROSSING TIME FOR UPSTREAM INTERSECTION VEHICLES STOPPED AT BEGINNING OF GREEN
(VS. AVERAGE QUEUE POSITION OVER A CYCLE)



(b) CROSSING TIME FOR UPSTREAM INTERSECTION VEHICLES NOT STOPPED AT BEGINNING OF GREEN
(VS. AVERAGE QUEUE POSITION OVER A CYCLE)

Figure D-10. Average crossing time vs. average queue.

the research agency's offices has indicated the reduction in congestion achieved by elimination of the all-red pedestrian phase.

IMPLICATIONS OF DATA ANALYSIS

The average crossing time (speed) of vehicles leaving the upstream intersection is significantly affected by the queue at the downstream critical intersection. On the basis of the analysis of data from Sites 1 and 2, when a queue extends beyond position 4 the average crossing time significantly increases.

Because the two sites have two different block lengths (530 and 180 ft), the MOE should be expressed as the distance between the end of the queue to the upstream stop line (i.e., clear distance). The distance is 230 ft, and is designated as the critical distance. The MOE should also be applicable to blocks of any length. However, further field verification would be necessary for blocks much shorter than 500 ft to make the finding more conclusive.

Start-up times appear to be similarly affected by the queue at the downstream intersection; again, the analysis is not conclusive.

Although the quality of operation is reduced, the productivity of the upstream intersection is not affected by a queue at the downstream section because the headway of vehicles leaving the upstream is not affected. The output will, of course, be reduced when spillback occurs.

Since time headway equals space headway divided by speed, and since headway remains the same, space headway decreases as speed decreases. This implies that as the queue reaches the critical length, the speed and space headways decrease so that the output remains the same. The decrease of space headways results in a more compressed traffic stream and a higher density.

TABLE D-5

AVERAGE CROSSING TIMES FOR VEHICLES BY LANES VS. AVERAGE QUEUE POSITION OVER A CYCLE

THOSE THAT STOPPED								
	LANE #1		LANE #2		LANE #3		APPROACH	
Queue Position	$\overline{X}$*	Sample Number	$\overline{X}$	Sample Number	$\overline{X}$	Sample Number	$\overline{X}$	Sample Number
1			7.27	24	3.77	11	6.17	35
2			3.50	3	5.19	13	4.88	16
4			2.85	31	2.14	22	2.56	53
4	2.63	4	2.25	6	2.25	27	2.29	37
5	2.43	35	2.28	38	1.57	30	2.13	103
6	2.28	52	2.07	18	2.16	24	2.21	94
7	2.13	4			1.80	5	1.94	9
8	2.45	11	2.70	5			2.53	16
THOSE THAT DID NOT STOP								
	LANE #1		LANE #2		LANE #3		APPROACH	
Queue Position	$\overline{X}$*	Sample Number	$\overline{X}$	Sample Number	$\overline{X}$	Sample Number	$\overline{X}$	Sample Number
1			2.57	50	3.05	21	2.71	71
2					2.81	34	2.81	34
3			2.49	84	2.52	63	2.50	147
4	1.93	14	3.30	10	1.89	83	2.09	107
5	2.18	73	1.72	71	1.72	102	1.86	246
6	1.69	128	1.24	67	1.33	78	1.48	273
7	1.36	11			1.23	15	1.29	26
8(no queue)	1.78	25	1.21	6			1.67	31

* Mean

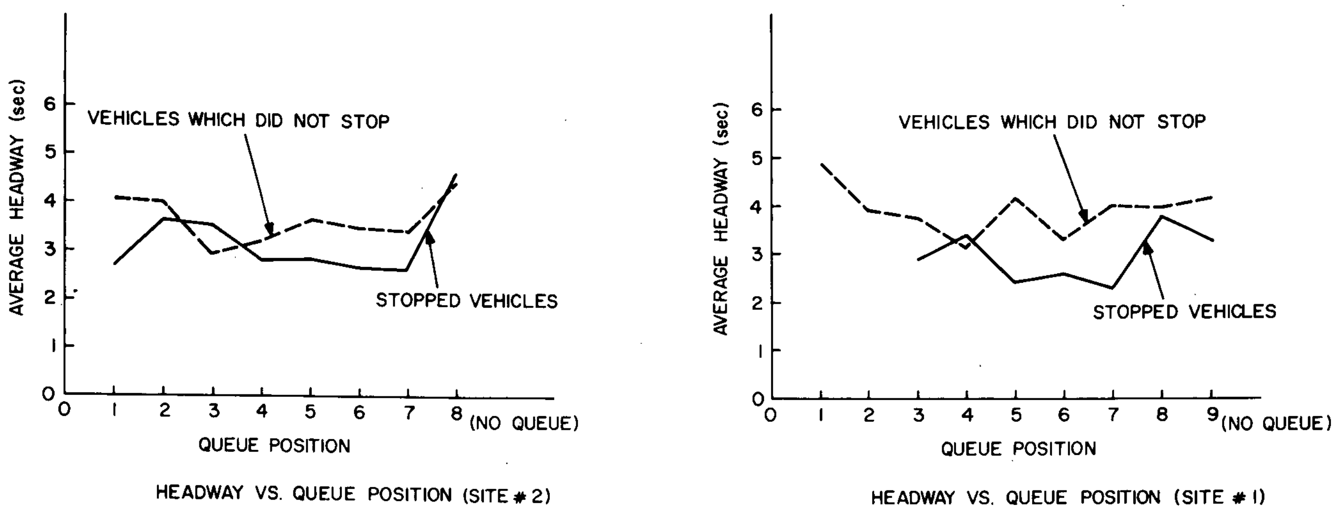


Figure D-11. Headway vs. downstream queue position.

Average queue length over a cycle is not sensitive enough to be used as the MOE. However, it is of use as an MOE in the absence of a detector system capable of yielding short-term measures of instantaneous queue length.

Regardless of the queue length at the downstream intersection, the status of the downstream signal indication affects the quality of operation (speed) at the upstream intersection. However, it will not affect the selected MOE.

No conclusive analysis is possible with the available data concerning the effects of left turns on the selected MOE; it does appear, however, that large percentages of left turns would be required for any significant effect.

APPENDIX E

A WORKING PAPER ON DETECTOR-OBSERVED TRAFFIC VOLUME CHARACTERISTICS

The regularity of daily traffic patterns is of basic concern when evaluating various candidate control strategies. At one extreme, if little variation exists from day-to-day, a time-of-day (minimal response) controller will undoubtedly suffice. At the other extreme, if no discernible pattern exists, or if it arrives randomly each day, a highly responsive control is appropriate.

Between these two extremes, there is the possibility of a basic underlying pattern with substantial variation about it. There is also the possibility that on a given day there will be a major deviation from an otherwise extremely regular pattern. To accommodate such possibilities, traffic data may be collected in real time and deviations noted, or a predictor may be established to estimate future values.

When considering the control of congestion, it is vital to consider (1) the regularity of daily patterns and (2) the regularity of the times at which various flow levels are reached. This will provide insight into the merits of minimal response control versus highly responsive control in this flow regime.

An extensive data base was acquired through the courtesy of the Metropolitan Toronto Department of Roads and Traffic, and was used to investigate the following topics:

1. The regularity of the weekday pattern.
2. The regularity of the time at which certain flow levels are reached.
3. The correlation of the flow between lanes on the same approach.
4. The potential for refining estimates of traffic volumes, based on observation of the deviations from the historic or nominal pattern.

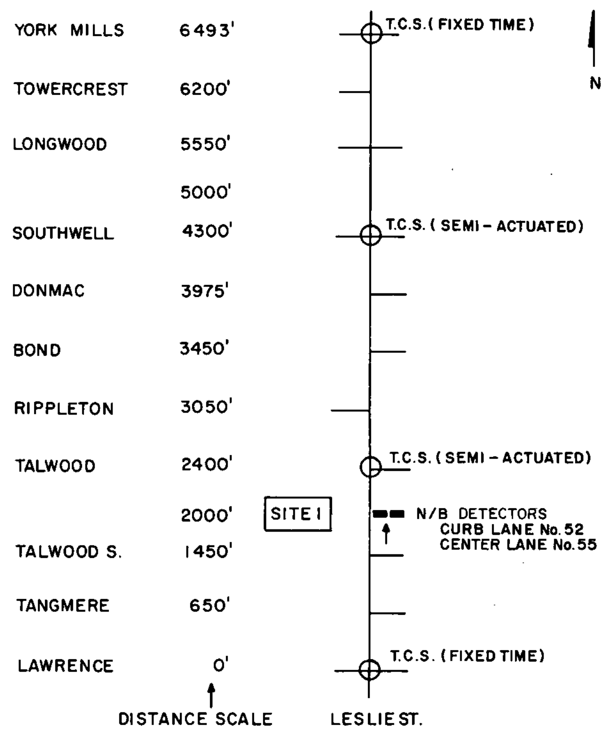


Figure E-1. Location of Site 1 (suburban).

DATA BASE

The Metropolitan Toronto Department of Roads and Traffic provided the researchers with an extensive data base, consisting of 5-min samples of volume by lane at each of 4 sites over a 77-day period. The data were collected by the Toronto computer system from September 24, 1973 to December 10, 1973.

Figures E-1 and E-2 depict the detector locations for each of the four sites. Figure E-3 shows photographs of the four sites to provide some indication of the streets on which the data were taken.

The data were acquired from the detectors whenever possible. Table E-1 gives the distribution of the samples by day of week and by hour of day. Note that all periods of common interest are well covered. The data were output in 5-min counts, with the first counting period initiating whenever the computer "came up." For simplicity, all data were shifted to standard 5-min periods, with the standard periods initiating at 0000 to 0005 hours (i.e., midnight to 5 min after midnight). The time shift thus introduced was uniformly distributed between −2.5 and +2.5 min. This could introduce a standard deviation of $\sigma = 1.44$ min in any time shift estimates.

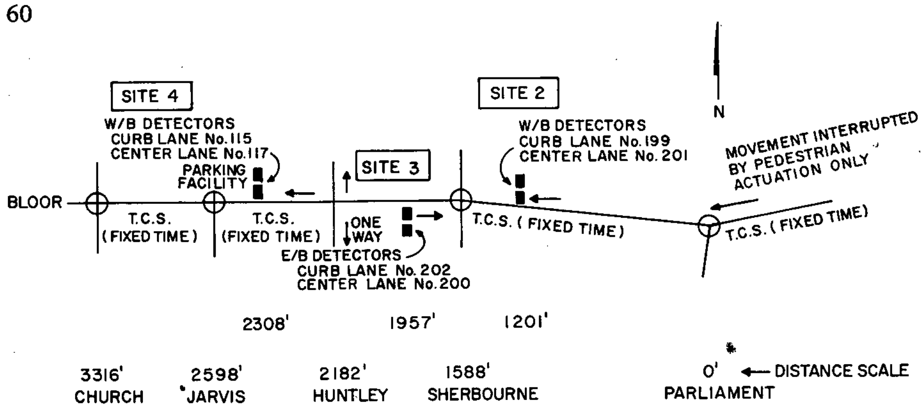


Figure E-2. Location of Sites 2, 3, and 4 (central area).

REGULARITY OF THE DAILY PATTERN

The average pattern was computed for each day by averaging the volume observed in each time slot. Because the weekdays did not differ substantially, a single average pattern was also computed aggregating all weekdays at each detector. Figure E-4 shows the average weekday pattern, and the Saturday and Sunday patterns, for Site 2. Also shown is the average pattern for Monday and Friday.

Figure E-5 illustrates the average weekday pattern at each of the four sites. The shaded area indicates the region within which one can expect 95 percent of the observations of volume to fall. This is not a confidence bound on the average. The confidence bound is much tighter. It is an estimate of the fluctuations from day-to-day within a specified time slot.

One may observe in Figure E-5 that, although there is substantial variation most of the time, the peaking is quite sharp, with relatively little variation in the time at which certain levels are reached. That is, a specified level X is reached at approximately the same time every day; this is variation in the horizontal dimension. One can even question whether the variation in the vertical dimension is as "substantal" as it appears: the 95 percent range is in the order of $\pm$ 120 vphpl, or $\pm$ 10 vehicles per lane in a 5-min period. This occurs when the average count is in the order of 40 vehicles per lane in a 5-min period. Further, the distribution is approximately normal, so that the deviations tend to be clustered near zero.

Figure E-6 shows the distribution of the times at which certain volumes are reached at Site 2. Note that the variation is rather small, leading to the conclusion that minimal response policies could be developed with some assurance that the onset of certain levels could be anticipated with some confidence. At the same time, the variation that does exist precludes extremely rapid preplanned switching of the control settings.

It should be noted that if one takes advantage of the regularity of pattern and does not have surveillance detectors, there is no protection against statistical "rare events"—major deviations from the average pattern. Generally, these can be associated with weather or special community activities.

CORRELATION OF FLOW BETWEEN LANES

If a volume is to be observed, it is necessary to determine if one lane will suffice, or if two or more lanes must be measured. For the data available, which contained only two-lane approaches, the question reduces to "one or both."

It was found that all weekdays could be aggregated for a given site and that a two-regime curve relating total volume to sub-lane volume should be established. The breakpoint was established at 360 vph in the curb lane, and no data for a curb-lane volume below 240 vph were considered.

From Table E-2, it can be seen that the total volume is rather strongly correlated to curb-lane volume in the higher flow regime, with correlation coefficients in the order of 0.9 and above.

A least-squares fit was executed on the data for each site, with the two curves tied to a common point at the boundary between regimes:

$$Y = \begin{cases} A + B\,(X - X_0) \\ A + C\,(X - X_0) \end{cases} \tag{E-1}$$

where:

$Y =$ total 5-min count;
$X =$ curb lane 5-min count; and
$X_0 =$ 30; boundary between regimes.

The results are summarized in Table E-2, and are plotted in Figure E-7. Note that the slopes of lines in the $X \geq 30$ regime are comparable, except for Site 1, the suburban site.

Clearly, the lane split tends to equalize as volume increases. The downtown sites tend to have greater concentrations in the outer (left-hand) lane, although data are certainly not sufficient to identify location as a causative factor.

On the basis of these fits, it is recognized that the total volume can in fact be computed with some confidence, if the calibrated lines are known. Indeed, for certain ranges, an assumption of 50 : 50 split will suffice. If a representative calibrated curve is used, it should be recognized that the actual split between lanes may depend on turning volume and other factors. Figure E-8 presents an approximation drawn from Figure E-7, and does not address

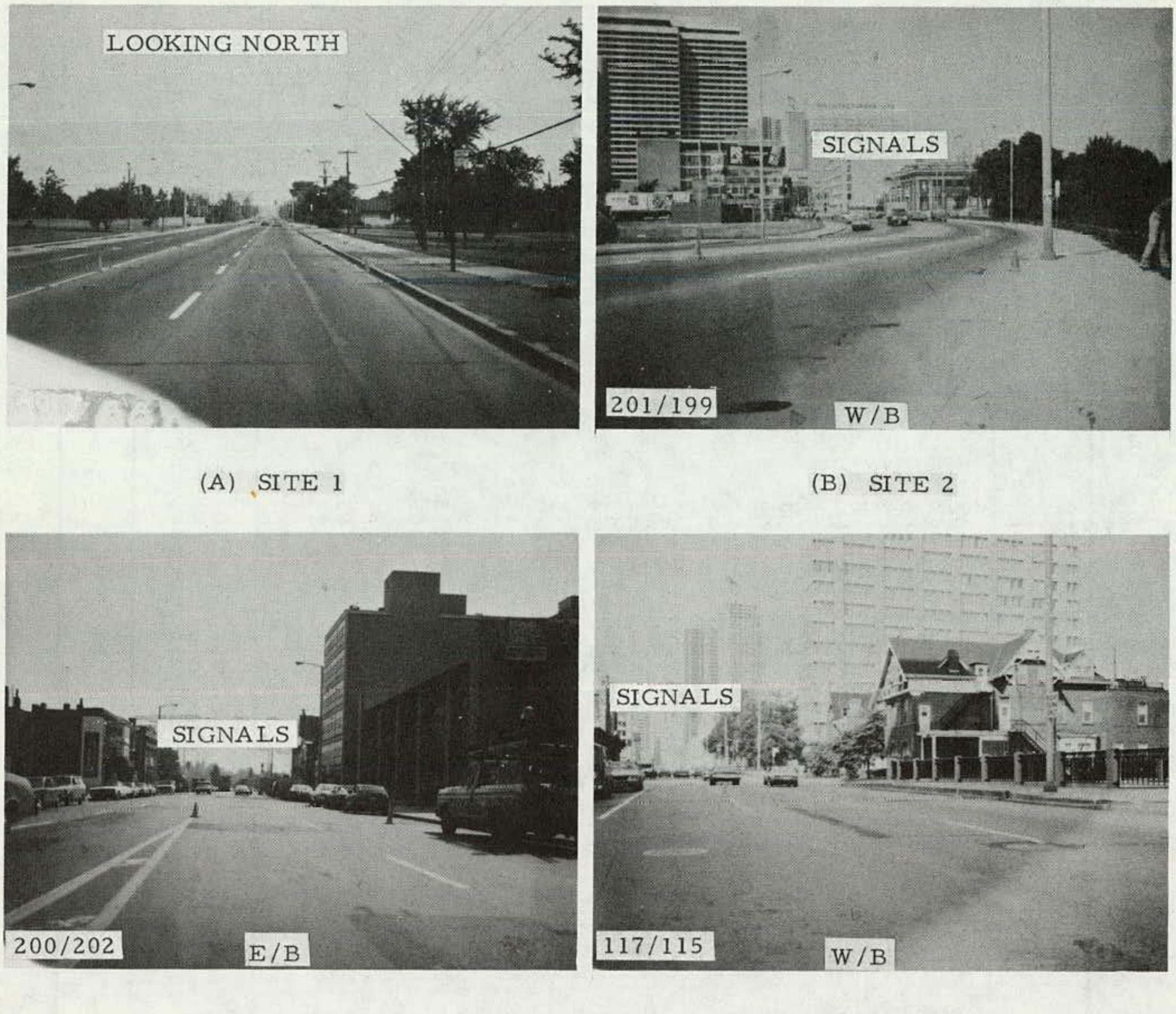


Figure E-3. Photographs of the four sites (note nos. shown are detector nos.).

these refining factors.

It was determined that the weekend curves are not dissimilar to the weekday curves within the range of the weekend data.

POTENTIAL FOR PREDICTION

On any given day, the actual pattern differs from the average pattern by some set of discrepancies, ϵ_i. At a particular site, it is assumed that a_i equals the average count for time period i, x_i equals the actual count for time period i, and ϵ_i equals $x_i - a_i$. If the set of ϵ_i is serially uncorrelated, there is no hope for predicting a future value x_i any better than simply saying that its expected value is a_i. If there is some correlation among the ϵ_i, however, there is information contained in the sequence of past values of ϵ_i. This information may be extrapolated into some future period, k, to get some better estimate (i.e., prediction) of a value x_{i+k}.

Predictors may be based on nothing more than historic patterns, or they may ignore the historic pattern and project forward on the basis of current trends, or they may project forward a refinement to the historic pattern on the basis of recent (i.e., real-time) deviations from the historic pattern. Because the weekday pattern itself is so regular (see Figures E-5 and E-6), it appears most fruitful to investigate whether there is any additional information in the set of discrepancies.

The set of $\{\epsilon_i\}$, which is the outcome on any given day, may be viewed as the outcome of a zero-mean process. It is of interest to compute the autocovariance values at one and two lags, $R(1) = E[\epsilon_i \epsilon_{i+1}]$ and $R(2) = E[\epsilon_i \epsilon_{i+2}]$, as well as the variance $\sigma^2 = E[\epsilon_i^2]$. If a sample of length N data points were available, these could be estimated by

TABLE E-1

DISTRIBUTION OF SAMPLES BY HOUR OF DAY AND DAY OF WEEK

Hour \ Day	1	2	3	4	5	6	7	Total	Percent
1	115.	69.	46.	38.	52.	92.	104.	516.	3.9
2	69.	0.	0.	12.	12.	26.	83.	202.	1.5
3	96.	26.	0.	12.	10.	0.	72.	216.	1.6
4	96.	56.	0.	35.	13.	14.	72.	286.	2.2
5	97.	80.	29.	68.	23.	36.	79.	412.	3.1
6	129.	116.	74.	89.	99.	75.	96.	678.	5.1
7	143.	113.	97.	109.	129.	104.	113.	808.	6.1
8	144.	132.	117.	120.	132.	108.	120.	873.	6.6
9	139.	133.	118.	120.	132.	94.	65.	801.	6.0
10	135.	128.	120.	120.	131.	82.	0.	716.	5.4
11	136.	132.	108.	104.	124.	57.	0.	661.	5.0
12	132.	131.	108.	96.	116.	43.	0.	626.	4.7
13	67.	44.	45.	47.	61.	48.	0.	312.	2.3
14	131.	117.	116.	81.	94.	48.	0.	587.	4.4
15	133.	120.	120.	86.	101.	48.	0.	608.	4.6
16	141.	120.	117.	97.	127.	57.	2.	661.	5.0
17	144.	132.	132.	120.	132.	77.	24.	761.	5.7
18	144.	131.	132.	120.	132.	89.	29.	777.	5.8
19	118.	114.	114.	100.	126.	95.	72.	739.	5.6
20	57.	32.	15.	45.	95.	103.	106.	453.	3.4
21	15.	0.	0.	1.	91.	108.	117.	332.	2.5
22	8.	0.	0.	7.	84.	103.	120.	322.	2.4
23	52.	30.	8.	45.	86.	96.	119.	436.	3.3
24	70.	46.	28.	57.	92.	101.	114.	508.	3.8
TOTAL	2511.	2002.	1644.	1729.	2194.	1704.	1507.	13291.	100.0
PERCENT	19.	15.	12.	13.	17.	13.	11.	100.	0.0

NOTE: Day 1 is Monday; Hour 1 is Midnight to 1 AM

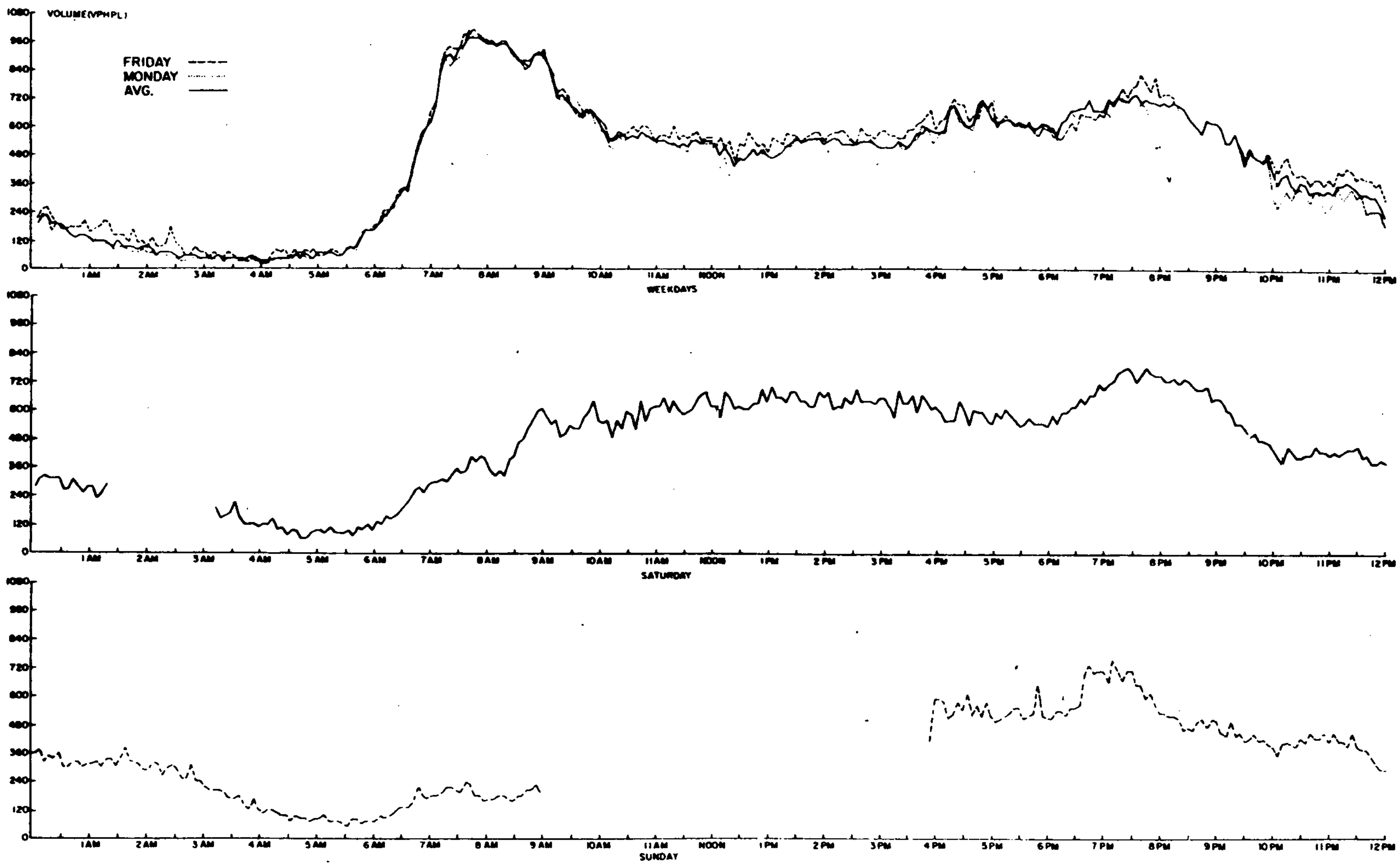
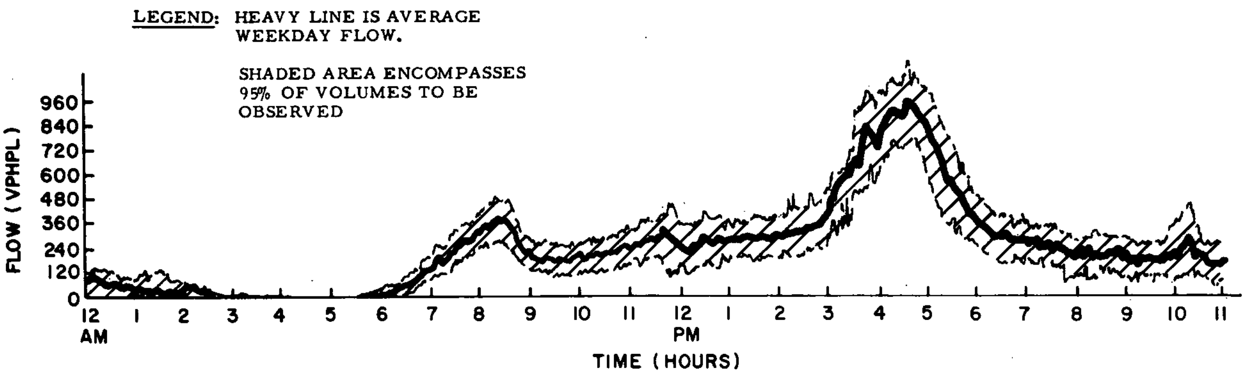
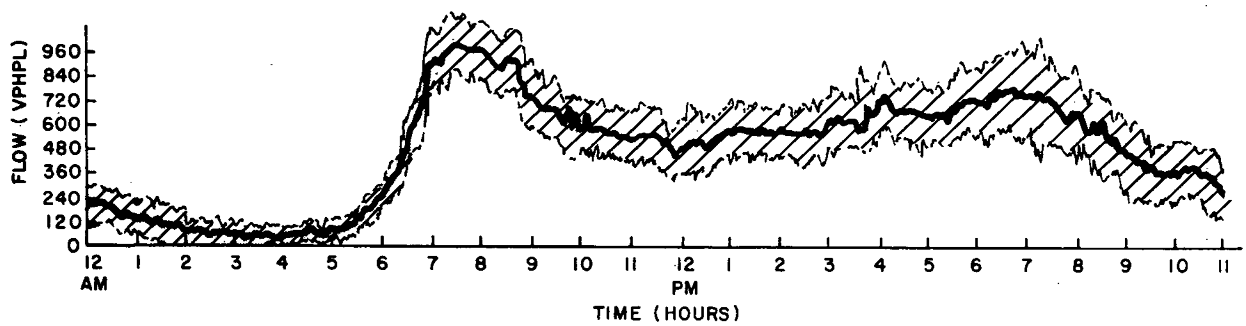


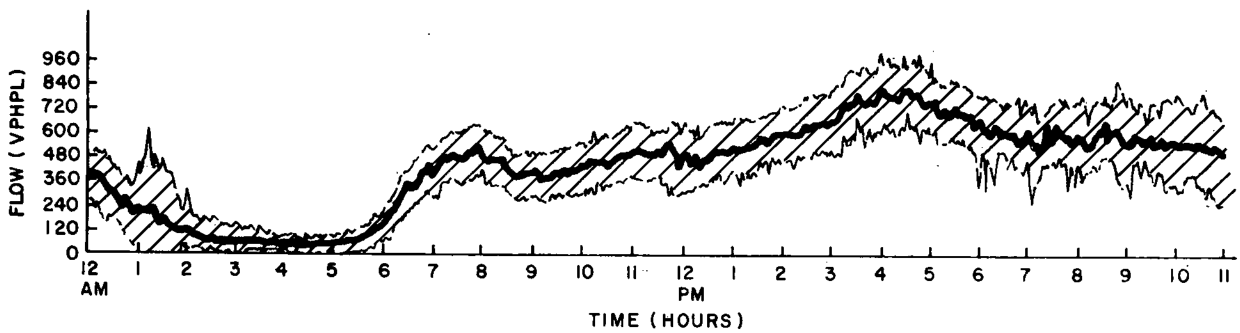
Figure E-4. Average patterns at Site 2.



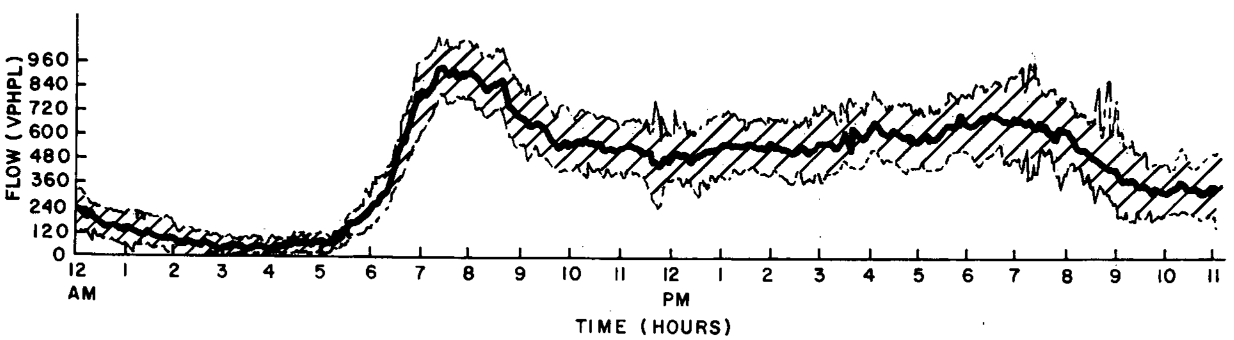
(a) SITE 1



(b) SITE 2



(c) SITE 3

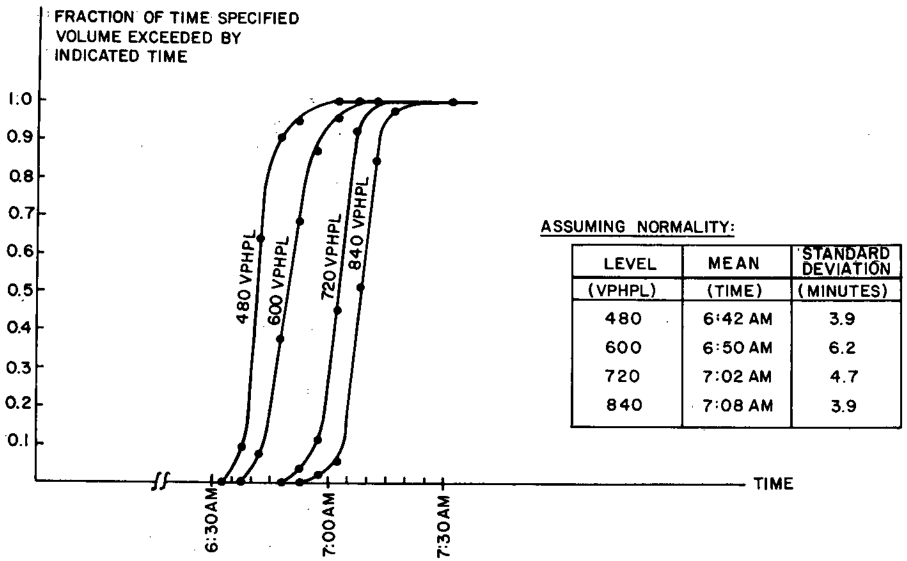


(b) SITE 4

NOTE: (1) SITES 2 AND 4 ARE ONE BLOCK APART ON SAME STREET, IN SAME DIRECTION

(2) ALL SITES ARE TWO MOVING LANES.

Figure E-5. Average weekday pattern and variation of individual flows.



LEVEL (VPHPL)	MEAN (TIME)	STANDARD DEVIATION (MINUTES)
480	6:42 AM	3.9
600	6:50 AM	6.2
720	7:02 AM	4.7
840	7:08 AM	3.9

Figure E-6. Times at which certain levels are reached (Site 2).

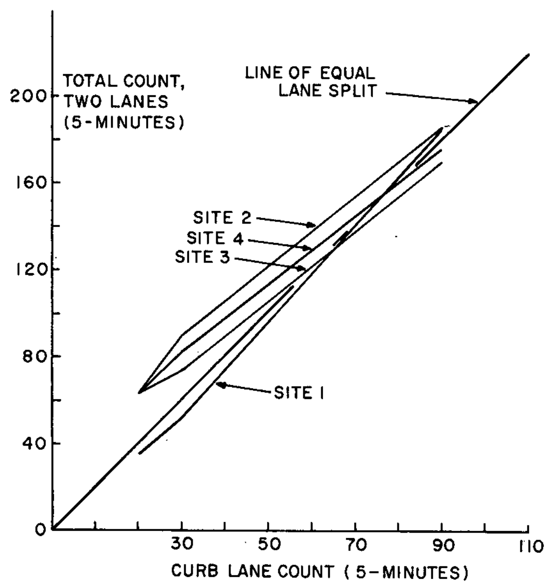


Figure E-7. Least-squares fit of total count as a function of curb lane count, weekdays only.

TABLE E-2

SUMMARY OF CORRELATIONS OF TOTAL VOLUME TO CURB-LANE VOLUMES, WEEKDAYS ONLY

Site		Correlation, Total to Curb	Correlation, Center Lane to Curb	Sample Size	B (Slope in Regime)	C (Slope in Regime)	A, the Intercept at Curb Vol = 30
1	LOW	0. 67	0. 37	2936	1. 7		52. 6
	HI	0. 98	0. 93	3416		2. 2	
2	LOW	0. 63	0. 46	1640	2. 5		89. 4
	HI	0. 95	0. 76	6311		1. 6	
3	LOW	0. 27	0. 08	1912	1. 2		74. 4
	HI	0. 88	0. 59	6124		1. 6	
4	LOW	0. 49	0. 32	1633	2. 1		83. 5
	HI	0. 92	0. 62	6400		1. 5	

NOTE:

$$Y = A + \begin{cases} B(X - 30) & X < 30 \text{ (Low)} \\ C(X - 30) & X \geq 30 \text{ (Hi)} \end{cases}$$

Y is total 5-min. count. X is curb lane 5-min. count.

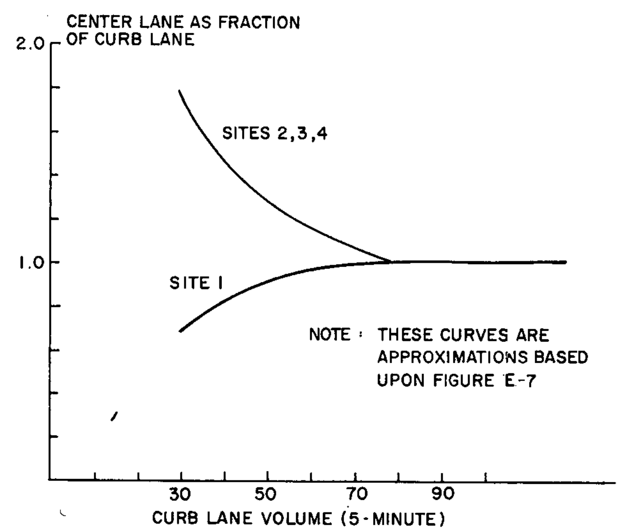


Figure E-8. Split between lanes, weekdays only, based on Figure E-7.

$$\hat{R}(1) = \frac{1}{N\text{-}2}\sum_{i=1}^{N-1} \epsilon_i\epsilon_{i+1} \qquad \text{(E-2a)}$$

$$\hat{R}(2) = \frac{1}{N\text{-}3}\sum_{i=1}^{N-2} \epsilon_i\epsilon_{i+2} \qquad \text{(E-2b)}$$

$$\hat{\sigma}^2 = \frac{1}{N\text{-}1}\sum_{i=1}^{N} \epsilon_i\epsilon_i \qquad \text{(E-2c)}$$

respectively. If there were M days available, the estimates obtained from Eq. E-2 (a, b, c) for each day could then be averaged to improve the estimate. The use of these equations assumes, however, that the quantities $R(1)$, $R(2)$, and σ^2 do not change as a function of time within the day, and this is not necessarily true.

To assure that an unnecessary (and improper) assumption is not made, the estimators

$$\hat{R}_i(1) = \beta\,\hat{R}_{i-1}(1) + (1-\beta)\,\epsilon_i\epsilon_{i-1} \qquad \text{(E-3a)}$$

$$\hat{R}_i(2) = \beta\,\hat{R}_{i-1}(2) + (1-\beta)\,\epsilon_i\epsilon_{i-2} \qquad \text{(E-3b)}$$

$$\hat{\sigma}_i^2 = \beta\,\hat{\sigma}^2_{i-1} + (1-\beta)\,\epsilon_i\epsilon_i \qquad \text{(E-3c)}$$

were used; thus, if the quantities of interest do change, it will be reflected in the estimators. The quantity β controls both the responsiveness of the estimator to change and the variance (and thus confidence) of the estimate. A β of 0.8 was used in this work.

It may be shown that most predictors of interest depend on the foregoing quantities. Indeed, if a predictor is being used such that one has past data through period (K-i), and is computing the values for period $(K+1)$ during period K (the data for which are not yet in hand), the quantity of interest is $E[\epsilon_{i+2}\,\epsilon_i]$, which indicates how much knowledge of period $(K+1)$ can be extracted from period $(K-1)$ and prior.

The quantities $\hat{\rho}_i(1)$ and $\hat{\rho}_i(2)$ may be computed from

$$\hat{\rho}_i(1) = \frac{\hat{R}_i(1)}{\sqrt{\hat{\sigma}_i^2\,\hat{\sigma}^2_{i-1}}} \qquad \text{(E-4a)}$$

$$\hat{\rho}_i(2) = \frac{R_i(2)}{\sqrt{\hat{\sigma}_i^2\,\hat{\sigma}^2_{i-2}}} \qquad \text{(E-4b)}$$

These reflect the correlations at one and two lags, respectively.

Figure E-9 is a plot of $\hat{\rho}_i(1)$ for two weekdays at Site 2. The average weekday pattern is also shown. Figure E-10 shows the plot of $\hat{\rho}_i(2)$ for the same two weekdays.

It may be concluded that the autocovariance values are indeed functions of the time of day and that they should be treated as such. Moreover, there are times when they are sufficiently high that they can be used effectively to refine the estimate of a future volume x_{i+k}.

Of course, it must be recognized that the effectiveness, although real, may not be cost effective. If $\hat{\rho}_i(2) \simeq 0.7$, the variance reduction is in the order of $(0.7)^2 \simeq 0.5$, or 50 percent. The standard deviation is thus reduced to $1/\sqrt{2}$ or 0.707 of its former value. As shown earlier (Fig. E-5), the 95 percent bounds were $\pm$ 120 vphpl for a mean of about 480 vphpl. They would now be reduced to $\pm$ 0.707(120) or $\pm$ 85 vphpl. The net "tightening" of the range is $(120-85)/12 = 3$ vehicles per lane per 5 min. One must ask if this added precision is worth the cost.

Figures E-9 and E-10 illustrate that the correlations not only are time varying but they also do not have comparable values at the same clock times from day-to-day. The indication is that, if a person decides to undertake a refined prediction of future volume levels, it should be undertaken with on-line estimation of the predictor coefficients. These coefficients vary with time. Moreover, this variation cannot be specified a priori from historic data as a simple—or even complex—function of time.

Note that this last conclusion can only be reached through a data base such as is used here. Many daily sets $\{\hat{\rho}_i(k)\}$ must be observed, some of which are represented in Figures E-9 and E-10. Also the averaging of many sets $\{\hat{\rho}_i(k)\}$ or sets $\{\hat{R}_i(k)\}$ will not produce a better estimate of the true temporal correlations. Indeed, it will obscure it, for the average will tend to zero or, at least, to small values.

DISCUSSION OF RESULTS

Within the context of the control of congested traffic flow, it is relevant to note that:

1. The daily pattern at a specific site is quite regular, so that minimal response (i.e., preplanned) signal control can be considered very seriously as a viable approach.
2. The times at which specific higher volumes are first reached is even more regular, and one can anticipate the initiation of those levels with some confidence. Although preplanned switching can thus be done with confidence, it should be recognized that the variabiltiy that does exist lessens the benefits of multiple, rapid switches.
3. Single-lane detection can provide good indications of the total approach volume, at least on two-lane approaches.
4. The differences about the average pattern are serially correlated, so that some information can be extracted from past differences in order to enhance volume predictions.

These results support the use of minimal response policies in many applications. Do they, in fact, indicate that minimal response policies may be best or optimal in some sense?

Addressing the use of on-line traffic information to aid or determine control settings, two types of systems must be distinguished: actuated equipment and system control algorithms. The time frame of the data studied (i.e., 5-min periods) is relevant only for the second type.

It should be observed that the information contained in this appendix relates only to the inherent variability of the traffic flow and to potential for refining traffic predictions. It was noted that a two-step predictor can sometimes reduce the variance by 50 percent. This can result in the deviation from the mean being reduced from 25 percent of the mean to 17.5 percent of the mean when considering the 95 percent range of volume values.

It does not follow, however, that this reduction is important. The control algorithm may just not need the

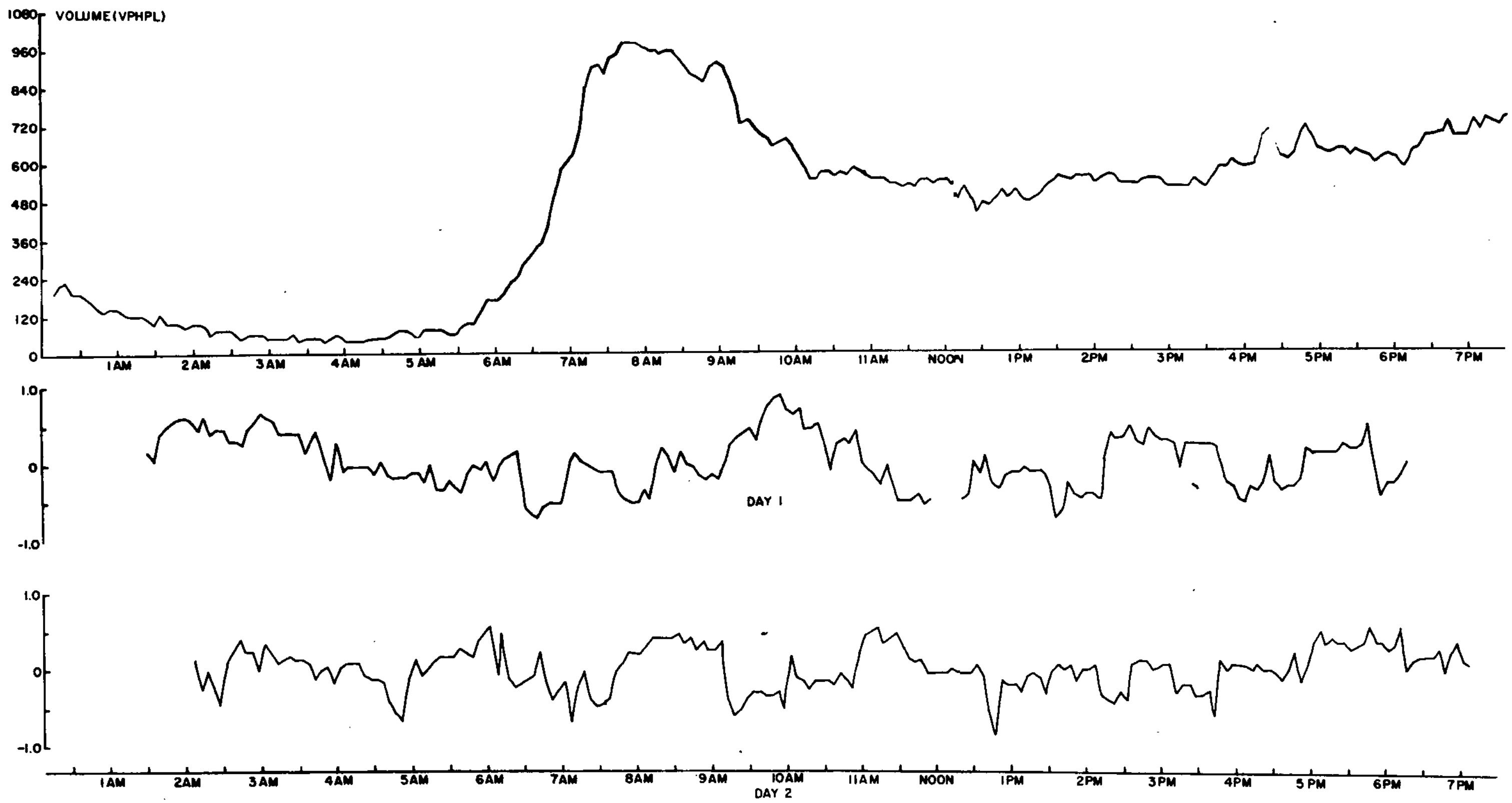


Figure E-9. $\hat{\rho}_i$ *(1) for two weekdays at Site 2, referenced to average pattern.*

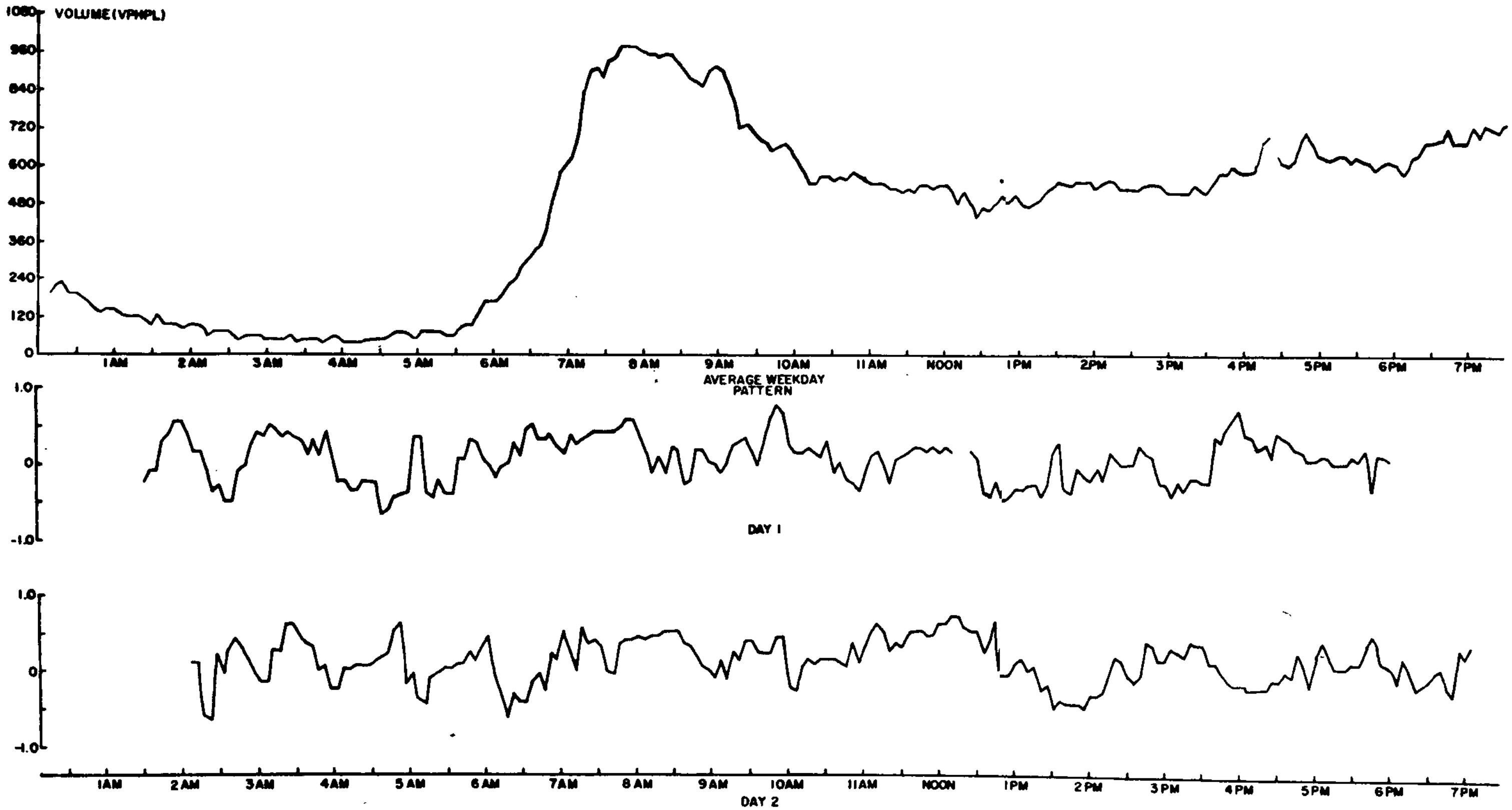


Figure E-10. $\hat{\rho}_i$ *(2) for two weekdays at Site 2, referenced to average pattern.*

volume to be predicted that precisely. Even if there is some benefit to such added precision, there is the question of cost effectiveness: the detectors cost money, as well as the maintenance of the detector system. Further, there are computation burdens on the computer system and storage requirements for intermediate results.

Because of these considerations, it is necessary to evaluate the use of such predictors as are possible only in terms of the cost effectiveness of the control function delivered, with predictor and control effectiveness integrally related. It is clear at this point that meaningful results can be obtained with only minimal response policies. It remains for advanced control projects, such as the UTCS (Urban Traffic Control System) program, to determine the cost effectiveness of various predictor-control policy combinations.

If minimal response policies are used, two problems remain: (1) potentially severe impacts due to the "rare event" that would be ignored; (2) determining the average pattern, which may change slowly over time (i.e., months or years). The first problem can be addressed by limited surveillance, but not for on-line adaptations but for a "quality control chart" type monitoring on whether the average pattern assumed is indeed applicable.

The second problem can likewise be addressed by detectors placed for surveillance, and not necessarily for on-line control purposes. Such detectors could be sampled systematically, not all necessarily on the same day, so as to update average stored patterns.

APPENDIX F

A WORKING PAPER ON A TACTICAL QUEUE-ACTUATED CONTROL

Congestion on urban street networks has become a familiar occurrence in central business districts throughout the United States and abroad. Urban street grids are commonly used for a variety of often conflicting purposes. They serve through traffic, as well as circulating traffic with destinations on or adjacent to the street grid. Land access is provided to abutting properties. Pedestrians, as well as vehicles, are serviced. Conflicting vehicular and pedestrian movements are interfaced at grade, with no, or minimal, separation.

This system of conflicting movements, modes, and purposes is controlled by a variety of common traffic control devices and techniques, principally signals. The emphasis of current traffic control schemes is mainly directed towards light traffic flows in which flexible coordination of signals will enable traffic platoons to progress along "green waves," and towards conditions of moderate congestion in which attention is directed to maximizing the capacity of "bottlenecks."

When volumes become excessively high, most of the present concepts of optimization appear ineffective or invalid (*F-1*). As traffic flows increase, the system (or points within the system) is subject to breakdown or congestion. The occurrence of saturation begins principally at intersections and points of maximum conflict and, if unchecked, spreads to infect other parts of the system.

In saturated conditions, congestion is unavoidable, and, therefore, the control policy should be aimed at postponing the onset and/or the severity of the secondary congestion, which is caused by the blockage of intersections that are not the originators of congestion.

This appendix describes a new, advanced control concept that is well suited under extremely heavy flow conditions—that is, under oversaturation.

Oversaturation represents a condition that occurs when queues fill entire blocks and interfere with the performance of adjacent upstream intersections. Sometimes cross-street traffic is blocked by extended queues. This kind of traffic performance may be experienced at single points, along major arterials, or throughout entire subsections of a network. If one observes such a system when it is operating in an oversaturated mode, the problem appears to be areawide in nature. If, however, one were able to observe the onset of the oversaturated state, one would see that instead of having an areawide problem, there exists one, or perhaps a few, critical locations at which congestion first begins to develop. Thus, the problem of oversaturation begins as a localized problem. If allowed to develop, its effects may become far more widespread, and may appear to be areawide in nature.

It is this initial, local nature of the potential for causing oversaturation that appears to be the key to both the characterization of oversaturation and the potential for control to delay its onset, or to avoid its appearance completely.

In the text and appendixes of this report, the various stages of traffic congestion were clearly defined using a queue formation mechanism; measures of oversaturation were devised and field validated, and extensive evaluation of various existing control measures were made to determine their effectiveness under oversaturated flow conditions. The following are conclusions based on the evaluation of various existing control measures that are well known to the current traffic engineering practitioners, as well as the evaluation of advanced concepts that are not fully tested in the field:

1. Some of the control measures evaluated (such as simultaneous offset plans) that are of strategic nature are very effective in delaying and/or eliminating oversaturation at critical intersections. They, however, tend to degrade the performance of adjacent intersections in the system.
2. Control measures of a strategic nature, such as smooth flow (*F-2*) theory and reverse progression, generally reduce queues at the critical intersection more than necessary to prevent oversaturation, and reduce the productivity of a critical intersection.
3. Control measures that are of tactical nature, such as modifying cycle length and/or split ratio, do not act positively to prevent the blockage of intersections, which is the prime concern during the period of oversaturation.
4. A new control concept is desired, which not only ensures the prevention of intersection blockage as much as possible but also minimizes the degrading of other intersections in the system.
5. A new control concept should be capable of maintaining high productivity at a critical intersection. Productivity is more important than the quality of operation during a period of oversaturation.

THE QUEUE-ACTUATED POLICY

It was apparent from the analysis performed that none of the control methods really act to completely prevent oversaturation or intersection blockage.

Prevention of intersection blockage becomes more important with arterials or networks where traffic flow is heavy in all directions. In such cases, intersection blockage not only increases the delay through the system, but also affects more segments of that system.

"Queue-actuated signal control" was developed to ensure, as much as possible, the prevention of intersection blockage. It was also intended to fully utilize all the available green time by creating continuous demand (queues) at all approaches. It is a control policy where an approach receives green automatically when the queues on that approach become equal to, or greater than, some predetermined length.

When the queue at one approach becomes equal to, or greater than, a given length ($Q_{\max}$), that approach will receive green regardless of the conditions on the conflicting approaches. Thus, how much green one approach would receive during a given time interval depends on the link length and the number of lanes, as well as on the flow rates of both approaches. This control responds to queues in a way similar to the manner in which the current traffic-actuated control responds to demand.

It is obvious that, when one approach has a relatively low flow rate, drivers on that approach would suffer long delays because the queues on that approach would take a longer time to reach the predetermined length. When volumes become extremely heavy on both approaches, the effective cycle length will become very short, because a shorter time is required for the queue to reach a given length. This problem can be somewhat reduced by selecting the proper maximum queue length allowed ($Q_{\max}$). As will be shown later, the effective cycle length is only a function of $Q_{\max}$, given the flow rates.

To avoid too long or too short cycle lengths, one may impose maximum or minimum values of green time, which will determine the upper and lower bounds of cycle length as well as of green time.

One of the advantages of this control policy is that it is a more positive way to prevent intersection blockage than any of the other methods reviewed thus far.

Isolated, as well as coordinated, intersections are examined under this control policy. Isolated intersections are tested primarily to acquire a basic understanding of the policy, which would be useful in applying this control approach to a coordinated system of signals.

DELAY MODEL FOR ISOLATED INTERSECTIONS

An analytical model was developed to study the characteristics of "queue-actuated control" at isolated intersections. The model was validated by use of the UTCS (Urban Traffic Control System) simulator. The analytical study includes construction of a delay model, minimization of the average delay through controlled intersections, and estimation of expected cycle length under queue-actuated control.

Basic Cases

Newell (*F-3*), and Sagi and Campbell (*F-4*), expressed the delay at a signalized intersection under saturated flows in the manner shown in Figure F-1 (A).

When the concept of queue-actuated signal control is applied to a signalized intersection, the delay at both approaches during the ith cycle can be expressed as in Figure F-1 (B), where:

Q_{m1}, Q_{m2} = maximum queue length allowed at Approaches 1 and 2, respectively;
Q_{1i} = queue length at Approach 1 at the end of green;
Q_{2i} = queue length at Approach 2 at the end of green;
X_1, X_2 = flow rates at Approaches 1 and 2, respectively, and X_1 and X_2 are assumed to have a Poisson distribution with expected values $E[X_1] = \lambda_1$ and $E[X_2] = \lambda_2$;
t_{1i} = effective red phase at Approach 1 during i^{th} cycle;
t_{2i} = effective green phase at Approach 1 during i^{th} cycle;
t_i = effective cycle length during i^{th} cycle ($t_i = t_{1i} + t_{2i}$); and
S = saturation flow (service rate).

S is assumed to be a constant, equal to or greater than X_1 and X_2. When $S < X_i$, the queues become infinite, and the system will break down because $(S - X_i)$ is the net discharge rate from an intersection.

In addition, it is assumed that loss time due to start-up delay and the amber phase is constant. Loss time is not included in the delay diagram, but is taken up separately.

Because of statistical fluctuations of flows, the following four different situations arise even under saturated flow conditions (see Table F-1).

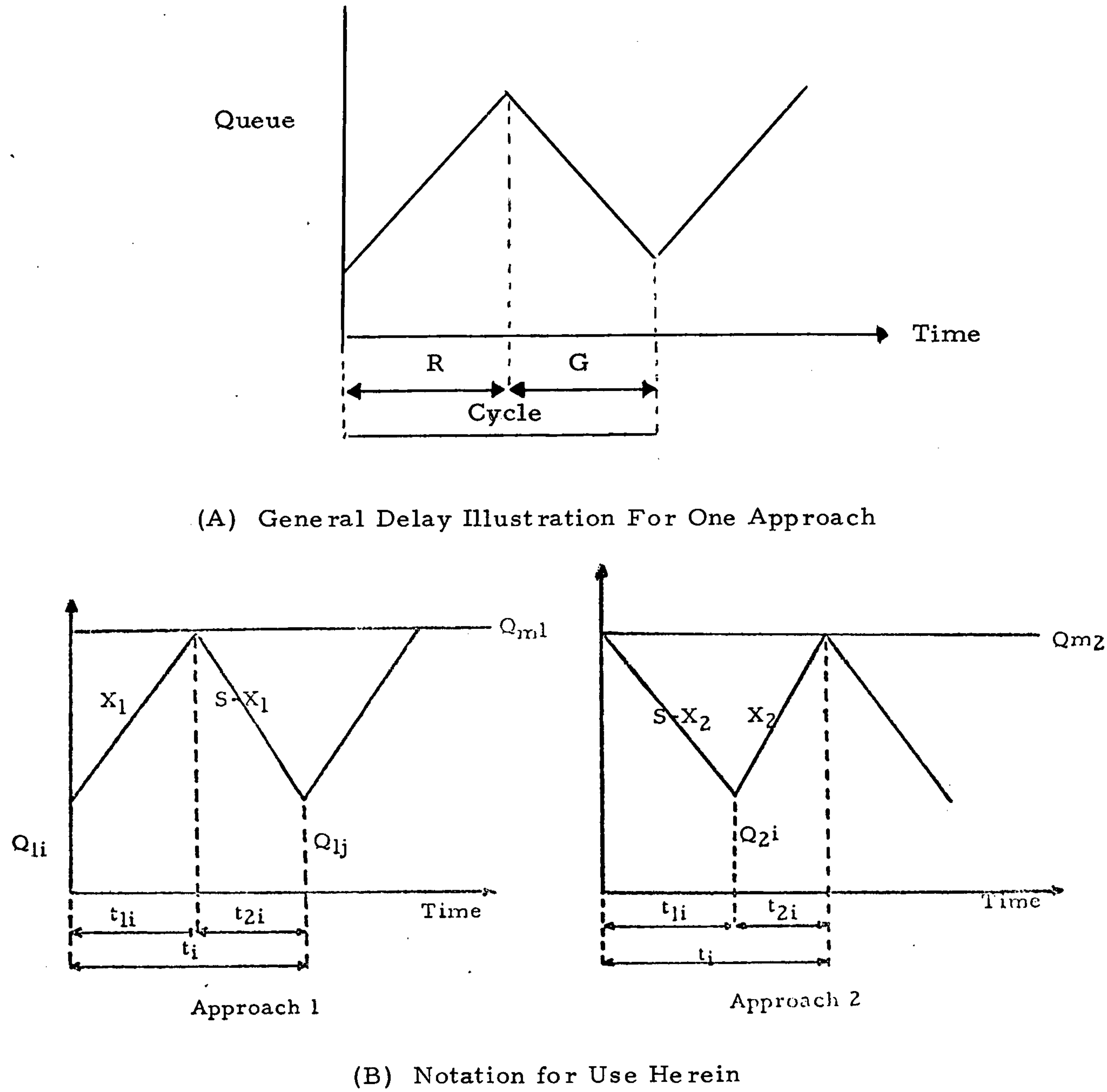


Figure F-1. Delay diagrams and notation.

Case 1: Residue Queues on Both Phases

Switching from one phase to another occurs when queues at one approach reach a predetermined length (Q_{m1} or Q_{m2}). Queues at the end of green (Q_{1i} or Q_{2i}) are greater than zero or become zero at the moment of the termination of green. Note the defined locations of Q_{1i}, Q_{2i}, and Q_{ij}.

Case 2: Approach 1 Empties Before Q_{m2} *is Reached*

Queues at Approach 1 are completely cleared before queues at Approach 2 reach Q_{m2} and, thus, before the green phase at Approach 1 is terminated.

Case 3: Approach 2 Empties Before Q_{m1} *is Reached*

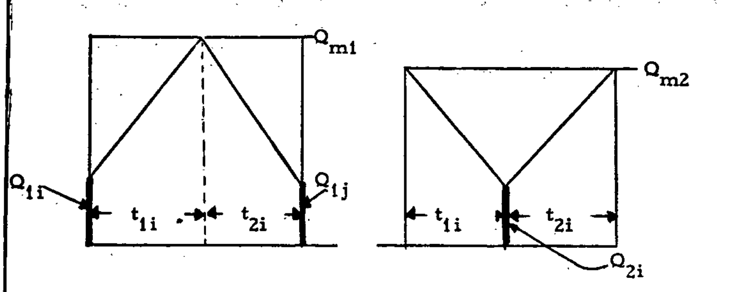
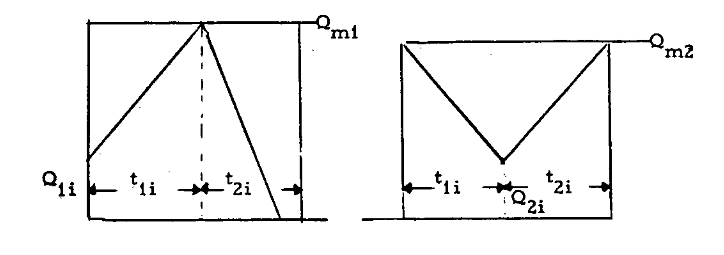
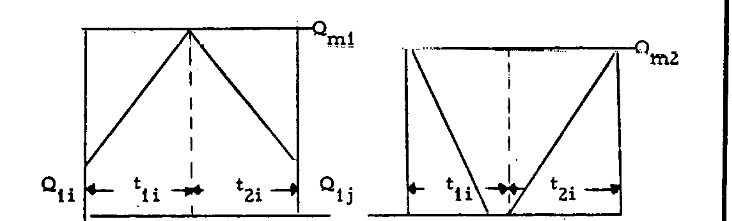
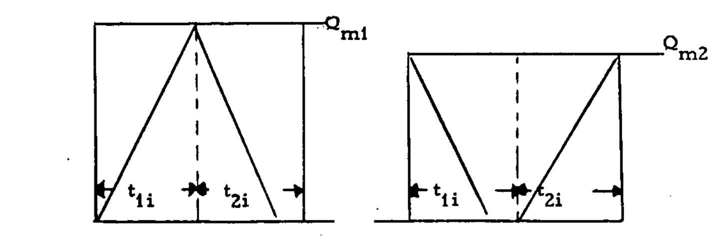
Queues at Approach 2 are completely cleared before queues at Approach 1 reach Q_{m1} and thus before the green phase at Approach 2 is terminated.

Case 4: Both Approaches Empty Before Other Q_m *is Reached*

This is a combination of cases 2 and 3. Since this case would be likely under light flows, Q_{1i} is assumed to be zero.

TABLE F-1

FOUR CASES OF QUEUE PATTERNS

CASE	APPROACH 1 / APPROACH 2	NOTES
1		$Q_{1i} \geq 0$ $Q_{2i} \geq 0$ $Q_{ij} \geq 0$
2		$Q_{1i} \geq 0$ $Q_{2i} \geq 0$
3		$Q_{1i} \geq 0$ $Q_{2i} \geq 0$
4		$Q_{1i} = 0$

Probability of Each Case

Let P_1 = probability of queues being cleared at Approach 1 before green is terminated, and P_2 = probability of queues being cleared at Approach 2 before green is terminated. Then, it follows that for case 1, probability $=(1 - P_1)\ (1 - P_2)$; for case 2, probability $= P_1\ (1 - P_2)$; for case 3, probability $=$ (1) $- P_1$) (P_2); and for case 4, probability $= (P_1)\ (P_2)$.

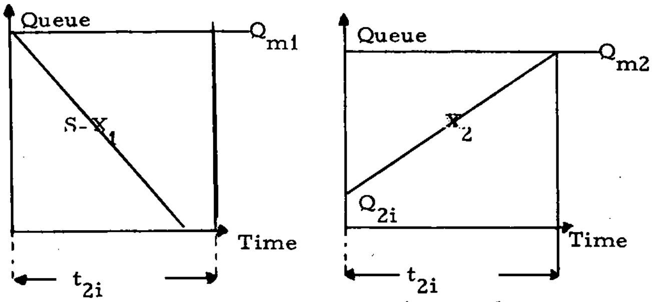
Under heavy flow conditions, cases 2, 3, and 4 are not likely to occur because:

1. Due to the very nature of the control policy, every vehicle is expected to stop at least once, and there will not be a period of green without queue.

2. Cases 2 through 4 will likely occur towards the end of saturated period, and the occurrence of these cases is the very indication that this control measure is not effective any longer.

Evaluation of P_1

From the sketch



Approach 1 Approach 2

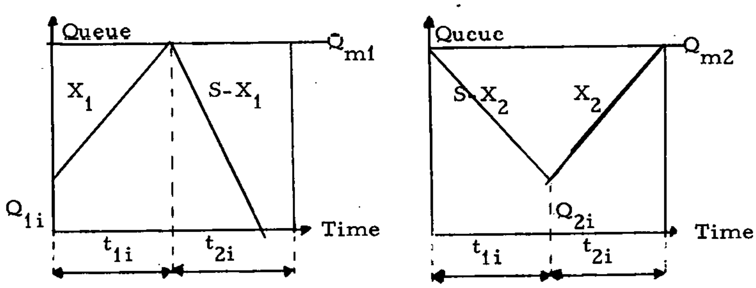
it follows that

$$P_1 = P\{(S - X_1)\, t_{2i} > Q_{m1}\} = P\left\{(S - X_1) > \frac{Q_{m1}}{t_{2i}}\right\}$$

Since $t_{2i} = (Q_{m2} - Q_{2i})/X_2$, as can be seen from Approach 2,

$$P_1 = P\left\{(S - X_1) > \frac{Q_{m1}\, X_2}{Q_{m2} - Q_{2i}}\right\} \qquad \text{(F-1)}$$

However, when one looks at the previous phase interval as well (i.e., when Approach 2 had the green), the delay diagram is as follows:



Approach 1 Approach 2

From the foregoing diagrams, $t_{1i} = (Q_{m1} - Q_{1i})/X_1$ and $t_{2i} = (t_{1i})\,(S - X_2)/X_2$. Then,

$$P_1 = P\left\{(S - X_1) > \frac{Q_{m1}}{t_{2i}}\right\}$$

$$= P\left\{(S - X_1) > \frac{(Q_{m1})\,(X_2)}{t_{1i}\,(S - X_2)}\right\}$$

Since $t_{1i} = (Q_{m1} - Q_{1i})/X_1$, as was just shown,

$$P_1 = P\left\{(S - X_1) > \frac{Q_{m1}\,(X_1)\,(X_2)}{(Q_{m1} - Q_{1i})\,(S - X_2)}\right\} \qquad \text{(F-2)}$$

While Eq. F-1 requires the estimation of Q_{2i}, Eq. F-2 requires the estimation of Q_{1i}. However, as seen from the two equations, $S - X_1$ is much less sensitive to any possible errors in the estimation of Q_{1i} than of Q_{2i}. Therefore, Eq. F-2 will be used. Note that the terms

$$(S - X_1) > \frac{(Q_{m1})\,(X_1)\,(X_2)}{(Q_{m1} - Q_{1i})\,(S - X_2)}$$

can be written as

$$X_1 < S\left\{\frac{(Q_{m1} - Q_{1i})\,(S - X_2)}{(Q_{m1} - Q_{1i})\,(S - X_2) + Q_{m1}\, X_2}\right\}$$

Let

$$A = \left\{\frac{(Q_{m1} - Q_{1i})\,(S - X_2)}{(Q_{m1} - Q_{1i})\,(S - X_2) + Q_{m1}\, X_2}\right\}$$

Then, the original expression can be written as $P_1 = P\,(X_1 < SA)$.

Since A is a function of X_2, which is also a random variable,

$$P_1 = P\,(X_1 < S\,A) = \sum_{\text{all } X_2} P\,(X_1 < S\,A/X_2)\; P\,(X_2)$$

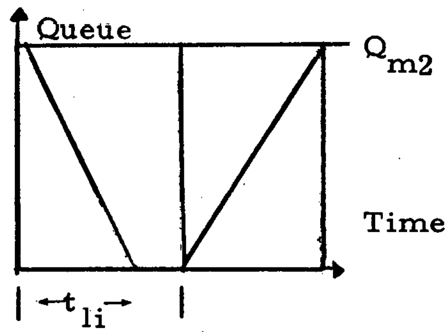
$$= \sum_{X_2=0}^{\propto} \left(\sum_{K=0}^{S\,A} P\,(X_1 = K/X_2)\right) P\,(X_2)$$

Thus,

$$P_1 = \sum_{X_2=0}^{\propto} \left(\sum_{K=0}^{S\,A} P\,(X_1 = K/X_2)\right) P\,(X_2) \qquad \text{(F-3)}$$

Evaluation of P_2

The development of the relation for P_2 follows along similar lines to the development of P_1. Note that, as illustrated,



$$P_2 = P\,\{(S - X_2)\, t_{1i} > Q_{m2}\}$$

$$P_2 = P\left\{(S - X_2) > \frac{Q_{m2}}{t_{1i}}\right\}$$

Because $t_{1i} = (Q_{m1} - Q_{1i})/X_1$ by consideration of the concurrent Approach 1 pattern, it follows that

$$(S - X_2) > \frac{Q_{m2}}{(Q_{m1} - Q_{1i})/X_1} = \frac{Q_{m2}\, X_1}{Q_{m1} - Q_{1i}}$$

Defining

$$B = \frac{Q_{m2}\, X_1}{Q_{m1} - Q_{1i}}$$

one can write

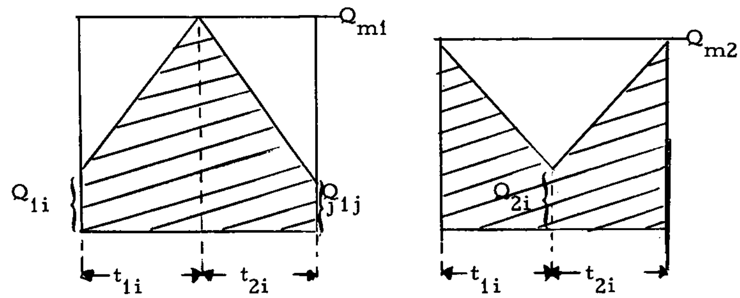
$$P_2 = P\,(S - X_2 > B) = P\,(X_2 < S - B)$$

$$P_2 = P\,(X_2 < S - B) = \sum_{\text{all } X_1} P\,(X_2 < (S - B)/X_1)\; P\,(X_1)$$

$$P_2 = \sum_{X_1=0}^{\propto} \sum_{K=0}^{(S-B)} P\,(X_2 = K/X_1)\; P\,(X_1) \qquad \text{(F-4)}$$

Delay Due to Queue

The total delay during a given cycle because of the existence of queues is the summation of the shaded areas of the delay diagram:



Delay Due to Loss Time

The total delay due to loss time, such as start-up delay and the amber phase not used by traffic, would be, for the i^{th} cycle,

$$D_L = \left(\frac{L}{t_i}\right)\left(\frac{1}{2}\right)(t_i\,(X_1 + X_2)) = \frac{L}{2}(X_1 + X_2) \qquad \text{(F-5)}$$

where L, the loss time per cycle, is the sum of all amber not used as green and all start-up delays.

Total Delay and Effective Cycle Length

The total delay experienced by traffic on both approaches of an intersection would be

$$D_T = (\text{Delay due to Queue}) + D_L \qquad \text{(F-6)}$$

This relation may be used in conjunction with the "area under the curve," as shown in each case in Table F-1, to compute the total delay in a cycle, as shown in Table F-2. In addition, the values of t_{1i}, t_{2i}, and the sum of these—denoted t_i—are also shown. Note that t_i is the cycle length and is a random variable. For simplicity, it was assumed that $Q_{m1} = Q_{m2}$.

Evaluation of Q_{1i}

Once Q_{1i} (the queue length at the end of green for Approach 1) is known, delay can be evaluated. Since Q_{1i} can be estimated only by a complete accounting of inputs and outputs at an approach, and the probability density function of Q_{1i} is unknown, the average Q_{1i} over a given control period is used, instead of the instantaneous Q_{1i}, for estimation of the delay.

In Newell's derivation of his delay model (*F-3*), for a fixed time signal,

$$\overline{Q}_{1i} = \frac{I\lambda_I}{2S}\left(\frac{G}{G+R} - \frac{\lambda_I}{S}\right)^{-1}$$

where:

$\overline{Q}_{1i}$ = average of Q_{1i};
λ_1 = average flow at Approach 1;

$$I = \frac{\text{VAR}(D)}{\overline{D}} + \frac{\text{VAR}(A)}{\overline{A}};$$

$\overline{A}$ = average arrival during a cycle;
$\overline{D}$ = average Departure during a cycle;

G = green phase; and
R = red phase.

However, under saturation conditions,

$$\text{VAR}(D) \rightarrow 0$$

$$\frac{\text{VAR}(A)}{\overline{A}} = 1 \text{ for Poisson arrival}$$

Thus,

$$I = \frac{\text{VAR}(D)}{\overline{D}} + \frac{\text{VAR}(A)}{\overline{A}} = 1$$

and

$$\overline{Q}_{1i} = \frac{\lambda_1}{2S}\left(\frac{t_{2i}}{t_{1i} + t_{2i}} - \frac{\lambda_1}{S}\right)^{-1}$$

$$= \frac{\lambda_1}{2S}\left(\frac{\left(\frac{S-\lambda_2}{\lambda_2}\right)}{\left(\frac{S}{\lambda_2}\right)} - \frac{\lambda_1}{S}\right)^{-1}$$

$$= \frac{\lambda_1}{2S}\left(\frac{S-\lambda_2}{S} - \frac{\lambda_1}{S}\right)^{-1}$$

$$= \frac{\lambda_1}{2S}\left(\frac{S-(\lambda_1+\lambda_2)}{S}\right)^{-1}$$

$$\overline{Q}_{1i} = \frac{\lambda_1}{2\,(S-(\lambda_1+\lambda_2))} \qquad \text{(F-7)}$$

Average Delay/Vehicle

The average delay/vehicle is expressed as:

$$\text{Average Delay/Vehicle} = E\left[\frac{D_T}{A}\right]$$

where D_T is the total delay and A is the total number of arrivals during a given time period.

Since D_T and A are positively correlated, the average delay/vehicle will be approximated as:

$$\text{Average Delay/Vehicle} = E\,[D_T]/E\,[A]$$

where $E\,[D_T]$ is the expected value of D_T and $E\,[A]$ is the expected value of A. Table F-3 shows $E\,[D_T]$ and $E\,[A]$ for each case.

TABLE F-2

TOTAL DELAY AND EFFECTIVE CYCLE LENGTH

Case	Total Delay	Effective Cycle Length, Green and Red Phases.
1	$D_T = (\frac{1}{2})(\frac{S}{X_2})(\frac{Q_m - Q_{1i}}{X_1})\left[2(Q_m + Q_{1i}) - \frac{(S^2 - SX_1 - SX_2)}{X_1 X_2}(Q_m - Q_{1i})\right] + \frac{L}{2}(X_1 + X_2)$	$t_{1i} = \frac{(Q_m - Q_{1i})}{X_1}$ $t_{2i} = t_{1i}\,(\frac{S - X_2}{X_2})$ $t_i = t_{1i}\,(\frac{S}{X_2})$
2	$D_T = (\frac{1}{2})(\frac{S}{X_2})(\frac{Q_m - Q_{1i}}{X_1})\left[2Q_m - \frac{(S - X_2)(Q_m - Q_{1i})}{X_1} + \frac{(Q_m + Q_{1i})}{S} X_2 + \frac{Q_m^2 X_1 X_2}{S(S - X_1)(Q_m - Q_{1i})}\right] + \frac{L(X_1 + X_2)}{2}$	Same as Case 1
3	$D_T = \frac{(Q_m^2 - Q_{1i}^2)}{2X_1} + \frac{2Q_m^2 X_2 - Q_m^2 (S - X_1)}{2X_2^2} + \frac{Q_m^2}{2(S - X_2)} + \frac{Q_m^2}{2X_2} + \frac{L(X_1 + X_2)}{2}$	$t_{1i} = \frac{(Q_m - Q_{1i})}{X_1}$ $t_{2i} = \frac{Q_m}{X_2}$ $t_i = \frac{(Q_m - Q_{1i})}{X_1} + \frac{Q_m}{X_2}$
4	$D_T = \frac{Q_m^2}{2X_1} + \frac{Q_m^2}{2(S - X_1)} + \frac{Q_m^2}{2(S - X_2)} + \frac{Q_m^2}{2X_2} + \frac{L(X_1 + X_2)}{2}$	$t_{1i} = \frac{Q_m}{X_1}$ $t_{2i} = \frac{Q_m}{X_2}$ $t_i = \frac{Q_m}{X_1} + \frac{Q_m}{X_2}$

Minimization of Delay with Respect to Q_m

Let

$$D = \frac{\text{Total Delay during } i^{\text{th}} \text{ Cycle}}{\text{Effective Cycle Length}} = \frac{\text{Delay}}{\text{Unit Time}}$$

Then, the Q_m that minimizes the delay can be found by

$$\frac{dE[D]}{dQ_m} = 0$$

However, mathematical derivation of the Q_m that minimizes the delay is trivial for cases 3 and 4 (i.e., $Q_m = 0$) and, therefore, the following two alternative solutions are suggested.

Alternative 1

The average delay/vehicle for all cases can be expressed as:

$$\text{Average Delay/Vehicle} = \sum_{i=1}^{4} P(i)\; d(i)$$

where $P(i)$ is the probability of case i, and $d(i)$ is the average delay/vehicle for case i. Then, by plotting the average delay/vehicle for all cases for all possible ranges of Q_m, the Q_m that minimizes the total delay can be found graphically (point X in Figure F-2).

TABLE F-3

E[D_T] AND E[A] FOR EACH CASE

Case 1	$E[D_T] = \frac{S(Q_m - \overline{Q}_{1i})}{2} [2(Q_m + \overline{Q}_{1i}) E[\frac{1}{X_1}] E[\frac{1}{X_2}] - (S^2 E[\frac{1}{X_1^2}] E[\frac{1}{X_2^2}] - SE[\frac{1}{X_1}] E[\frac{1}{X_2^2}] - SE[\frac{1}{X_1^2}] E[\frac{1}{X_2}]) (Q_m - \overline{Q}_{1i})] + \frac{L(\lambda_1 + \lambda_2)}{2}$ $E[A] = S(Q_m - \overline{Q}_{1i}) (E[\frac{1}{X_1}] + E[\frac{1}{X_2}])$
Case 2	$E[D_T] = \frac{S(Q_m - \overline{Q}_{1i})}{2} [2Q_m E[\frac{1}{X_1}] E[\frac{1}{X_2}] - (Q_m - \overline{Q}_{1i}) E[\frac{1}{X_1^2}] E[\frac{S-X_2}{X_2}] + \frac{Q_m + \overline{Q}_{1i}}{S} E[\frac{1}{X_1}] + \frac{Q_m^2}{S(Q_m - \overline{Q}_{1i})} E[\frac{1}{S-X_1}]] + \frac{L(\lambda_1 + \lambda_2)}{2}$ $E[A] = S(Q_m - \overline{Q}_{1i}) (E[\frac{1}{X_1}] + E[\frac{1}{X_2}])$
Case 3	$E[D_T] = \frac{Q_m^2}{2} (E[\frac{1}{X_1}] + 2E[\frac{1}{X_2}] + (\lambda_2 - S) E[\frac{1}{X_2^2}] + E[\frac{1}{S-X_2}]) - \frac{\overline{Q}_{1i}^2}{2} E[\frac{1}{X_1}] + \frac{L(\lambda_1 + \lambda_2)}{2}$ $E[A] = (2 + \lambda_2 E[\frac{1}{X_1}] + \lambda_1 E[\frac{1}{X_2}]) Q_m - (1 + \lambda_2 E[\frac{1}{X_1}]) \overline{Q}_{1i}$
Case 4	$E[D_T] = \frac{Q_m^2}{2} (E[\frac{1}{X_1}] + E[\frac{1}{X_2}] + E[\frac{1}{S-X_1}] + E[\frac{1}{S-X_2}]) + \frac{L(\lambda_1 + \lambda_2)}{2}$ $E[A] = (2 + \lambda_1 E[\frac{1}{X_2}] + \lambda_2 E[\frac{1}{X_1}]) Q_m$

Alternative 2

An alternative to the foregoing graphical solution is to determine the optimum Q_m based on case 1 only, because case 1 is the most likely under heavy flows. Because the delay curve is generally rather flat near the optimum, this should suffice. As shown by Webster (*F-5*), the delay for cycle length within the range ¾ to 1½ times the optimum value is only slightly higher than that given by the optimum cycle.

From Table F-2, note that D_T (total delay for case 1) is given by:

$$D_T = \frac{1}{2}\left(\frac{S}{X_2}\right) t_{1i}\left[2(Q_m + Q_{1i}) - \frac{S^2 - SX_1 - SX_2}{X_1 X_2}(Q_m - Q_{1i})\right] + \frac{L(X_1 + X_2)}{2}$$

Let

$$D = (\text{Total Delay/Effective Cycle Length}) = (\text{Delay/Unit Time})$$

Because $t_{2i} = t_{1i}\ [(S - X_2)/X_2]$ and t_i, the effective cycle length, equals (t_{1i+2i}), it follows that $t_i = t_{1i}\ (5/X_2)$. Then, dividing by t_i, it follows that

$$D = \frac{1}{2}\left(2(Q_m + Q_{1i}) - \frac{S^2 - SX_1 - SX_2}{X_1 X_2}(Q_m - Q_{1i})\right) + \frac{L\,(X_1 + X_2)\, X_2 X_1}{2S\,(Q_m - Q_{1i})}$$

Now taking an expectation and noting that $Q_{1i} \simeq \overline{Q}_{1i}$ and that X_1 and X_2 are independent,

$$E\,[D] = \frac{1}{2}\left[2\,(Q_m + \overline{Q}_{1i}) - (Q_m - \overline{Q}_{1i})\left(S^2\, E\left[\frac{1}{X_1}\right]\right.\right.$$

$$E\left[\frac{1}{X_2}\right] - SE\left[\frac{1}{X_1}\right] - SE\left[\frac{1}{X_2}\right]\Big)\Big]$$
$$+ \frac{L}{2S\,(Q_m - \bar{Q}_{1i})}\,(\lambda_1\, E\,[X_2^{\,2}] + \lambda_2\, E\,[X_1^{\,2}])$$

Let

$$W = \left\{S^2 E\left[\frac{1}{X_1}\right] E\left[\frac{1}{X_2}\right] - SE\left[\frac{1}{X_2}\right] - SE\left[\frac{1}{X_1}\right]\right\}$$

and

$$B = \{\lambda_1\, E\,[X_2^{\,2}] + \lambda_2\, E\,[X_1^{\,2}]\}$$

Then

$$E\,[D] = \frac{1}{2}[2Q_m + 2\bar{Q}_{1i} - (Q_m - \bar{Q}_{1i})\; W] + \frac{LB}{2S\,(Q_m - \bar{Q}_{1i})}$$
$$= \left(Q_m - \frac{WQ_m}{2}\right) + \left(\bar{Q}_{1i} + \frac{W\bar{Q}_{1i}}{2}\right) + \frac{LB}{2S\,(Q_m - \bar{Q}_{1i})}$$
$$E\,[D] = \left(1 - \frac{W}{2}\right) Q_m + \left(1 + \frac{W}{2}\right) \bar{Q}_{1i} + \frac{LB}{2S\,(Q_m - \bar{Q}_{1i})} \qquad \text{(F-8)}$$

$$\frac{dE\,[D]}{dQ_m} = \left(1 - \frac{W}{2}\right) - \frac{LB}{2S\,(Q_m - \bar{Q}_{1i})^2} = 0$$
$$\frac{LB}{2S\,(Q_m - \bar{Q}_{1i})^2} = \left(1 - \frac{W}{2}\right)$$
$$Q_m = + \bar{Q}_{1i} \pm \sqrt{\frac{LB}{(2 - W)\, S}} \qquad \text{(F-9)}$$

The form of the solution is shown in Figure F-3, together with the two components of the total delay.

However, the Q_m that minimizes the expected delay exists only if $(2 - W) \geq 0$.

$$2 - W = 2 - S^2 E\left[\frac{1}{X_1}\right] E\left[\frac{1}{X_2}\right] + SE\left[\frac{1}{X_1}\right] + SE\left[\frac{1}{X_2}\right] \geq 0$$

$$2 + SE\left[\frac{1}{X_2}\right] - S\left(SE\left[\frac{1}{X_2}\right] - 1\right) E\left[\frac{1}{X_1}\right] \geq 0$$

$$E\left[\frac{1}{X_1}\right] \leq \frac{2 + SE\left[\frac{1}{X_2}\right]}{S\left(SE\left[\frac{1}{X_2}\right] - 1\right)}$$

Thus, the Q_m that minimizes the expected delay exists only if

$$E\left[\frac{1}{X_1}\right] \leq \frac{2 + SE\left[\frac{1}{X_2}\right]}{\left(SE\left[\frac{1}{X_2}\right] - 1\right) S} \qquad \text{(F-10)}$$

The region that this relation describes is shown in Figure F-4. When $(2 - W) < 0$, the relationship between delay and Q_m is as shown in Figure F-5.

When $Q_m = [(2 + W)/(W - 2)]\,\bar{Q}_{1i}$, $E\,[D]$ due to queue becomes zero, and the optimum Q_m should be $[(2 + W)/W - 2)]\,\bar{Q}_{1i}$. However, no delay due to queues implies that there are no queues, and this invalidates the whole delay diagram. Therefore, the area defined as "area

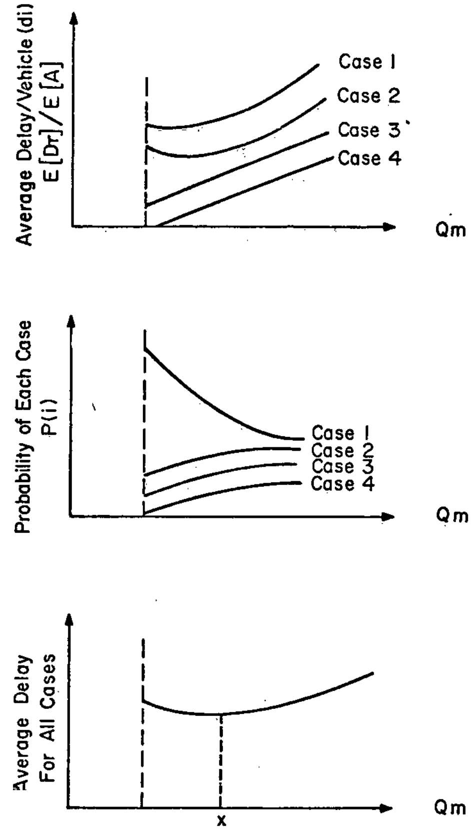


Figure F-2. Illustration of the graphical solution of optimum Q_m.

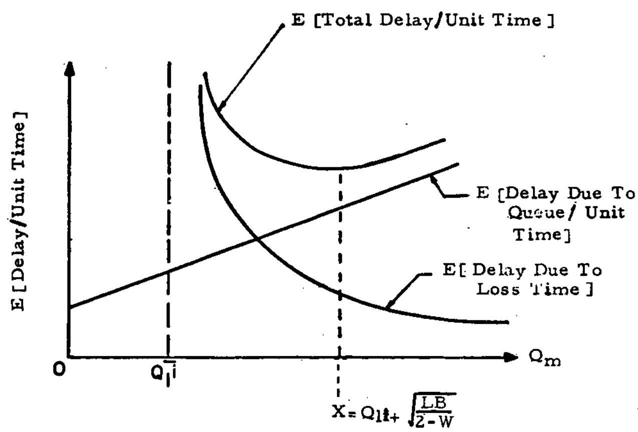


Figure F-3. E[delay/unit time] vs. Q_m.

feasible for minimization of delay" is the only area where this control concept should be applied.

Sample Problem

Consider the flows shown in Figure F-6. Also, a queue-activated control is desired. Then determine the gueue, Q_m, that should be used as a switching point so that the expected delay per unit time is minimized. The solution follows.

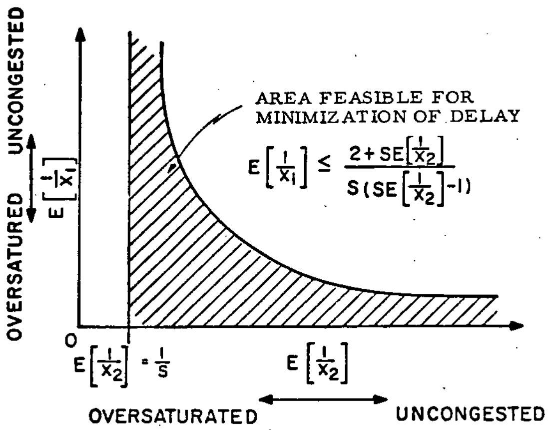


Figure F-4. Feasible region of volume combination.

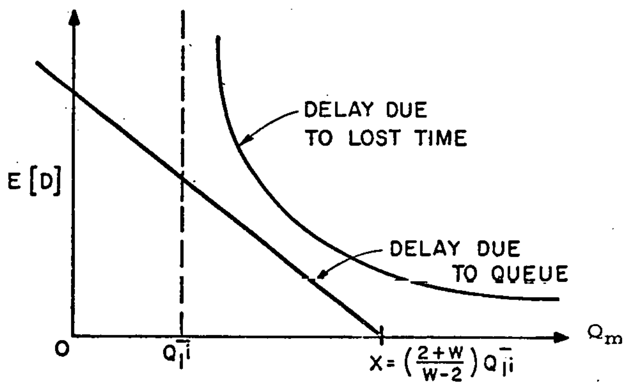


Figure F-5. E[delay/unit time] vs. Q_m for low volumes.

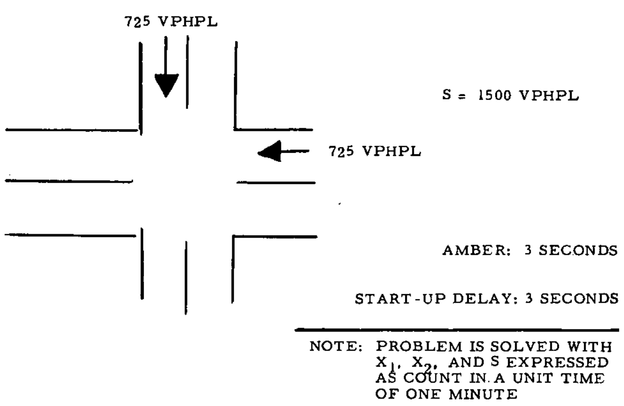


Figure F-6. An illustrative problem.

Given an intersection and traffic volumes, one would usually wish to know the average delay and the expected cycle length under queue-actuated control.

1. *Step 1*—Check whether the given volumes warrant queue-actuated control and complete various mathematical manipulations by determining several different expected values of the random variables X_1 and X_2 (flow rates):

a. Means: $X_1 = X_2 = 12.1$.

b. $E\left[\frac{1}{X_1}\right] = E\left[\frac{1}{X_2}\right] = 0.095$ by direct computation; where Poisson probabilities are used,

$$E\left[\frac{1}{X_1}\right] = \sum_{X_1=0}^{\infty} P(X_1) \cdot \frac{1}{X_1} \text{ and } E\left[\frac{1}{X_2}\right] = \sum_{X_2=0}^{\infty} P(X_2)\frac{1}{X_2}$$

c. $E[X_1^2] = E[X_2^2] = 158.5$ by direct computation;

$$E[X_1^2] = \sum_{X_1=0}^{\infty} P(X_1) \cdot X_1^2 \text{ and } E[X_2^2] = \sum_{X_2=0}^{\infty} P(X_2) \cdot X_2^2,$$

or by the always applicable relation that $E[X^2] = \sigma^2 + \mu^2$ and $\sigma^2 = \mu$ for a Poisson.

2. *Step 2*—Check volumes to see if queue-actuated control should be applied, by satisfying the following equation:

$$E\left[\frac{1}{X_1}\right] \leq \frac{2 + SE\left[\frac{1}{X_2}\right]}{\left(SE\left[\frac{1}{X_2}\right] - 1\right) S}.$$

With:

$$0.0915 \leq \frac{2 + (25)(0.0915)}{((25)(0.0915) - 1)(25)} = \frac{4.28}{32.0} = 0.134$$

queue-actuated control may be applied.

3. *Step 3*—Determine Q_m based on case 1. Before the average delay and cycle length are computed, the maximum queue length to be allowed on each approach must be determined. Many criteria could be used to determine the desired queue length, including average delay and pedestrian crossing requirements. Using the average delay as the criteria, the formula developed previously should be used to determine the Q_m that minimizes the average delay. Thus:

$$Q_m = \bar{Q}_{1i} \pm \sqrt{\frac{LB}{(2-W)\,S}}$$

where:

$$\bar{Q}_{1i} = \frac{\lambda_1}{2(S - \lambda_1 - \lambda_2)} = \frac{12.1}{(2)(0.8)} = 7.5;$$

$$W = S^2 E\left[\frac{1}{X_1}\right] E\left[\frac{1}{X_2}\right] - SE\left[\frac{1}{X_1}\right] - SE\left[\frac{1}{X_2}\right] = 0.66;$$

$B = \lambda_1 E[X_2^2] + \lambda_2 E[X_1^2] = 3835.9;$

$L = \text{loss time} = 12 \text{ sec} = 0.2 \text{ min};$ and

$$Q_m = 7.5 + \sqrt{\frac{(0.2)(3835.9)}{(2 - 0.66)(25)}} = 12.25.$$

4. *Step 4*—Check in each case the probability of queue being cleared before the end of green:

$$P_1 = P_2 = \sum_{X_2=0}^{\infty}\left(\sum_{K=0}^{SA} P(X_1 = K/K_2)\, P(K_2)\right) = 0.204$$

where $P_1 = P_2$ because of the equal arrival rates. Recall the following:

$P(i)$ = probability of case i
$P(1) = (1 - 0.204)(1 - 0.204) = 0.633$
$P(2) = (1 - 0.204)(0.204) = 0.163$
$P(3) = (0.204)(1 - 0.204) = 0.163$
$P(4) = (0.204)(0.204) = 0.041$

Thus, the assumption that case 1 is the most likely case was correct in this example.

5. *Step 5*—Determine the average delay and effective cycle length using the formulas given in Tables F-2 and F-3, and now knowing that $Q_m = 12.25$; it follows that:

CASE	$E[D_T]$ (min)	$E[A]$ (veh)	AVERAGE DELAY (sec/veh)	EFFECTIVE CYCLE LENGTH (sec), WITHOUT AMBERS
1	18.6	21.7	51.3	59.6
2	21.7	21.7	60.0	59.6
3	17.8	35.8	29.8	93.0
4	28.9	51.6	33.6	134.4

6. *Step 6*—The average delay for all cases is:

$$d = \sum_{i=1}^{4} P(i)\, d(i)$$

where:

$P(i)$ = probability of case i;
$d(i)$ = average delay of case i; and
d = 48.5 sec.

7. *Step 7*—The average effective cycle length for all cases is:

$$t = \sum_{i=1}^{4} P(i)\, t(i) = 68.0$$

where:

$t(i)$ = effective cycle length of case i; and
t = 68.0 sec + 6 sec amber = 74 sec.

However, a constraint of a maximum green could eliminate case 4, whose effective cycle length appears to be too long.

To examine whether or not Q_m based on case 1 only would really result in a minimum average delay for all cases, the average delay for all cases is plotted as a function of Q_m, as shown in Figure F-7. The results indicate that the Q_m which minimizes the average delay for all cases (point M in the figure) is approximately 13.25, as opposed to $Q_m = 12.25$ based on case 1 only. The increase in the average delay is less than 2 percent; therefore, it is considered that the use of Q_m determined from case 1 would be adequate most of the time.

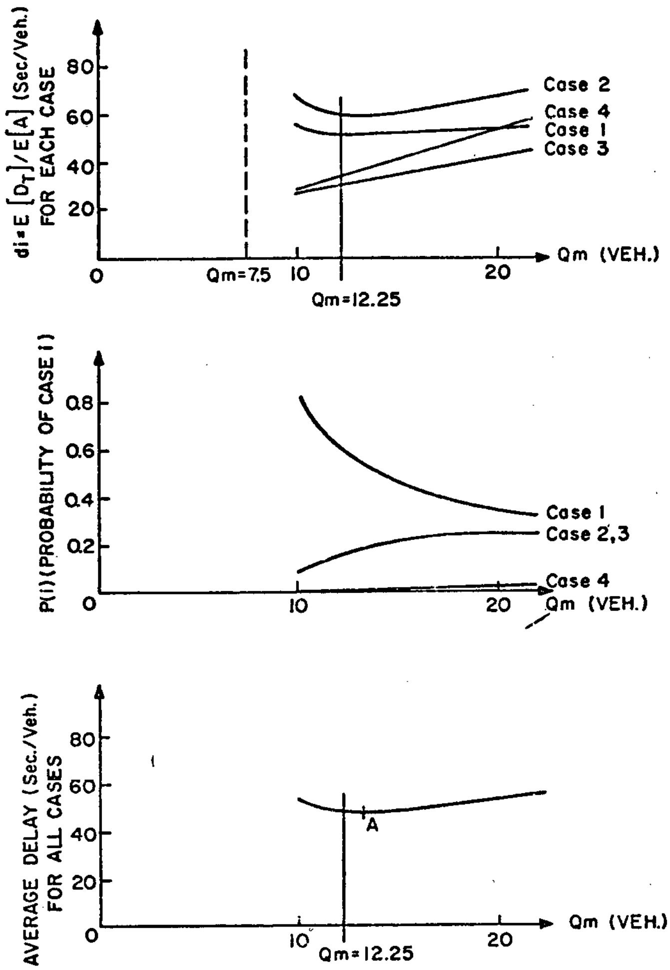


Figure F-7. Graphical solution of Q_m for example problem.

TABLE F-4

COMPARISON OF COMPUTED AND SIMULATED RESULTS

	Computed	Simulated
$\overline{Q}_{1i}$	7.5	5.6
Average Delay/ Veh. (Sec.)	48.5	51.0
Average Effective Cycle Length (Sec.)	74.0	78.0

Simulation Results of the Sample Problem

A UTCS-1 simulation was run of the control policy for the sample problem just presented. The control algorithm was programmed, and the detective locations previously computed were used.

As can be seen from the comparison in Table F-4, the difference between the computed and the simulated values of $\overline{Q}_{1i}$ is approximately 25 percent; however, the differences are less than 5 percent for both the average delay per vehicle and the average effective cycle length. The comparison also shows that the computed values of the average delay and average cycle length are smaller than the simulated values. This was expected, because the delay model did not explicitly take into consideration arrivals during the loss time.

COORDINATED SIGNALIZED INTERSECTIONS

Queue-Actuated Control in Midst of Coordinated Intersections

When coordinated systems of signals are considered, the underlying philosophy is generally one of establishing signal timing to facilitate the uninterrupted movement of through vehicles along a roadway. In saturated or oversaturated operations, however, development of signal timing for such a purpose has little justification in that all vehicles must join queues and stop at least once somewhere upstream of the critical intersections. In practical terms, there is no "through" traffic.

Introduction of the "queue-actuated control policy," in the midst of a coordinated system of signals, is functionally similar to introduction of any of the more conventional traffic-responsive control equipment (e.g., full-actuated, volume-density), in that it places an essentially cycle-free traffic controller in the midst of a set of fixed-time signals.

It is worthwhile, however, to consider how the queue-actuated control policy performs when implemented at a critical intersection, which itself is in the midst of a coordinated system of signals. As this policy is queue-responsive (i.e., green time allocation is dependent on some maximum queue length), the signal at the critical intersection will be "driven" by vehicle inputs from the immediate upstream intersections. Two extremes of conditions can exist.

One extreme is where vehicle inputs are all provided by the through-link upstream of the intersections upstream of the critical intersection. In this case, green at the critical intersection will initiate as soon as sufficient additional through vehicles cause the queue to reach its maximum allowed value. Thus, the relative offset between these two intersections "looks" like a forward progression. The other extreme is when all vehicles feeding the queue at the critical intersection are turn-ins from the upstream intersection. In such a situation, the maximum allowable queue length will be reached before the initiation of green at the upstream intersection, and the relative offset will "look" like a reverse progression. For some combination of through and turn-in traffic, the relative offset will "look" to be simultaneous.

Thus, one sees that the actual initiation time and duration of green times at the critical intersection under the queue-actuated control policy is not only dependent on the predetermined value for maximum allowable queue on the competing approaches and their respective volumes, but also on the way in which these volumes input into the competing links.

The situation downstream of an initial intersection operating under this policy is different, in that a fixed-time signal with predetermined green time allocations receives traffic from an essentially cycle-free signal. In such a situation it is possible, in theory, for a green phase at the critical intersection to occur that will provide more traffic to the next downstream signal than it can handle, and thus create a saturated condition there. In such a situation the problem of saturation will not have been eliminated, only relocated. This can only occur, however, when traffic on the cross street at the critical intersection is very light, thus allowing the "main" street an unusually long green time. In practice, however, this is quite unlikely, because one of the inherent features of the critical intersection is the presence of heavy traffic on both of the competing approaches.

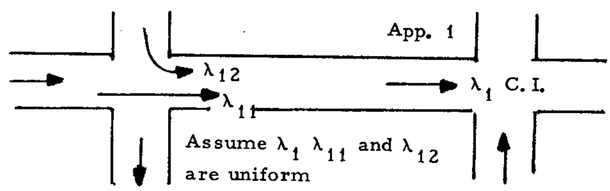
Effects of Upstream Signals on the Delay Model

Poisson arrivals were assumed in computing P_1, P_2, $E[1/X_1]$, $E[1/X_2]$, and Q_{1i}. These quantities can be recomputed for other distributions. For the last three, which are the only relevant ones when case 1 computations are used, it is sufficient to know the variance-to-mean ratio—for example, in computing Q_{1i} (see earlier discussion under "Evaluation of Q_{1i}").

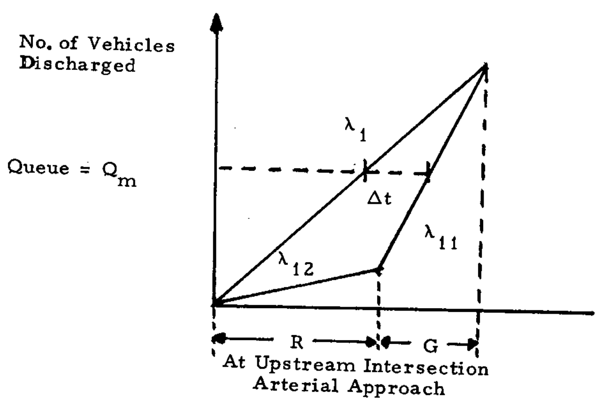
In addition to this, however, it has been assumed in the development that a single pattern of arrivals holds over an entire cycle, with length to differ from the expected value previously computed. The delay estimate will also be in error, but not by as detectable (or significant) an amount. The material to follow estimates the change in expected cycle length due to the arrival pattern driving the switching.

Referring to Figure F-8(A), for a given λ_1, there could be numerous combinations of λ_{11} and λ_{12}. Since switching occurs at a critical intersection, when queues reach a predetermined length, the switching time based on the assumption of a single constant λ_1 would be incorrect by time Δt, as depicted in Figure F-8(B).

Because the signal at the critical intersection is cycle-free, the switching time could fall at any point within the cycle of the upstream signal. Therefore, the expected Δt would be:



(A) Creation of λ_1



(B) Development of Queue

Figure F-8. Error in switching time.

$$E[\Delta t] = \frac{1}{2}(R+G) - \frac{\lambda_1 (R+G)}{\lambda_{11}} = \frac{1}{2}\left(\frac{C_1 \lambda_{11} - C_1 \lambda_1}{\lambda_{11}}\right) \quad \text{(F-11)}$$

where C_1 is the cycle length at the upstream intersection of Approach 1 of the critical intersection.

Note that Δt could be positive or negative, depending on the flow patterns at the upstream intersections. Since the same type of adjustment has to be made for Approach 2 of the critical intersection, the total adjustment, Δt, required to obtain the correct cycle length would be:

$$E[\Delta t] = \frac{1}{2}\left(\frac{C_1 \lambda_{11} - C_1 \lambda_1}{\lambda_{11}}\right) + \frac{1}{2}\left(\frac{C_2 \lambda_{21} - C_2 \lambda_2}{\lambda_{21}}\right) \quad \text{(F-12)}$$

where:

C_i = cycle length at the upstream intersection of Approach i of a critical intersection;
λ_i = arterial flow at the upstream intersection; and
λ_{i1} = cross-street flow at the upstream intersection.

Some Special Aspects

Switching

When all queues are cleared at one approach before queues reach Q_m at the other approach, one could either terminate green as soon as there are no queues (Fig. F-9 (A)), or wait until queues at the other approach reach Q_m (Fig. F-9(B)).

When Alternate 1 is chosen, it is no longer queue-actuated control, and it is more like "basic queue control" (*F-1*). As discussed previously, this would likely occur when flows are light, case 1 then not being the most likely.

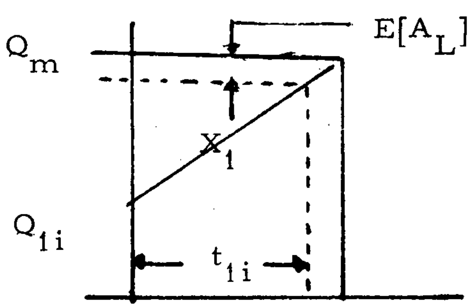
Arrivals During Loss Time

In the development of the delay model, arrivals during loss time were not taken into consideration. If the arrivals during the loss time are considered, the delay diagram would be as shown in Figure F-10 instead of Figure F-1.

Therefore, to be more precise, t_{1i} should be

$$t_{1i} = \frac{Q_m - Q_{1i} - E[A_L]}{X_1}$$

where $E[A_L]$ is the expected arrival during start-up delay:



However, consideration of $E[A_L]$ is not critical unless Q_m approaches the link length. In such cases, actual physical blockage of the upstream intersection could arise.

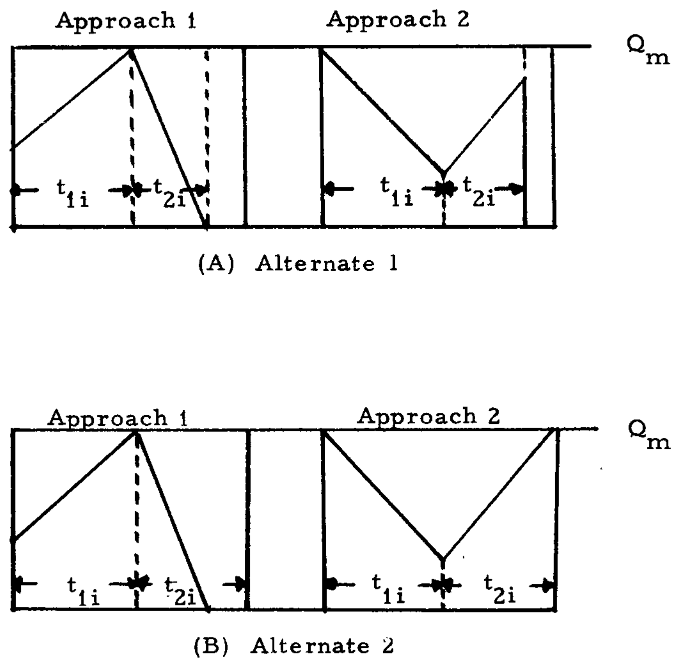


Figure F-9. Two alternatives for switching time.

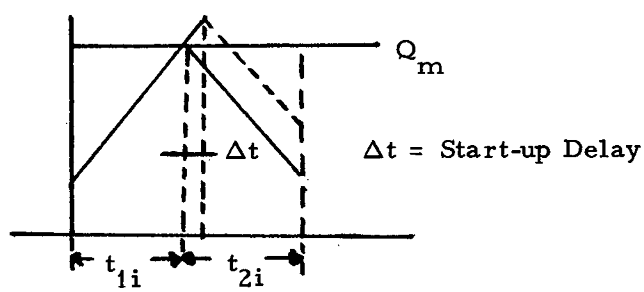


Figure F-10. Delay diagram with arrivals during loss time.

Sensitivity of Effective Cycle Length with Respect to Q_m

Since

$$t_i = \frac{S(Q_m - \overline{Q}_{1i})}{X_1 X_2}$$

$$\overline{Q}_{1i} = \frac{\lambda_1}{2(S - \lambda_1 - \lambda_2)}$$

it follows that

$$t_i = \frac{S\left(Q_m - \dfrac{\lambda_1}{2(S-\lambda_1-\lambda_2)}\right)}{X_1 X_2} = \left(\frac{S}{X_1 X_2}\right)\left(\frac{\lambda_1}{2(S-\lambda_1-\lambda_2) X_1 X_2}\right)$$

Because the slope of the line is $S/X_1 X_2$, the heavier the flow, the less sensitive is t_i with respect to Q_m. When flows are heavy, the error in effective cycle length due to error in Q_m would be smaller than when flows are light.

Figure F-11 illustrates the sensitivity and also the practical limits on t_i and Q_m.

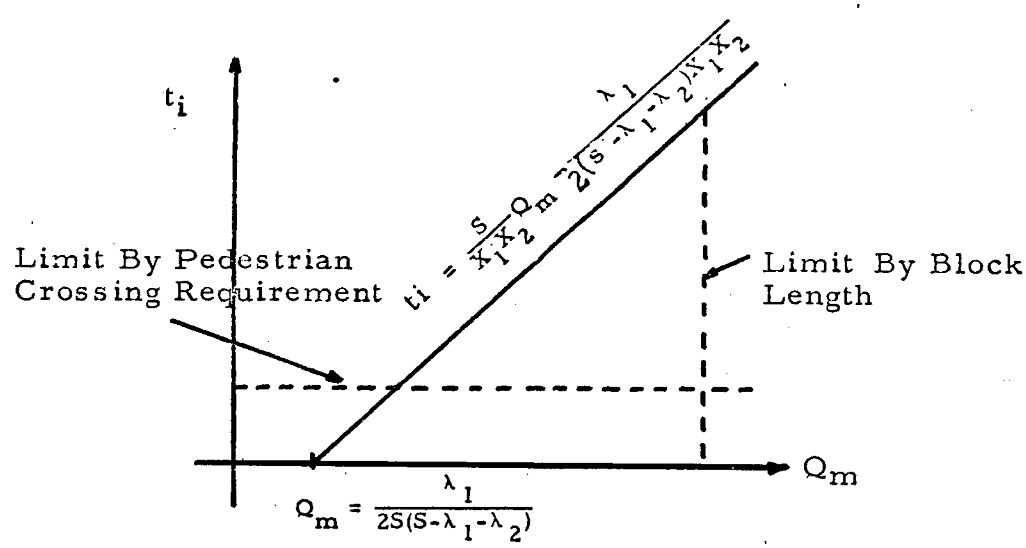


Figure F-11. Sensitivity of t_i with respect to Q_m.

ADDITIONAL WORK

Further research on the following aspects is recommended:

1. Since queue-actuated control must rely on accurate detection of queue length, the effects of errors in detection must be studied considering the limited accuracy of the detector system presently available.
2. The more desirable form of any control measure is one that could be used in oversaturated periods as well as in undersaturated periods. Therefore, the possibility of extending queue-actuated control to the period of undersaturation should be investigated.
3. An analytic model for the effects of queue-actuated control on adjacent intersections would help to better understand the effects of this control in the midst of coordinated signals.

REFERENCES

F-1. WAGNER, F. A., BARNES, F. C., and GERLOUGH, D. L., "Improved Criteria for Traffic Signal Systems in Urban Networks." *NCHRP Report 124* (1971) 86 pp.

F-2. YAGODA, H., "The Control of Arterial Street Traffic." Paper presented at the 1967 TFAC Control Conference, Haifa, Israel.

F-3. NEWELL, G., "Approximation Methods for Queues with Application to Fixed-Cycle Traffic Light." *SIAM Review*, Vol. 8 (1965).

F-4. SAGI, G., and CAMPBELL, L., "Vehicle Delay at Signalized Intersections—Theory and Practice." *Traffic Engineering* (Feb. 1969).

F-5. WEBSTER, F. V., "Traffic Signal Settings." *Road Research Laboratory Technical Paper No. 39* (1961).

APPENDIX G

SOME ANALYTICS ON TURN BAYS AND DELAY

If a turn bay is provided on a link that has L through lanes, a total flow, FLOW, and a right-turn flow, RTFLOW, some savings in delay will be realized and some increase in intersection productivity will be achieved.

This appendix presents the analytics that support a first-order estimate of the interaction of turn-bay length, flow and flow parameters, cycle length, and delay and productivity. As a result of this work, it is recognized that there is some benefit that can be achieved by matching the queue size to the available bay length. The queue size can be managed by control of the cycle length.

APPROACH CHARACTERISTICS

Consider that there is a total flow, FLOW; a right-turn flow, RTFLOW; and L through lanes. One might expect queues to form in the L lanes proportional to (FLOW/L), with all right turners concentrated in the right-hand lane.

There are no left turns. The percentage of turners in the right lane is thus

$$p_0 = \frac{\text{RTFLOW}}{(\text{FLOW})/L} \qquad \text{(G-1)}$$

(see Fig. G-1(A)).

However, if there is a turn bay, one might well expect the queues in the through lanes to be proportional to (FLOW-RTFLOW)/L. The vehicles arriving in the rightmost lane also include all the turners, so that the percentage of turners in that lane is

$$p_1 = \frac{\text{RTFLOW}}{\dfrac{\text{FLOW-RTFLOW}}{L} + \text{RTFLOW}}$$

$$p_1 = \frac{L(\text{RTFLOW})}{\text{FLOW} + (L-1)\,\text{RTFLOW}} \qquad \text{(G-2)}$$

(see Fig. G-1(B)).

Table G-1 summarizes p_1 for common values of L and turn percentage.

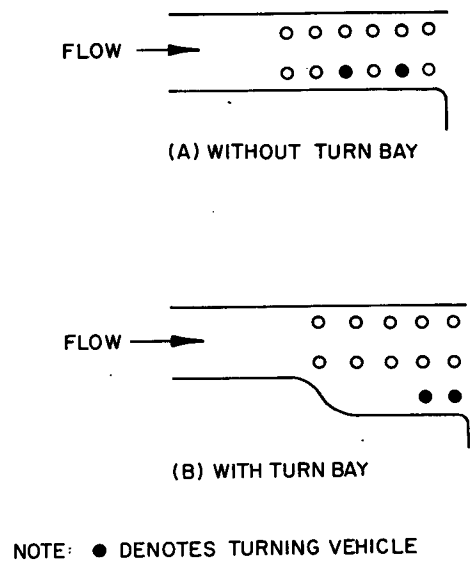


Figure G-1. Distribution of queue.

CYCLE EFFECTS

Given a cycle length of C seconds per hour, there are

$$N = \text{FLOW}\left(\frac{C}{3600}\right) \qquad \text{(G-3)}$$

vehicles per cycle, of which

$$M = \left\{\frac{\text{FLOW}}{L} + \left(\frac{L-1}{L}\right)\text{RTFLOW}\right\}\frac{C}{3600}$$

or

$$M = \left(\frac{p_T}{p_1}\right)N \qquad \text{(G-4)}$$

vehicles are in the rightmost lane, including the turn bay. If there is no turn bay, $M = N/L$.

Note that M may be decreased simply by decreasing the cycle length.

TABLE G-1

VALUES OF p_1 FOR VARIOUS LANES L

p_T TURN PERCENTAGE	NUMBER OF LANES		
	1	2	3
0.05	0.05	0.095	0.136
0.10	0.10	0.182	0.250
0.15	0.15	0.261	0.346
0.20	0.20	0.333	0.429

DELAY WITH NO TURN BAY

Assume for simplicity that there are n vehicles every cycle, of which exactly M are in the right-hand lane. As noted, $M = N/L$.

Let X be the number of right turners among the M vehicles. In a given cycle, this number is binomially distributed with probability p_1 describing the right-turn event. Thus

$$p(x) \triangleq P(X = x) = \binom{M}{x} p_1^{\,x}(1-p_1)^{m-x} \qquad \text{(G-5)}$$

Let D_i be the delay of the i^{th} vehicle. Assume

$$D_i = 3.5 + 2.0\,i$$

wherein all the start-up delay and some loss time is loaded onto the first vehicle ($i = 1$). The total delay is thus

$$D = \sum_{i=1}^{M} D_i = 3.5\,M + M\,(M+1) \qquad \text{(G-6)}$$

and the average delay is

$$\text{AVGD} = \frac{D}{M} = 4.5 + M \qquad \text{(G-7)}$$

DELAY WITH A TURN BAY

Assume that all turners who exist in any given cycle are free to enter the bay. If there are M vehicles in the right-hand lane and X of them are turners, they will form two separate files for discharge. The delays of the two files are given by

$$D_{\text{thru}}(X) = 3.5\,(M-X) + (M-X)\,(M-X+1)\,, \qquad \text{(G-8a)}$$

$$D_{\text{turn}}(X) = 3.5\,X + X\,(X+1) \qquad \text{(G-8b)}$$

$$D\,(X) = D_{\text{thru}}(X) + D_{\text{turn}}(X) \qquad \text{(G-8c)}$$

where the delay is now a function of the random variable X, which is binomially distributed. Note that

$$E\,[D\,(X)] = \sum_{x=0}^{M} p\,(x)\,D\,(X) \qquad \text{(G-9)}$$

and likewise for $E\,[D_{\text{thru}}(X)]$ and $E\,[D_{\text{turn}}(X)]$. Also,

$$E\,[\text{AVGD}\,(X)] = \frac{E\,[D\,(X)]}{M} \qquad \text{(G-10)}$$

But

$$E\,[\mathrm{AVGD_{thru}}\,(X)] = \sum_{x=0}^{M-1} \frac{D_{\mathrm{thru}}\,(X)}{M-X}\, p\,(x) \qquad \text{(G-11a)}$$

and

$$E\,[\mathrm{AVGD_{turn}}\,(X)] = \sum_{x=1}^{M} \frac{D_{\mathrm{turn}}\,(X)}{X}\, p\,(x) \qquad \text{(G-11b)}$$

if one is interested in the average delay experienced by persons in each category.

PRODUCTIVITY

A simple deterministic computation will be done of potential productivity increase.

Note that if N vehicles arrive in a given cycle and there is no turn bay, the queue in each lane is $M = N/L$ and it requires

$$T_0 = 3.5 + 2.0 \left(\frac{N}{L}\right)$$

seconds to discharge these vehicles, using the foregoing numbers.

If there is a turn bay, there are

$$\left(\frac{\text{FLOW-RTFLOW}}{L}\right)\frac{C}{3600} = \frac{N}{L}\,(1 - p_T)$$

vehicles in each lane, requiring

$$T_1 = 3.5 + 2.0\,\frac{N}{L}(1 - p_T)$$

seconds to discharge these vehicles. The effective green required is thus reduced by the percentage of turners, p_T.

COMPUTATIONS FOR DELAY

Figure G-2 shows the delay with and without a turn bay for the single-lane case ($L = 1$), assuming that all turners enter the bay. Other cases may be derived from this by adjusting for the differing queue lengths by lane. However, the case shown is representative. The adjustment for $L = 2$ is particularly small.

From the results in Figure G-2, two features of the turn bay design have emerged:

1. The delay benefits are realized primarily by the turners, not the through vehicles.
2. There can be substantial increases in the productivity, or decreases in the required green, whichever is more beneficial.

Note that the delay indicated is the component of delay during queue discharge. The component of delay due to the red phase is not shown, nor need it be considered because it cannot be changed without changing the red phase.

Note also that if right-turn-on-red (RTOR) is permitted, some localized delay (e.g., Fig. G-2(B)) can be avoided, although it is most probable that in a congested environment, there will be no substantial impact on total system delay.

FINITE TURN BAYS

In the foregoing development, it was assumed that all turners were in fact able to enter the turn bay and prepare for discharge. This may not occur for one of the following reasons:

1. The number of turners, X, exceeds the turn-bay length, even though they would have access to the bay.
2. Some or all of the turners do not have access to the bay because the through vehicles are in sufficient number to block access to the bay.

Figure G-3 illustrates the probability of blockage (including both possibilities) for three bay lengths. To assure

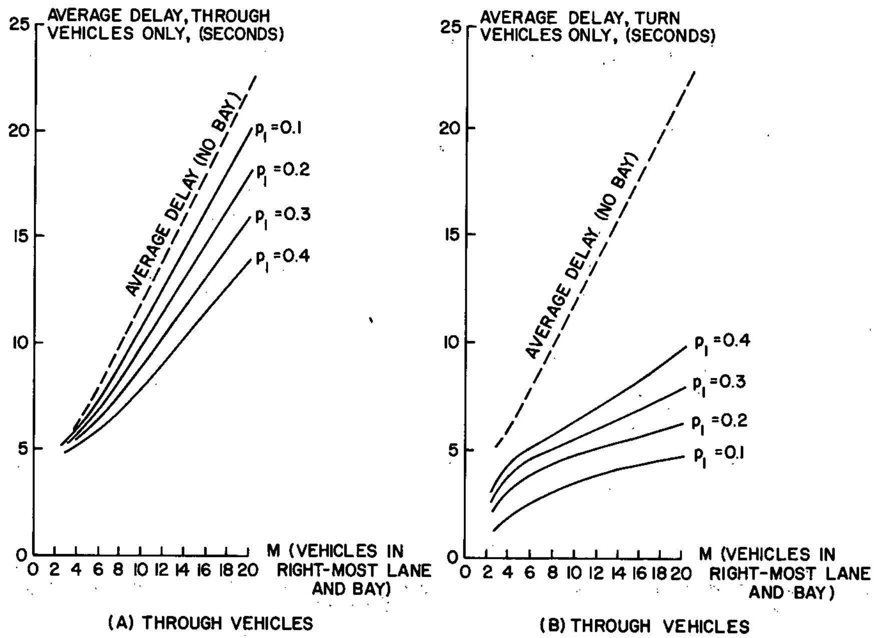


Figure G-2. Delay savings for single-lane case with turn bay.

APPENDIX H

SIMULATION OF SIGNAL POLICIES

This appendix reports on the simulation results obtained with various highly responsive signal control policies, including some advanced control concepts.

The advantages or disadvantages of three relevant advanced control concepts selected from the literature (smooth flow theory, Gazis minimum delay policy, and Longley's real-time queue proportionality policy) are cited in Chapter Two, as is a basic definition of the policies. A fourth concept, the "queue-actuated control policy" proposed by the research agency, is addressed in Appendix F. Subroutines were written for the UTCS-1 simulator for the real-time control policies, and are detailed in Appendix I.

The following compendium of results is based on several case studies using different test networks. Some notes on the intent or significance of the tests conducted are included in the discussion. The interpretations and application of these results are contained in the text and in Appendix J.

NETWORKS TO BE TESTED

Four networks were used as case studies: (1) Flatbush Avenue Extension, the location of two sites reported in Appendix D (Figs. H-1, H-2, and H-3), and (2) the three networks shown in Figure H-4. Table H-1 summarizes the situations studied and documented in this appendix.

PRELIMINARY DISCUSSION ON THE MORE RESPONSIVE POLICIES

Queue-actuated control, as previously noted, is addressed in Appendix F. Gazis' and Longley's policies are covered in the literature. Longley's strategy has been modified in this study for practical purposes as described in the following. It was also extended to signalized and non-signalized intersections.

Let X_1, X_2, Y_1, and Y_2 equal the queue length noted at the four approaches at the end of the red phases of the previous cycle, and LX_1, LX_2, LY_1, and LY_2 equal the respective link lengths (see Fig. H-5). The percent split in the X direction is given by:

$$p = \left\{ \frac{\text{Min}\ \{LY_i - Y_i\}}{\text{Min}\ \{LY_i - Y_i\} + \text{Min}\ \{LX_i - X_i\}} \right\} 100\% \tag{H-1}$$

Thus, the split is a function of the clear distance between the maximum queue and the upstream intersection.

To determine which technique of basing splits should be used—queue detected only over the previous cycle or some sort of averaging—simulations were run for an identical network under the two different methods. Exponential smoothing over three previous cycles was used as the averaging technique. As shown in Figure H-6, the two methods yield almost identical values of system delay;

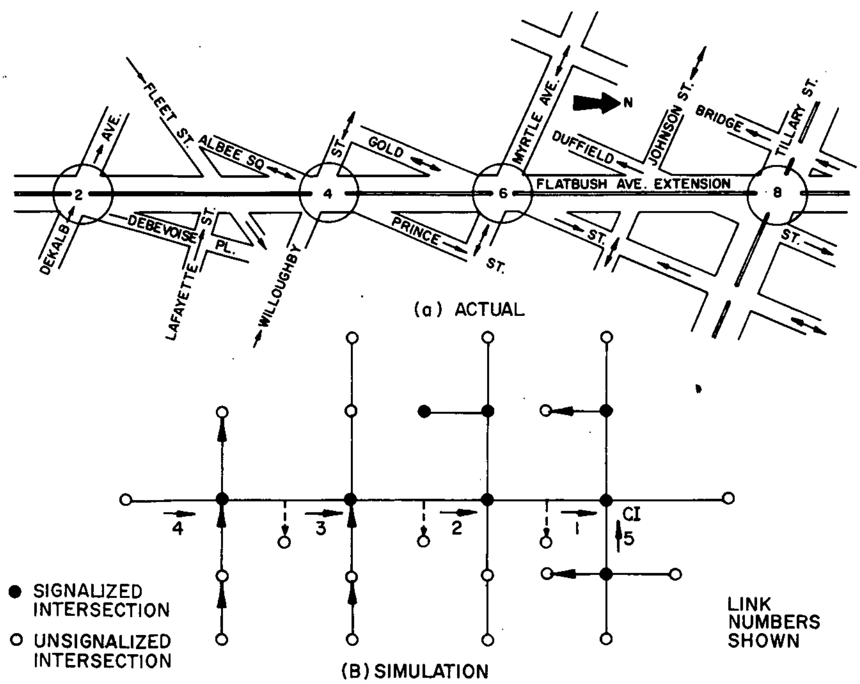


Figure H-1. Existing Flatbush Avenue Extension.

that such blockage does not occur, the queue M can only exceed the available turn bay by one vehicle. (A block probability ≤ 0.1 was assumed in making this statement.) Only then can the full productivity increase be assured, and the complete anticipated delay savings be realized.

LEFT-TURN BAYS

It should be noted that the productivity references for right turners is predicated on the formation of two parallel files, each of which is capable of discharging at essentially the same rate. If a right turner is caught in the "through file" by virtue of not having access to the turn bay, he will leave a "hole" in the through file when he switches to the turn bay, leaving a productivity gap.

For left turners, however, the bay serves a different function: left turners must generally wait for a turning opportunity, delaying all those behind them. The bay serves to eliminate such delays to the through vehicles.

SUMMARY

The work developed in this appendix has indicated the potential increases in productivity due to right-turn bays and the distribution of delay benefits among turners and through vehicles. It has also shown that the queue allowed should not be too much larger than the available bay (or the bay should be about as long as the queue, if this option exists). To assure blockage (as defined) no more than 50 percent of the time, one should make certain that

$$M \leq \text{BAY}\,(1 + 2p_1) \qquad \text{(G-12)}$$

as a rule-of-thumb derived from Figure G-3, where BAY is the length of the turn bay in vehicles. Using this in conjunction with Eq. G-3 and Eq. G-4, one may show that when $p_T = 0.10$, the ratio of C in seconds to BAY in vehicles should be about 7.2 for 600 vphpl and 4.8 for 900 vphpl. This would require a bay for either 8 veh or 13 veh. In many blocks, this is a very substantial portion of the block.

However, this computation does not mean that some or even most of the productivity increases and delay savings would not be realized with shorter bays. It does indicate that short cycle lengths aid in achieving such benefits, for shorter C values imply lower M values as needed in Eq. G-12.

In retrospect, some trends have been identified, and they may seem rather straightforward. The question now arises as to whether this analytic modeling should be refined to estimate productivity for various fixed values of flow and cycle length, or whether simulation should be used instead, or whether field work is needed. Within the scope of this project, it was decided to reinforce those trends with simulation results (see Appen. H and J).

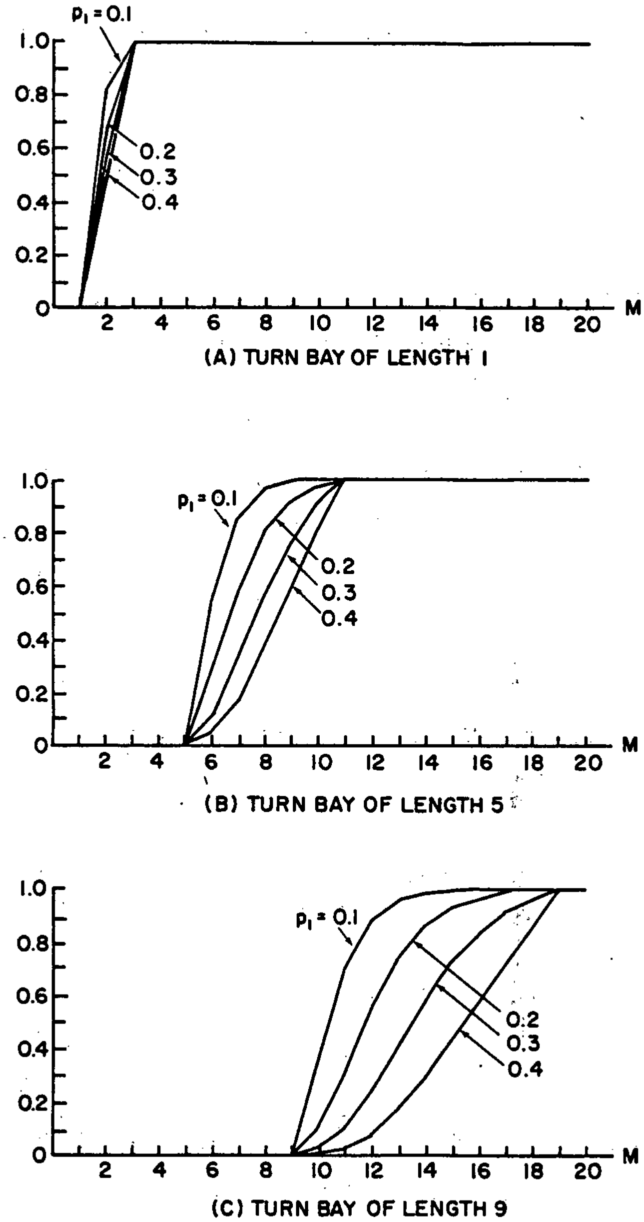


Figure G-3. Probabilities of blocking the turn bays.

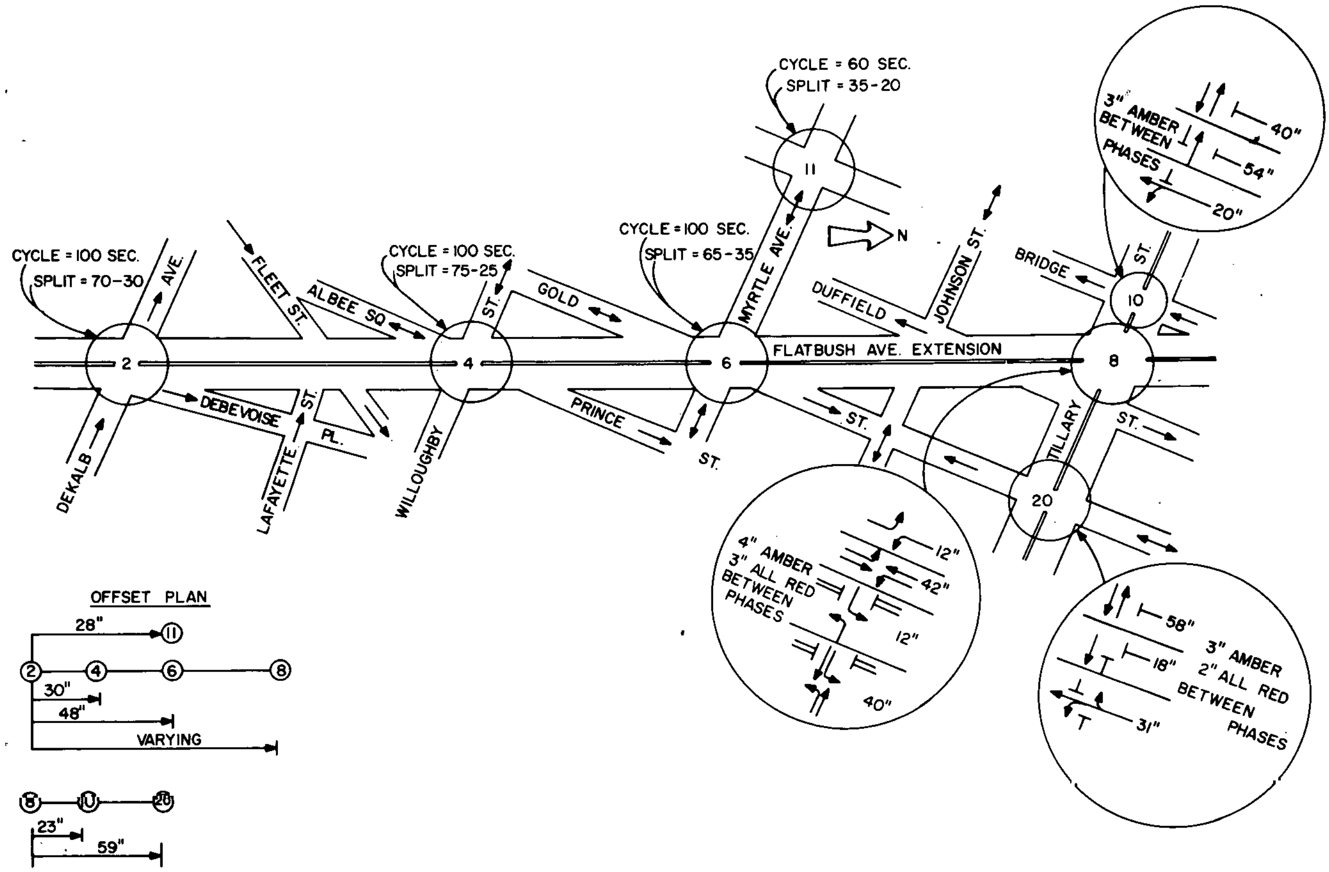


Figure H-2. Existing Flatbush Avenue Extension signal timing plan.

neither technique has an advantage over the other, when compared with the actual queue formation at the critical intersection. This is probably because of the perfect detection only possible in simulation.

Throughout this study, the technique of basing splits on queue lengths determined in the previous cycle is used, because it was believed to yield a more responsive control. Furthermore, a constant cycle length was used for all signals in the simulated system when this control policy was evaluated to minimize the effects of offset disruption resulting from the use of different cycle lengths.

The following three conditions must be considered to implement this control strategy:

1. When queues at both critical approaches are less than link length, the split is determined as based on Eq. H-1.

2. When the queue at one of the approaches (approach A in this case) is at least equal to link length (spillback), approach A is assigned the maximum green and approach B the minimum green. The maximum and minimum greens are predetermined on the basis of pedestrian crossing requirements as well as other factors.

3. When spillbacks occur on both approaches, the split at the critical intersection is based on the queue/link length ratios of the links upstream of intersections A and B and the same technique as in (1) is used. Signal splits at intersections A and B are adjusted to have the same amount of green as at the critical intersection and their respective upstream intersections. The excess green at intersections A and B is assigned to cross streets assuming turn-ins are not a major cause of oversaturation at the critical intersection.

CASE STUDIES

Test Set 1

To study the effect of cycle length on queue build-up, Flatbush Avenue Extension was simulated under two cycle lengths of 100 and 60 sec for all intersections on Flatbush Avenue and Tillary Street. Only a two-phase signal was used for the 60-sec cycle length at the critical intersection because four phases were considered impractical for a cycle length as short as 60 sec. The existing offset was used for both cycle plans.

Queue length at both approaches to the critical intersection was compared for the two cycle lengths. Figure H-7 shows that, along Flatbush Avenue, the maximum queue length is least for a 60-sec cycle and the onset of oversaturation is somewhat delayed. Along the other approach, oversaturation no longer exists for both 60- and 100-sec cycles during the simulation period.

The data in Table H-2 indicate that a shorter cycle produces better operation of the critical intersection regard-

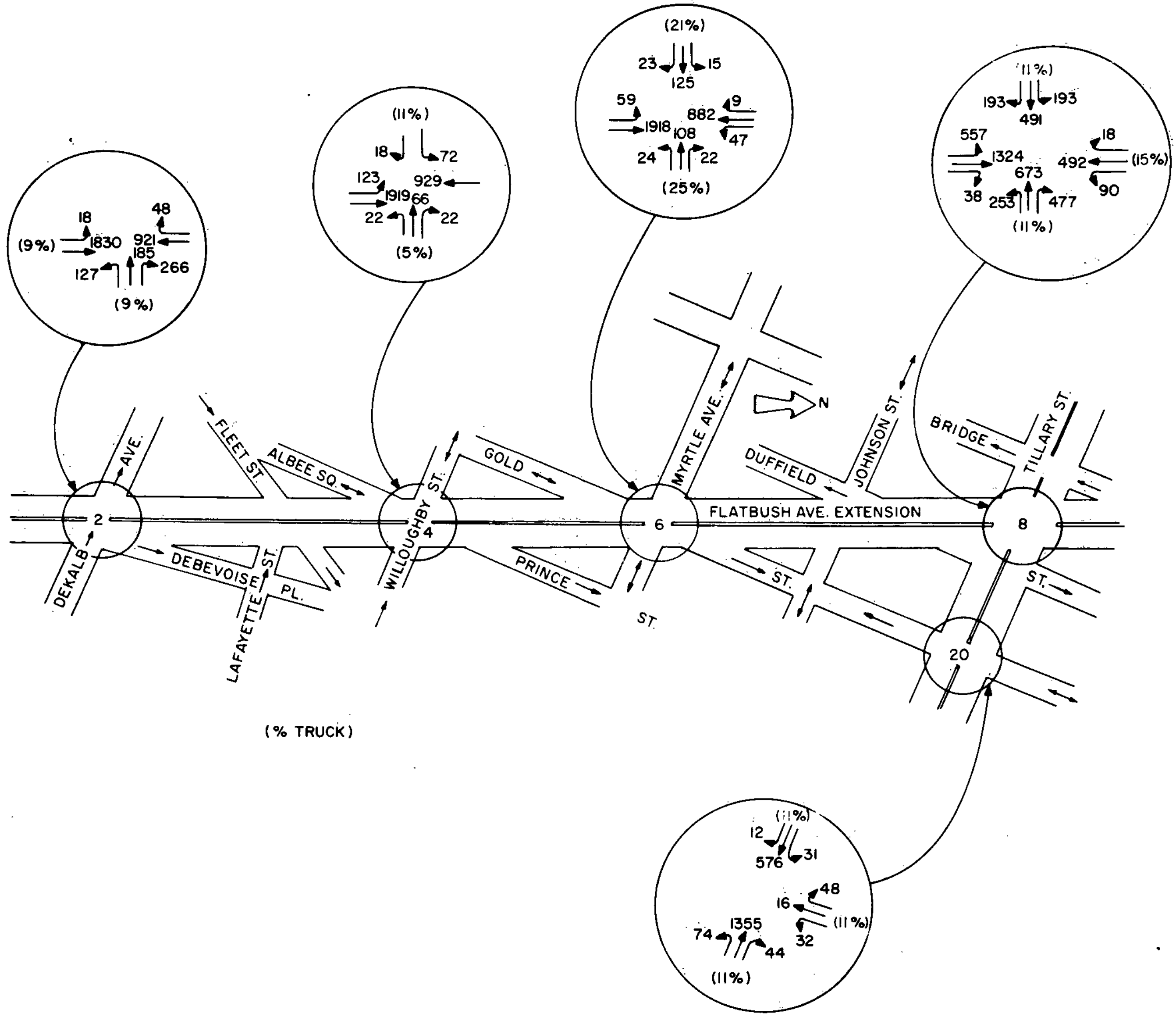


Figure H-3. Existing Flatbush Avenue Extension volume counts.

less of the measure of intersection performance used. The same table also shows that a greater improvement was achieved in the performance of cross streets than in that of the arterial (Flatbush Ave.). This results from the elimination of spillback along the arterial and is an example of how congestion management of one component of traffic can benefit other components of traffic in the system.

A long cycle length is often associated with a multiphase signal, which normally produces intersection oversaturation of Types I or IV (as defined in Chap. Two). When a critical intersection has a cycle length longer than the system cycle due to a multiphase, reducing the number of phases makes possible the use of a shorter cycle length as well as the use of a common cycle length for an entire system or any part thereof. A common cycle length reduces the problem of oversaturation of Type IV.

Even with a common cycle length, oversaturation can develop at a critical intersection that has a smaller green/cycle ratio than does the upstream intersection. This is an intersection oversaturation of Type I.

Test Set 2

In this test, Flatbush Avenue Extension was simulated under three different signal plans: (1) existing signal timing, (2) existing cycle lengths and splits with simultaneous offset for Flatbush Avenue, and (3) existing split with a common cycle length and simultaneous offset for Flatbush Avenue.

Existing Flatbush Avenue represents intersection oversaturation of Types I and IV. The critical intersection has a 120-sec cycle, and all upstream intersections have a 100-sec cycle. Figure H-8 and Table H-3 show that a simultaneous offset for three upstream signals not only significantly reduces the maximum queue length at the critical intersection, but also delays the onset of oversaturation of the critical intersection. The simultaneous offset plan also significantly reduces the average delay through the system. The common cycle length further delays the onset of oversaturation, although the average delay is slightly higher.

Test Set 3

In test set 3, existing Flatbush Avenue Extension exhibits intersection oversaturation of Types I and IV. The critical intersection has a smaller green/cycle ratio and capacity than the upstream sections, and different cycle lengths create varying offsets.

The findings showed that short cycle length, common cycle length, and simultaneous offset were quite efficient tools by which to reduce queue length at the critical intersection and the average delay through the system. The plan that produced the best results of the set used a 60-sec common cycle length with simultaneous offset and a two-phase signal at the critical intersection (see Fig. H-9 and Table H-4).

Test Set 4

One simulation study reported that, once saturation is present, minimum system delay occurs at two different cycle lengths: (1) when the critical intersection cycle length is extended just enough to provide the required arterial green and (2) when the critical intersection cycle is made equal to twice the system cycle length.

To test these findings, a one-way arterial was simulated under three different cycle plans:

1. Cycle length of all intersections was 60 sec with the critical intersection having a smaller green/cycle ratio than the other arterial signals.
2. Critical intersection cycle was increased to 84 sec to provide green time equal to that of other intersections.
3. Critical intersection cycle was increased to 120 sec.

From Table H-5, it can be seen that the results are not in agreement with the findings cited. An increase in average delay per vehicle was experienced by arterial traffic as well as cross-street traffic. The apparent differences, however, appear to be primarily due to differences in the two simulators used; the UTCS-1 simulator describes traffic behavior more realistically, especially under heavy flow conditions. Figure H-10 shows a comparison of the maximum queue length observed at the critical link under the three plans. This figure reveals that the common cycle (60 sec) produced almost continuous oversaturation in the critical link, whereas the other cycles (84 and 120 sec) yielded lower frequencies of occurrence of oversaturation of the link. Offsets could relieve some of this oversaturation problem.

Test Set 5

Yagoda et al. (Ref. (*41*) of Appen. C) present a method of timing signals on a two-way arterial that explicitly takes into account excess green time at several intersections along the arterial. They suggest that propagating the red wave down the street at a constant velocity

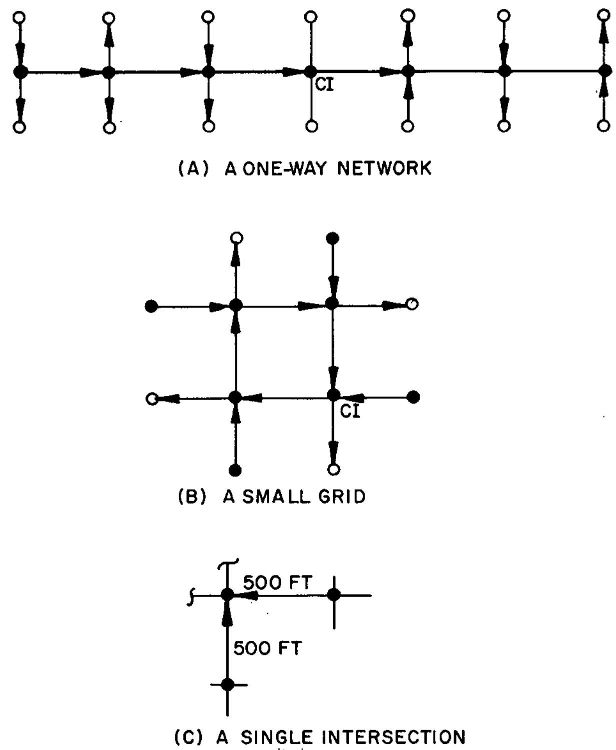


Figure H-4. Other networks studied.

may be more desirable than propagating the green wave. Accordingly, a one-way arterial was simulated under green-wave and red-wave progressions. The maximum queue lengths observed under both plans at the critical intersection are compared in Figure H-11. No significant differences can be noted. The average delay for red progression is about 9 percent higher than that for green progression. Thus, it would appear that under conditions of extremely heavy flows little is to be gained from such a progression.

Test Set 6

Another proposal for improving the problem of oversaturation at the critical intersection is by a gradual increase of bandwidth from the critical intersection upstream. The distance over which the change of bandwidth should be made will depend on the traffic as well as on the physical characteristics of the system. This technique will improve the critical intersection, but it will increase the queue and delay at the upstream intersections. As shown by the results in Figure H-12 and Table H-6, where the technique of "gradual change of bandwidth" was applied to a one-way arterial and compared with the standard progression which had the lower green/cycle ratio only at the critical intersection, the critical intersection is oversaturated during the period of initialization of the simulator (usually 600 sec) under heavy flow, but the gradual change of bandwidth significantly delays the onset of oversaturation. There are no differences in system average delay, but the technique significantly reduces delay at the critical intersection. An increase is experienced at the upstream intersections, as

TABLE H-1

A SUMMARY OF CASES STUDIED

TEST SET	SITE	CASES	INTENT OR COMMENTS
1	FLATBUSH AVE. EXT.	Existing Existing, with C = 100 (common) Existing, with C = 60 (common), 2 Phase at CI	Cycle length, effect
2	FLATBUSH AVE. EXT.	Existing Existing, with simultaneous up-stream on Flatbush Existing, with simultaneous and common cycle	Offset, and offset with common cycle
3	FLATBUSH AVE. EXT.	Existing Common C = 100, simultaneous Common C = 60, simultaneous, 2 phase at CI	Offset, and cycle length.
4	ONE-WAY ARTERIAL	C = 60 common C = 84 at CI C = 120 at CI	Common Cycle length vs. expanded at CI vs. multiple at CI
5	ONE-WAY ARTERIAL	Red Progression	An idea from the literature
6	ONE-WAY ARTERIAL	Progression to the CI Gradual change of bandwidth approaching the CI	An idea from the literature. Favorable results.
7	FLATBUSH AVE. EXT.	Existing Existing, but with uniform bandwidth Common cycle and uniform bandwidth	Reallocation of available time; organization of flow
8	INTERSECTION	Convention split Gazis Minimum delay Longley	local highly responsive vs. optimal delay minimization vs. conventional
9	FLATBUSH AVE. EXT.	Existing C=60 common, only change Queue-actuated at CI only change Longley at CI only change Simultaneous upstream intersections only Simultaneous all, C=60, 2ϕ at CI Simultaneous, C=60, but Longley at CI	Comparisons of highly responsive and minimally responsive policies
10	SMALL GRID	Simple progression Simple progression, Longley at CI Simultaneous Simultaneous, with Longley at CI Queue-Actuated at CI	Both low and major turn races considered; System lock-up observed.
11	ONE-WAY ARTERIAL	Progression Queue-actuation at CI	Adverse impact of queue-actuation except at heaviest flows.

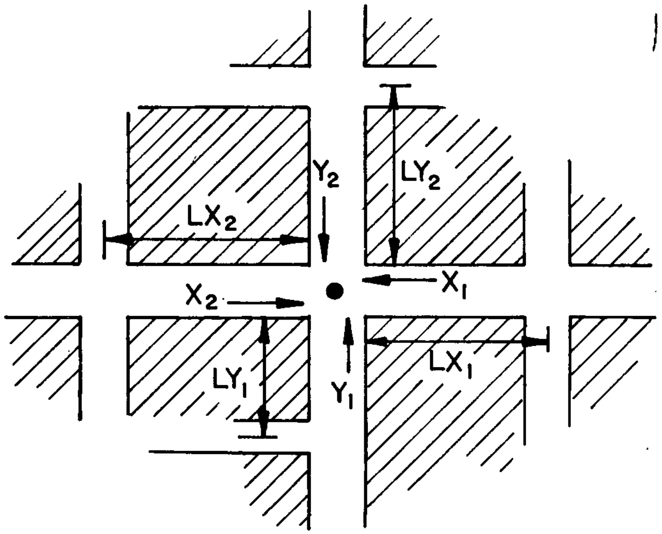


Figure H-5. Notation for Longley's control.

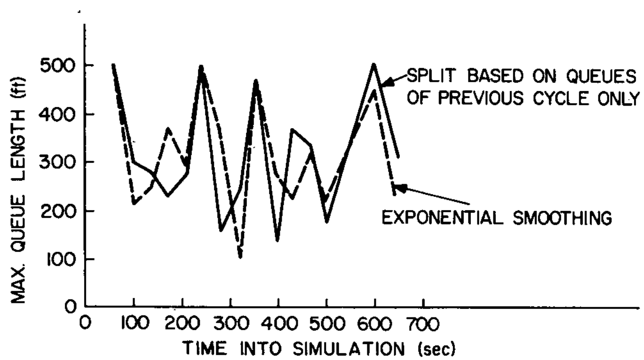


Figure H-6. Maximum queue length (link 4–5) for exponential smoothing and no smoothing.

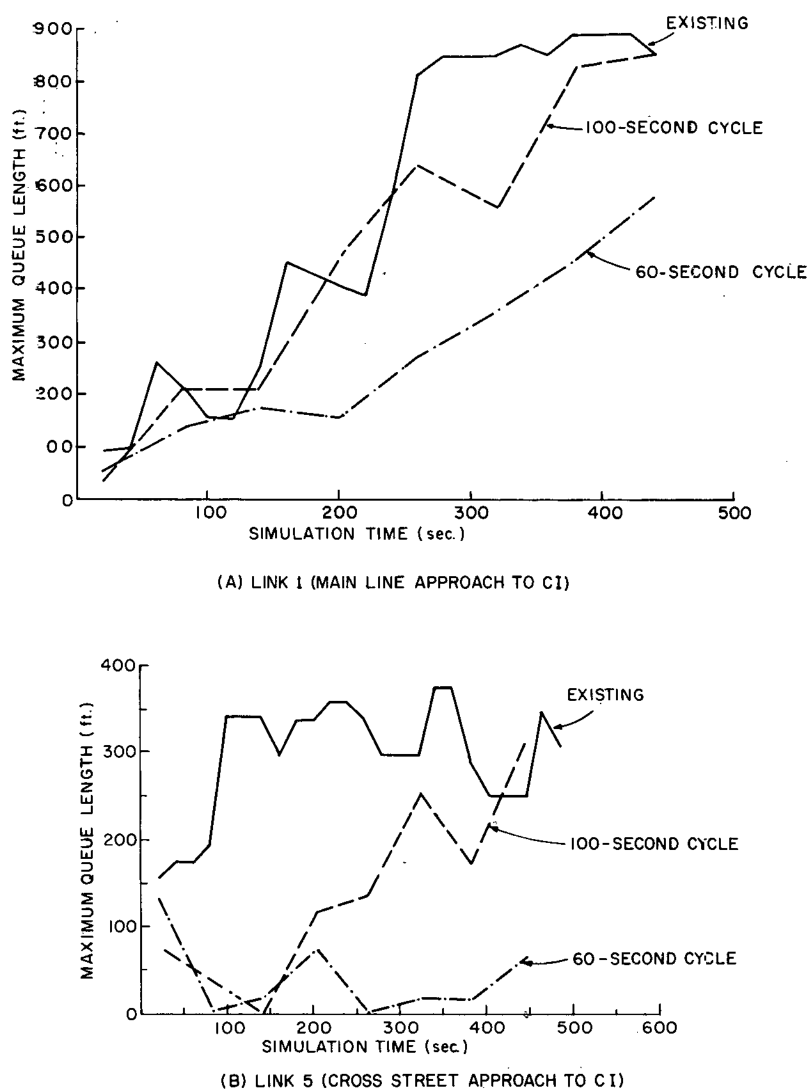


Figure H-7. Test set 1—impact of cycle length.

TABLE H-2

SYSTEM PERFORMANCE FOR TEST SET 1

ALTERNATE	STOP/VEH	M/T	AV. SPEED	MEAN OCC.	AV. DELAY/VEH.	M/T ART	M/T X-ST	MEAN OCC. ART	MEAN OCC. X-ST
Existing	1.33	0.293	8.22	153.6	80.13	0.386	0.165	89.1	64.3
C=100 common	1.40	0.350	9.84	137.0	64.50	0.391	0.256	95.1	41.9
C=60 common	1.24	0.442	12.66	106.9	42.12	0.508	0.306	71.4	35.5

M/T = MOVING TIME / TRAVEL TIME

AV. SPEED = AVERAGE SPEED THROUGH SYSTEM

MEAN OCC. = MEAN OCCUPANCY (VEHICLES)

AV. DELAY/VEH. = AVERAGE DELAY/VEHICLES (SECOND)

ART = ARTERIAL

X-ST = CROSS STREET

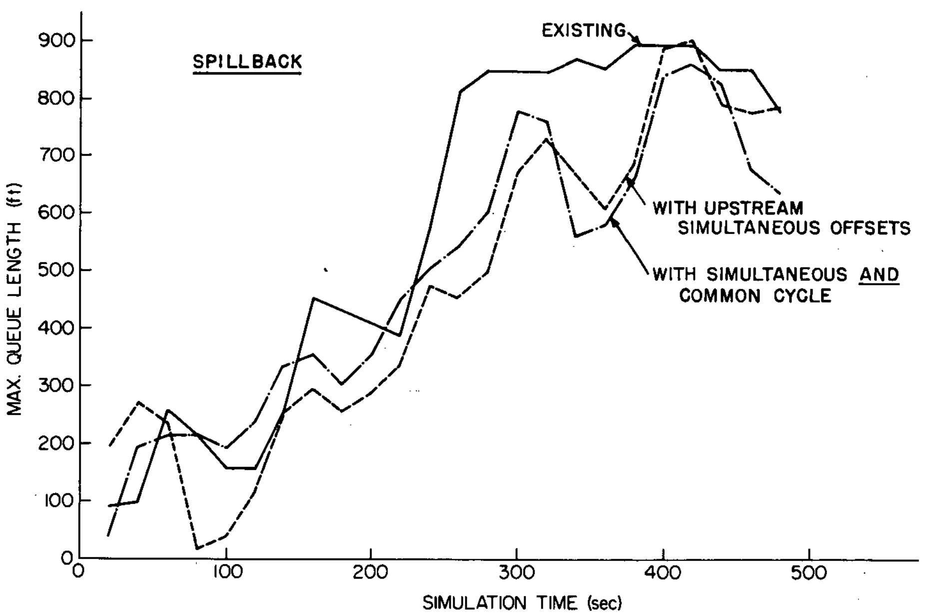
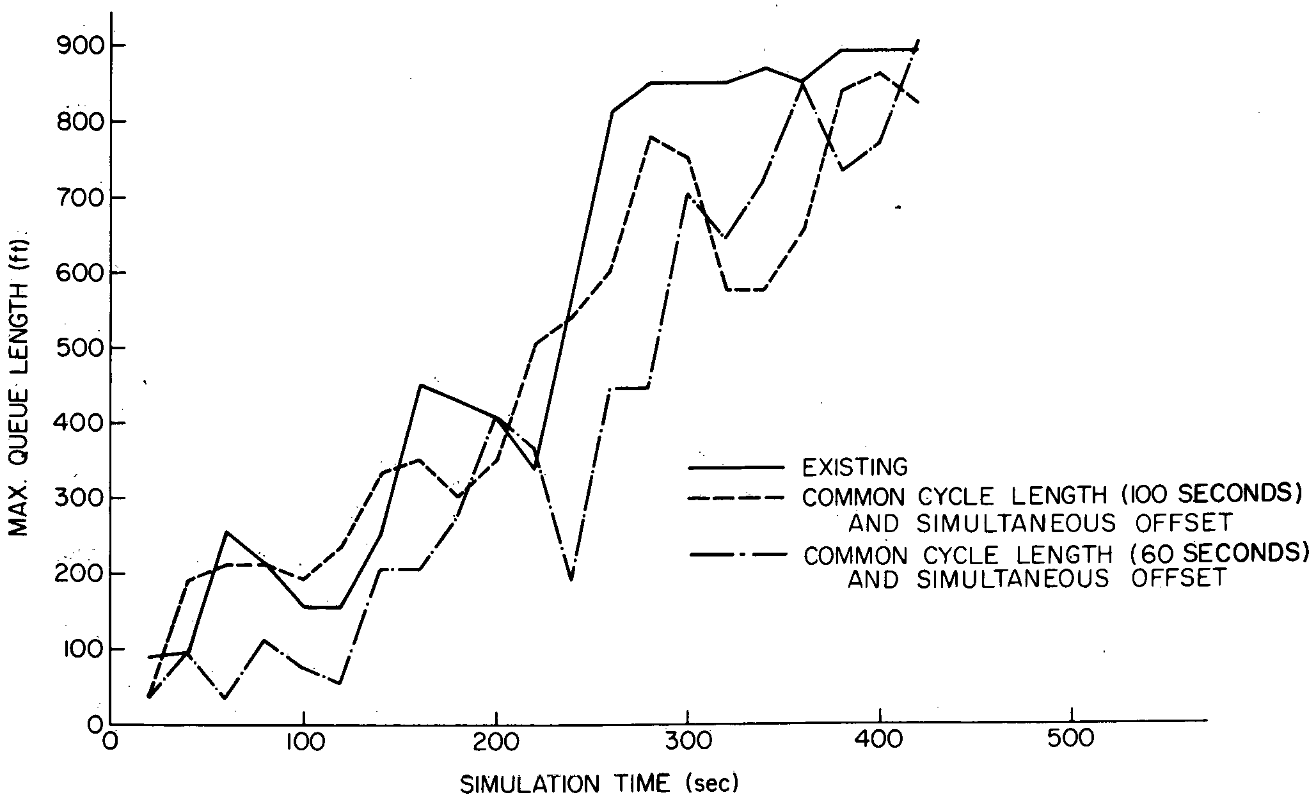


Figure H-8. Test set 2—impact of 3 different signal plans (link 1, main line approaching CI).

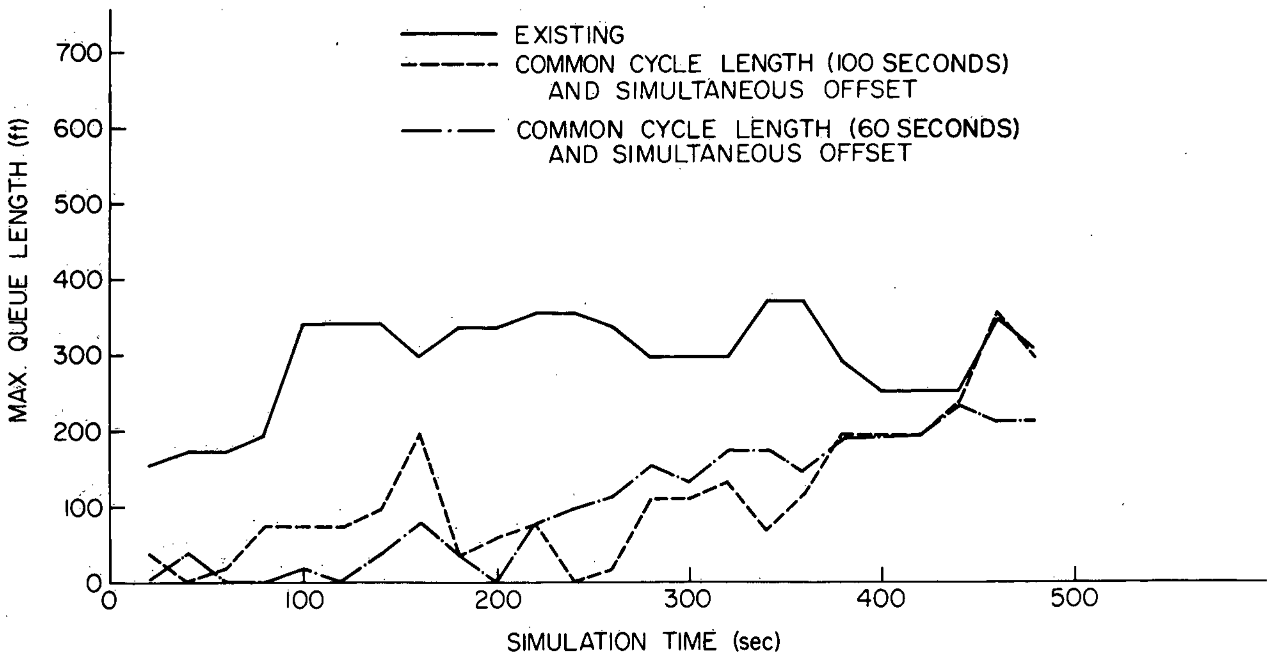
TABLE H-3

SYSTEM PERFORMANCE FOR TEST SET 2

ALTERNATE	STOP/VEH.	M/T	AV. SPEED	MEAN OCC.	AV. DELAY/VEH.	M/T ART	M/T X-ST	MEAN OCC. ART	MEAN OCC. X-ST
Existing	1.33	0.293	8.22	153.6	80.13	0.386	0.165	89.1	64.3
Simultaneous Upstream only	1.20	0.362	10.30	126.3	57.71				
Simultaneous <u>and</u> common cycle	1.17	0.361	10.22	127.9	60.29	0.419	0.245	85.0	43.0



(A) LINK 1 (MAIN LINE APPROACH TO CI)



(B) LINK 5 (CROSS STREET APPROACH TO CI)

Figure H-9. Test set 3—approaches to the CI.

TABLE H-4

SYSTEM PERFORMANCE FOR TEST SET 3

ALTERNATE	STOP/VEH.	M/T	AV. SPEED	MEAN OCC.	AV. DELAY/VEH.	M/T ART	M/T X-ST	MEAN OCC. ART	MEAN OCC. X-ST
Existing	1.33	0.293	8.22	153.6	80.13	0.386	0.165	89.1	64.3
Alternate I	1.17	0.361	10.22	127.9	60.29	0.419	0.245	85.0	43.0
Alternate II	1.18	0.435	12.15	109.9	44.99	0.456	0.349	77.8	32.1

Alternate I - 100 second cycle at all notes with left turn phase at CI

Alternate II - 60 second cycle at all notes without left turn phase at CI

TABLE H-5

SYSTEM PERFORMANCE FOR TEST SET 4

ALTERNATE	STOP/VEH.	M/T	AV. SPEED	MEAN OCC.	AV. DELAY/VEH.	M/T ART	M/T X-ST	MEAN OCC. ART	MEAN OCC. X-ST
Common C=60	1.41	.393	12.22	89.7	43.28	.355	.450	54.0	35.7
C=84 at CI	1.53	.356	11.26	98.4	49.81	.311	.429	61.2	37.4
C=120 at CI	1.37	.305	9.54	99.4	58.52	.249	.386	69.2	40.3

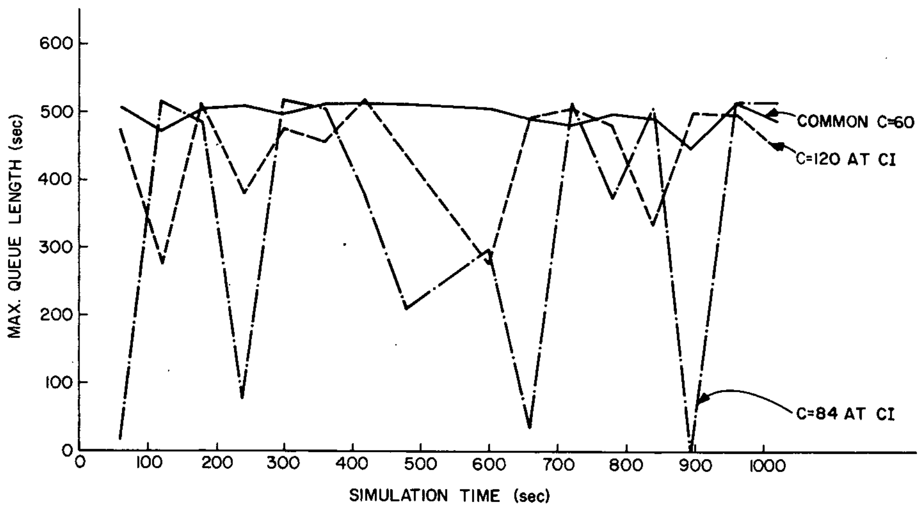


Figure H-10. Test set 4—maximum queue length at CI.

expected. However, a reduction of delay is experienced by the upstream cross-street traffic because of the additional green and the delay of oversaturation.

Test Set 7

The concept of uniform bandwidth was also tested on Flatbush Avenue Extension. When applied to the existing signal plan, only slight improvements were achieved in queue length and average delay; when applied to a common cycle length, however, a significant reduction was achieved in both queue length and average delay (Fig. H-13 and Table H-7). For either equal bandwidth or gradual change of bandwidth to be effective, a common cycle length appears necessary to eliminate arrivals of vehicles out of a

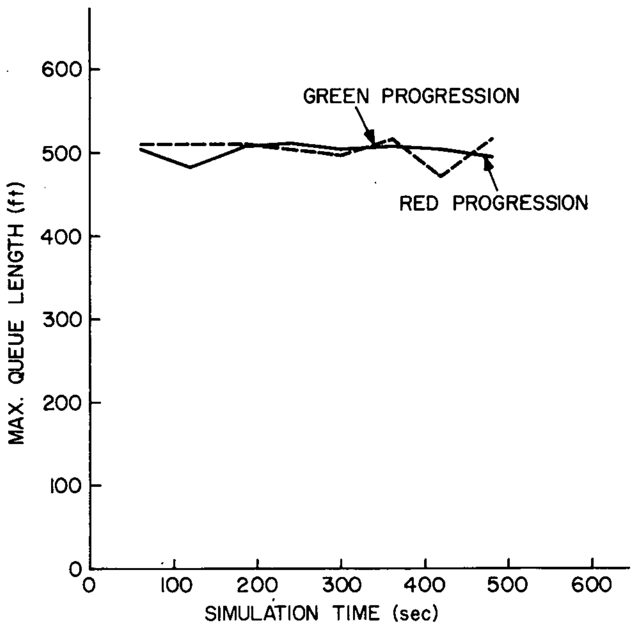


Figure H-11. Test set 5—queue length at the CI.

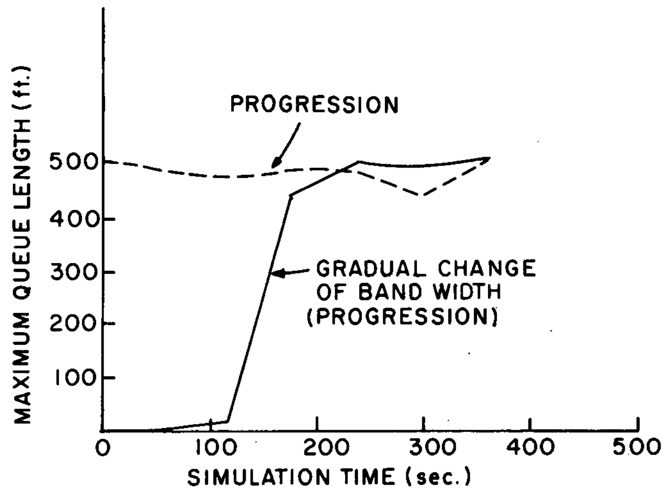


Figure H-12. Test set 6—queue length at the CI.

TABLE H-6

DELAY COMPARISON FOR TEST SET 6

	AVERAGE DELAY/VEH. (SEC)			
	SYS-TEM	UPSTREAM INT.	C. I.	UPSTREAM CROSS-STREET
Standard Progression	43.3	71.1	26.9	24.2
Gradual Change of Bandwidth	41.4	75.4	35.1	20.9

phase. The reason that the curve for uniform bandwidth with the existing cycle plan is much smoother than the curves for the other two plans is because the simulator used a larger interval within which to check the queue length.

Test Set 8

This test set compares the application of conventional and Gazis optimal settings to the problem of oversaturation. The intersection depicted in Figure H-4(C) is used (500 ft approach links). The fabricated demand functions and the minimum (total) delay settings are shown in Figure H-14.

As expected, neither the conventional nor the Gazis settings bring relief to the problem of oversaturation. Figure H-15 shows that the queue is not reduced, and spillback occurs. A summary of the system statistics is given in Table H-8.

Test Set 9

This set includes the testing of highly responsive control policies in an attempt to improve the critical intersection specifically (queue-actuated and Longley) and the system generally (smooth flow). The test situation is again Flatbush Avenue Extension.

Table H-9 summarizes the system performance for test set 9. For convenience, some results from test sets 1, 2, and 3 are included. It is apparent from the shading in this table that the saturation problem at this particular site can also be addressed by good minimal response policies.

Figure H-16 illustrates queue development approaching the CI for some of the policies cited. Figure H-17 presents the total system productivity for some of the policies. Note the very low productivity resulting from the smooth flow policy. To pursue this further, some comments are in order.

The smooth flow technique might appear to vastly improve the quality of traffic operations by almost total elimination of queues (this result was observed), but does so (at least in certain classes of situations, as shown in Fig. H-17) at the penalty of a markedly reduced productivity. The reasons for this can be explained by referring to Figure H-18.

The queues at the critical intersection are vehicles arriving at the end of the green at the critical intersection, due to staggering of the bandwidth. However, the V/C ratio at the upstream intersection is far less than unity; thus, the bandwidth contribution to queues is underutilized, and the result is smaller queues.

Delay to the nonpeak direction is greatly reduced because a progression is obtained for the nonpeak direction as a result of the reverse offset for the controlled direction. However, in this case, smooth flow has failed to aid the greatest system concern during congested operation: productivity. Apparently, the shifts in upstream links throttle the throughput, and this adverse impact "trickles down" to the CI in terms of fewer vehicles.

Queue-actuated control produces substantially lower delay through the system than the existing policy, as well as through the critical intersection. Examining the average delay for individual links, one sees that a greater reduction of delay is achieved at the critical intersection than at all other intersections combined. Reduction of delay at the critical intersection can be attributed to:

1. Shorter cycle length due to a short block length (link 5).
2. Increase of effective "green" due to the change from a four-phase to a two-phase signal. The existing signal at the CI has a 120-sec cycle with four phases. By applying "queue-actuated control" with $Q_{max} = 0.9$ (link length) and maximum green of 83 sec and minimum green of 30 sec, the effective cycle length varies between 70 and 100 sec except at the beginning of simulation.

Test Set 10

The network of Figure H-4(B) was considered, in this test, with saturation flow approaching the CI in both directions and with 0.8 saturation flow on the other routes. Blocks lengths are 600 ft. The turn conditions are 10 percent and 35 percent. Table H-10 summarizes the results for five signal policies, three of which are highly responsive. Figures H-19 and H-20 summarize the queue performance in the system under two of the cited policies, at the two turn rates.

When turn-ins are minor, queue-actuated control produces as much delay as does a conventional progression, while Longley's control strategy and a simultaneous offset plan are equally effective in reducing delay. Use of simultaneous offset, however, is more effective, on the average, in reducing queues and delaying the onset of oversaturation.

When turn-ins create "lock-in" of the successive links, queue-actuated control produces the least delay, because turn-ins always actuate "green" in favor of link 1 before any intersection blockage occurs. In other words, turn-ins cause an earlier initialization of "green" for link 1. Longley's control and simultaneous offset are more effective

TABLE H-7

SYSTEM PERFORMANCE FOR TEST SET 7

ALTERNATE	STOP/VEH.	M/T	AV. SPEED	MEAN OCC.	AV. DELAY/ VEH.
Existing	1.33	0.293	8.22	153.6	80.13
Uniform Band-width with Existing Cycle Length	1.34	0.299	8.33	148.2	76.16
Uniform Band-width with a common Cycle Length	1.22	0.352	10.01	133.1	62.76

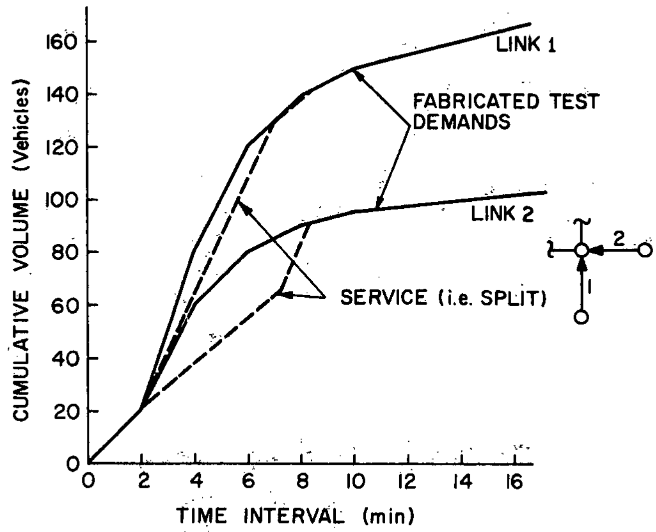


Figure H-14. Test set 8—demand functions and service pattern.

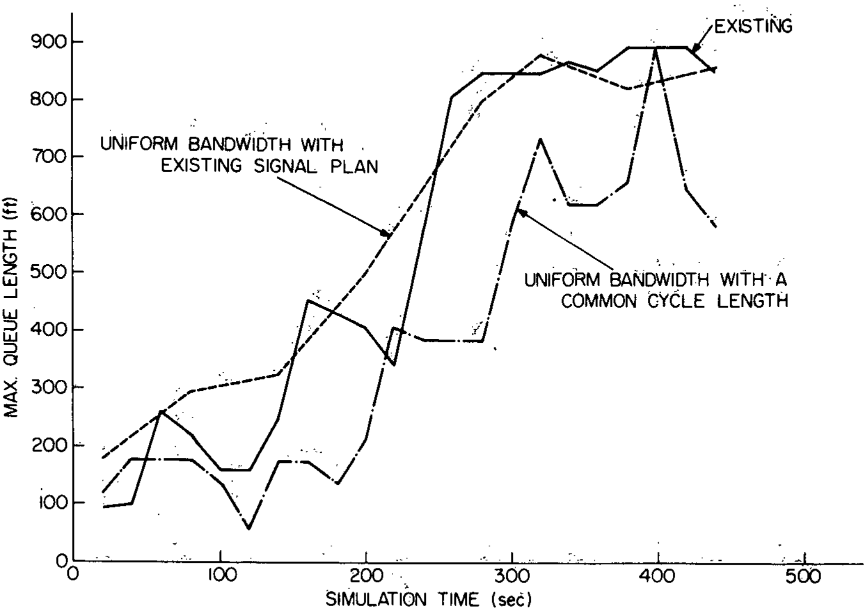


Figure H-13. Test set 7—queue length at main line approach to CI.

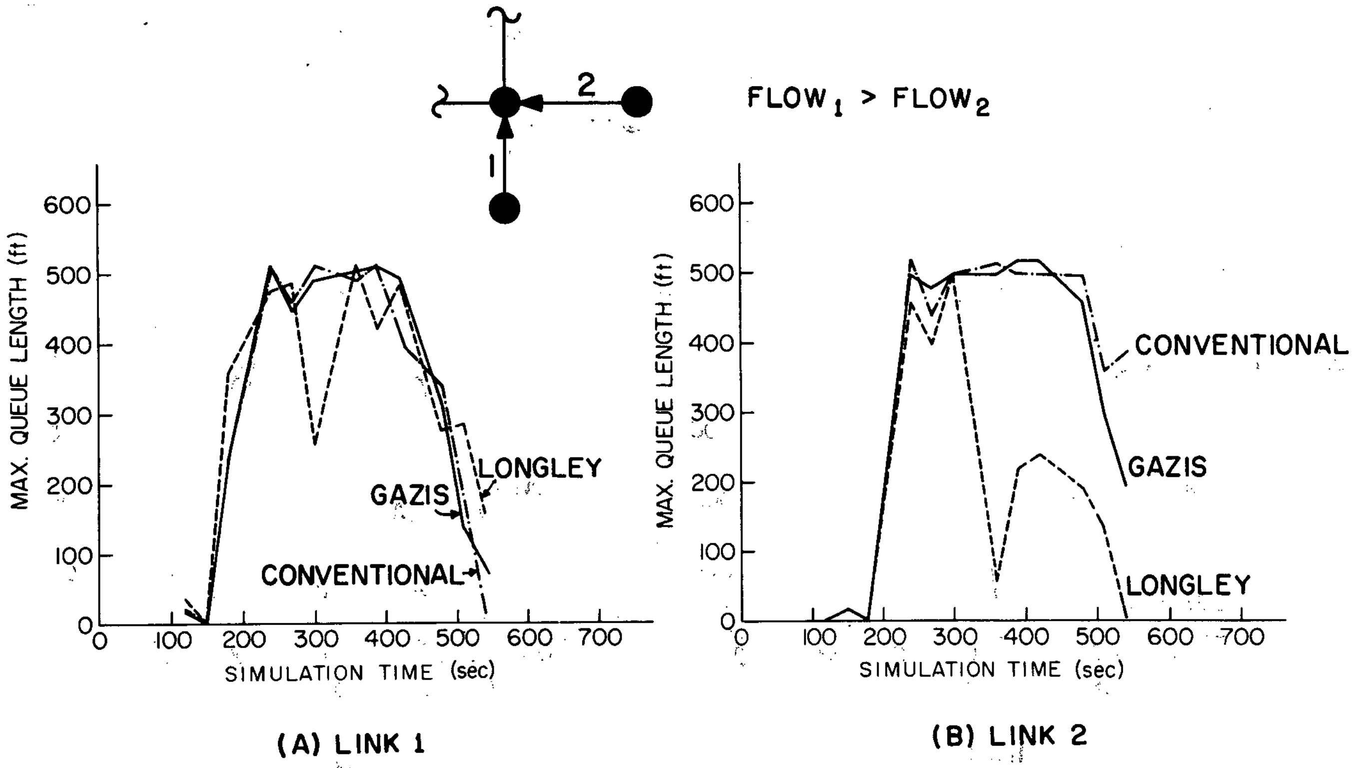


Figure H-15. Test set 8—approach queues.

TABLE H-8

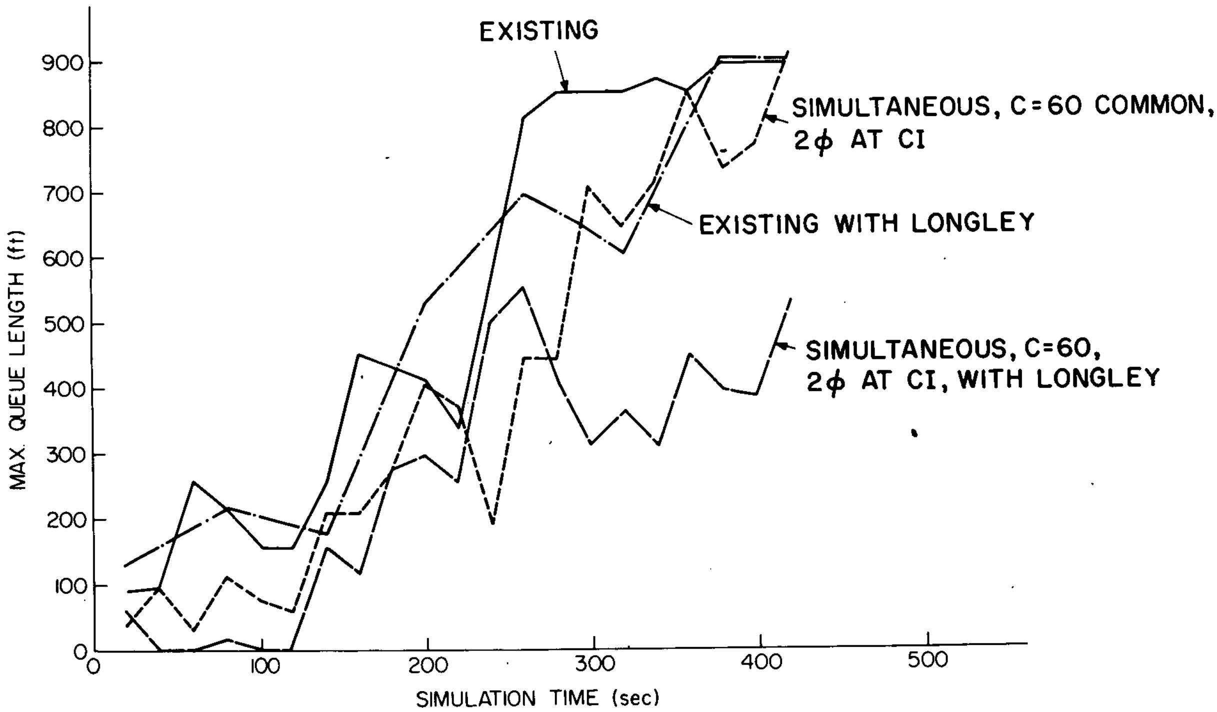
SYSTEM PERFORMANCE FOR TEST SET 8

ALTERNATE	STOP/VEH.	M/T	AV. SPEED	MEAN OCC.	AV. DELAY/VEH.
Conventional	0.91	.154	4.52	25.9	63.21
Gazis	0.92	.161	4.89	24.4	58.49
Longley	0.88	.182	5.64	21.2	49.49

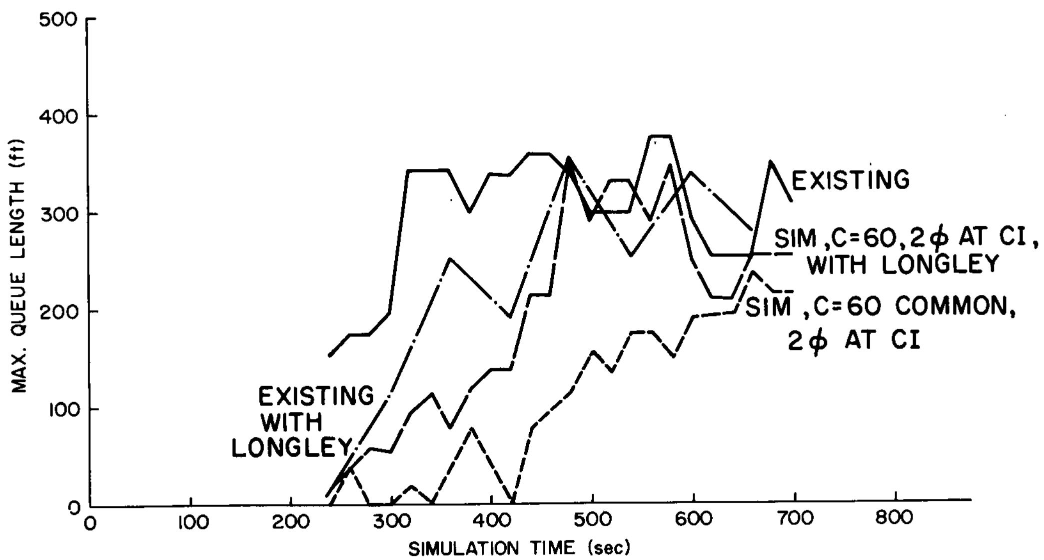
TABLE H-9

SYSTEM PERFORMANCE FOR TEST SET 9

ALTERNATES	DELAY/VEHICLE (SECONDS SYSTEM	CI	SYSTEM EXCLUDING CI	Highly Responsive	Minimal Response
Existing	80.1	77.4	21.1		▨
C = 60 Common only Change	42.1	-	-		▨
Queue-Actuated at CI only Change	49.1	39.5	15.2	▨	
Longley at CI only change	68.1	62.5	18.4	▨	
Simultaneous upstream Intersection only	57.7	55.1	22.6		▨
Simultaneous all, c = 60, 2ϕ at CI	45.0	-	-		▨
Simultaneous, C = 60, but Longley at CI	43.1	-	-	▨	



(A) LINK 1 (MAIN LINE APPROACH TO CI)



(B) LINK 5 (CROSS STREET APPROACH TO CI)

Figure H-16. Test set 9—approaches to the CI.

than is progression in reducing delay, but are less so than queue-actuated control.

Test Set 11

Flow on a one-way arterial was simulated to further study the effect of the queue-actuated control policy on other intersections. The network was simulated under a perfect progression with ideal traffic conditions (no turns, no trucks).

It is noted from Table H-11 that the average delay at the critical link is somewhat increased. As seen in other results, the queue-actuated control policy generally increases average delay except under extremely heavy flow

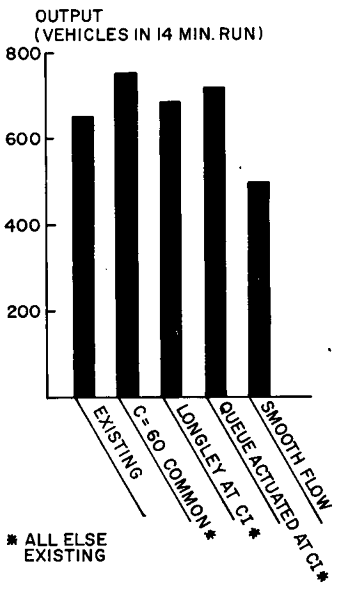


Figure H-17. Productivity of some policies.

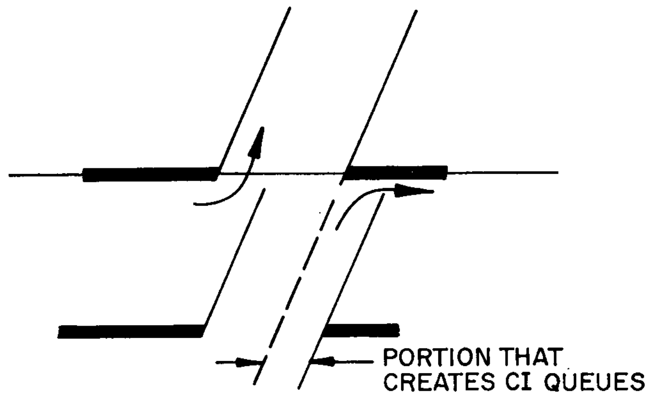


Figure H-18. Smooth flow mechanism.

TABLE H-10

SYSTEM PERFORMANCE FOR TEST SET 10

ALTERNATE	DELAY/VEHICLE (SECONDS) 10% TURNS SYSTEM	10% TURNS CI	35% TURNS SYSTEM	35% TURNS CI
Simple Progression	117.4	62.8	277.4	164.0
Simple Progression, with Longley at CI	98.4	55.7	168.2	94.1
Simultaneous	92.3	53.2	152.2	90.2
Simultaneous, with Longley at CI	98.2	-	169.6	-
Queue-Actuated at CI	169.2	101.6	134.1	79.6

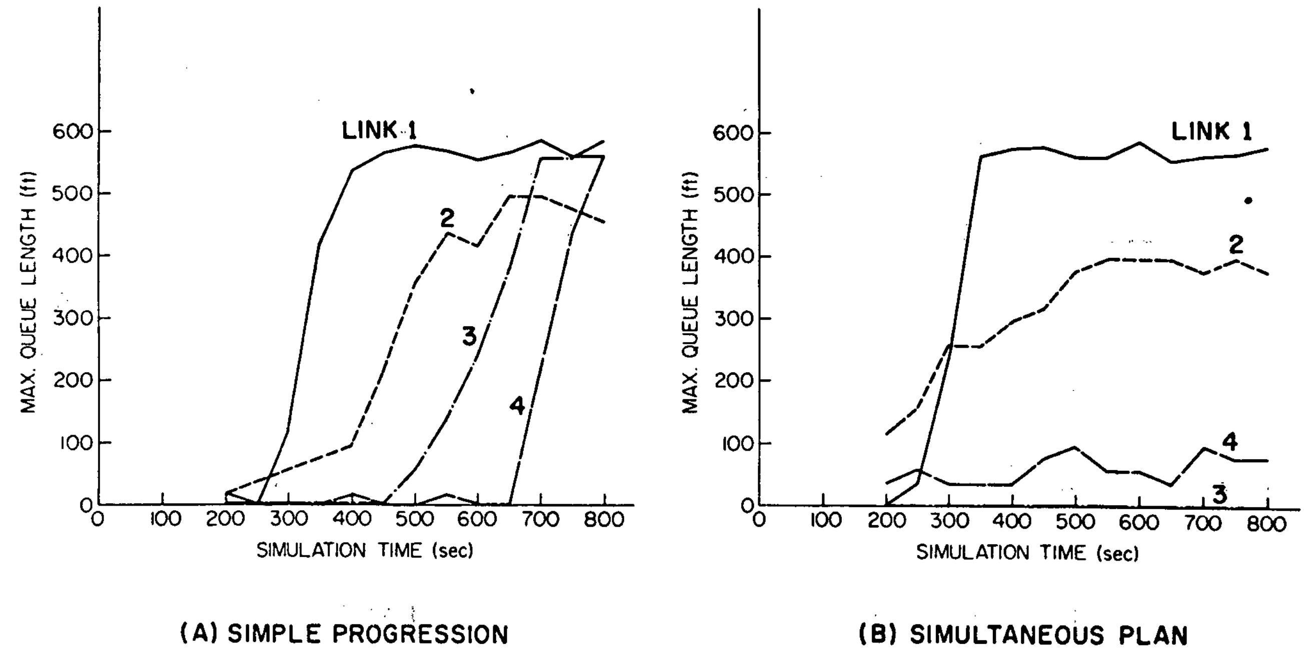


Figure H-19. Test set 10 with 10% turns.

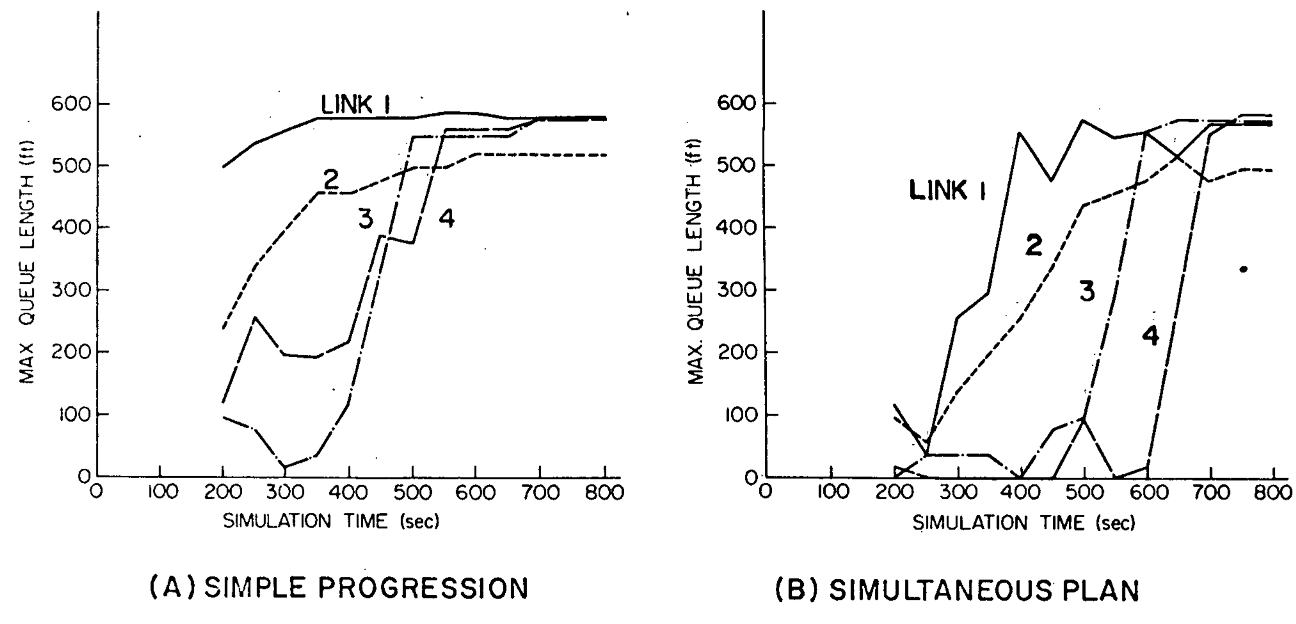


Figure H-20. Test set 10 with 35% turns.

conditions. Delay can be reduced by selecting the proper values of Q_{max}, and maximum and minimum green. This will, in effect, determine the effective cycle length. Table H-11 presents results for $Q_{max} = 0.9$ link length, and maximum and minimum green values of 44 and 20 sec, respectively. Use of $Q_{max} = 0.5$ link length does not reduce delay because the cycle length becomes too short to be practical.

TABLE H-11

COMPARISON FOR SELECTED LINKS OF TEST SET 11

Link	Average Delay/Vehicle (Second)	
	Progression	Queue-Actuated
Upstream	37.2	33.8
	33.9	33.1
C. I.	35.1	39.3
Downstream	3.5	16.4
	2.4	6.5
	2.4	2.9
System	43.3	53.6

APPENDIX I

SOME PROGRAMS DEVELOPED TO AID ANALYSIS OF THE UTCS-1 SIMULATOR RESULTS

In addition to the standard outputs of the UTCS-1 simulator (NETSIM) and of the associated post-processor routine, the option exists to output some statistics every few seconds. It was found useful to output such statistics every 5 sec and to write special programs to manipulate that output.

DATA AVAILABLE

Figure I-1 is a flow diagram of the UTCS-1 outputs and the summaries generated from the periodic output statistics tape file. The card formats of the periodic output statistics are shown in Figure I-2.

It should be noted that the tape on which these statistics are stored can, and commonly does, contain outputs from many runs, separated only by the distinctive header cards ending in "99." There are no distinct files (one for each run) on this tape. The creation of the WRMXXX. SAT file, where XXX denotes a case number (i.e., a UTCS-1 run), allows one to look at the time history of a particular link, within a particular case, in a convenient fashion.

PROGRAM A—QUEUE HISTORY

Program A was written in the interest of studying the queue development on a given link for the two control policies (signal or nonsignal remedies). A typical output of the program is presented in Figure I-3. It represents the results of two runs. Note, also the following:

1. The numbers represent the lane of the queue, as shown in Figure I-2: "1" for curb lane, "4" for left-turn bay, and "5" for right-turn bay.
2. The release signal state is indicated by symbols on the horizontal axes: (-) for red, (1) for green, and (/) for amber. Other symbols are used for special phases.
3. The maximum queue is delimited by a row of asterisks.
4. The queue cannot overflow 30, by virtue of an interval limit in the program.

The flow diagram (Fig. I-4) shows how this program is used. The input format is described in Table I-1. A program listing is contained in Figure I-5. Figure I-6 shows the input structure for all examples contained in this appendix.

PROGRAM B—OCCUPANCY HISTORY

Queues sometimes do not indicate the extent to which the link is occupied because queues are defined in terms of virtually stopped vehicles. Accordingly, in response to this need, Program B was written as the occupancy counterpart to Program A, the queue history.

Figure I-7 is a typical output. Figure I-8 indicates the use of the program. Table I-2 describes the input format. Figure I-9 is a listing of Program B.

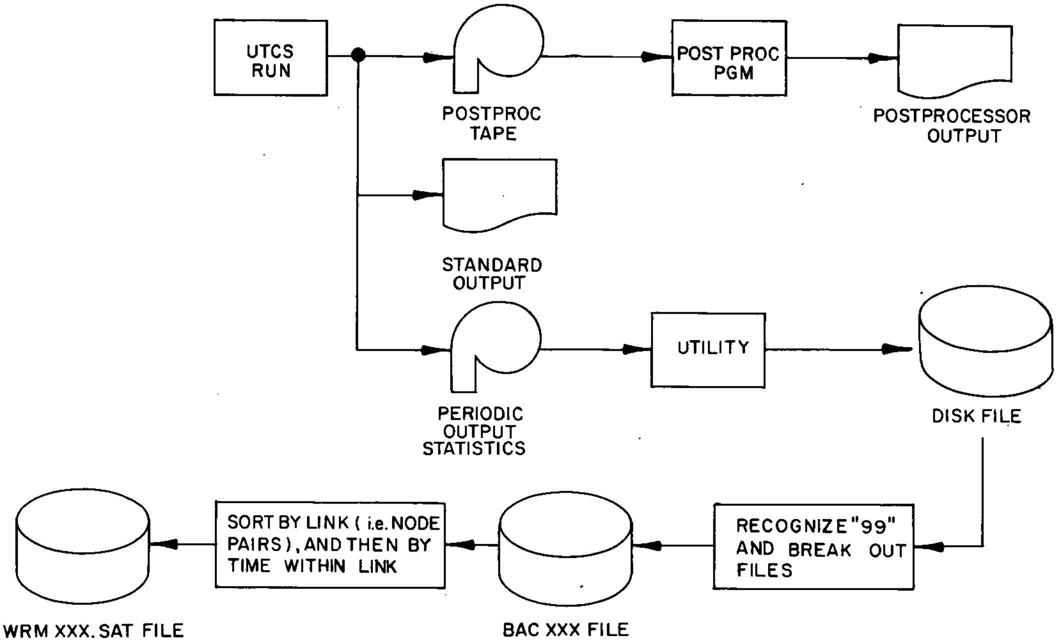


Figure I-1. UTCS-1 simulator outputs, including options, and use thereof.

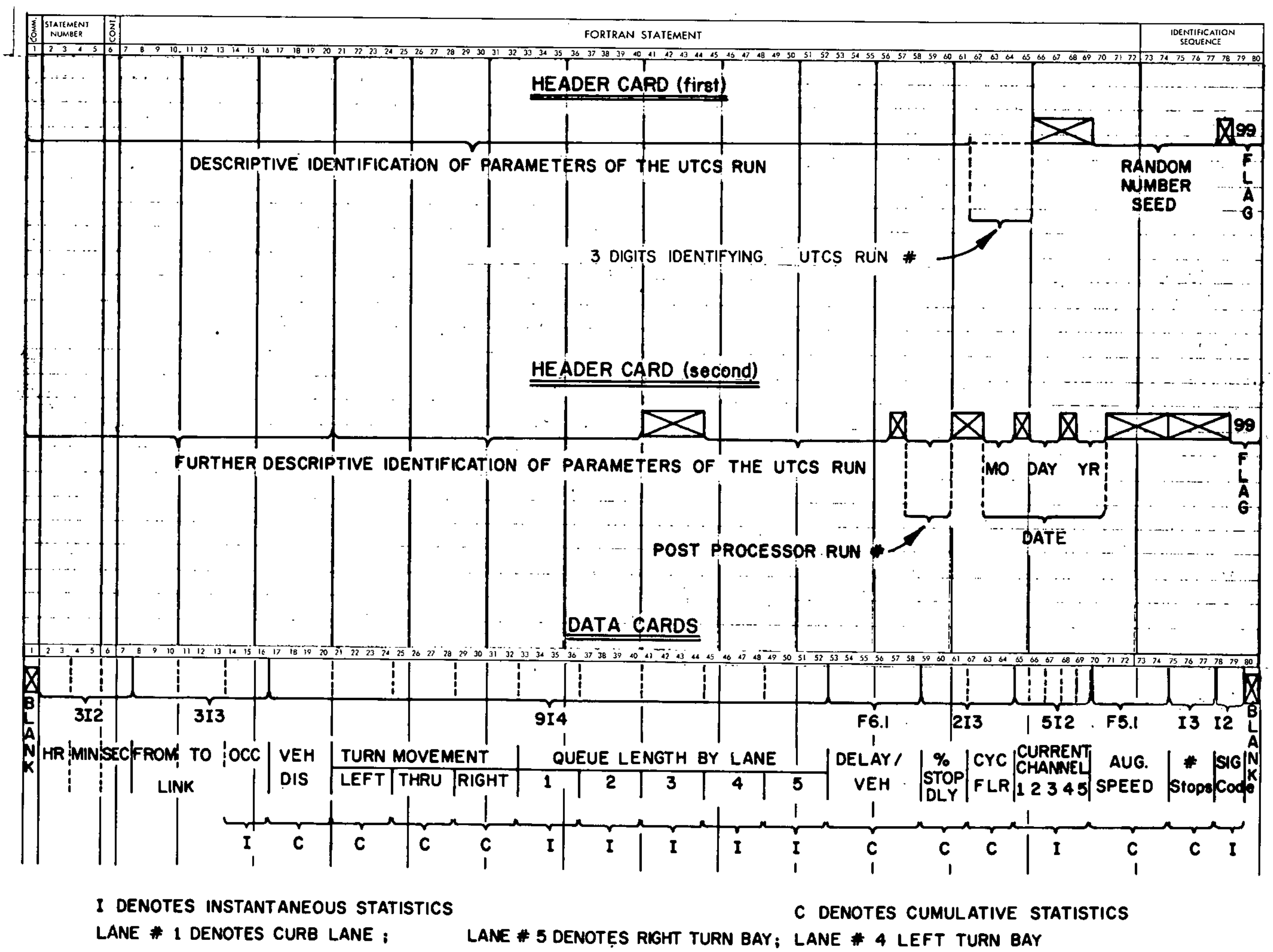


Figure I-2. UTCS periodic output statistics input data format.

```
CCMPARISON OF QUEUES....TWO CASES

CYCLE LENGTH=    60  SECONDS                COMPARISON RUN  SEPT 27 1974

STEPS OF     5  SECONDS

LINK( 33, 23) IS UNDER STUDY

CASE 1 IS RUN  304                          CASE 2 IS RUN  314

000000000011111111112222222222           000000000011111111112222222222
012345678901234567890123456789           012345678901234567890123456789
|2                      *                |2                      *
| 12                    *                |12                     *
|   1 2                 *                |  2 1                  *
|   1  2                *                |  2 1                  *
|   1  2                *                |  2 1..                *
|   1  2                *                |  2 1                  *
|   1  2'               *                |  2 1                  *
-  1  2                 *                - 2 1                   *
-1   2                  *                -  1                    *
-   2                   *                -                       *
-2                      *                -                       *
|                       *                |                       *
|    2                  *                |2                      *
|    1 2                *                | 2 1                   *
|     1 2               *                |    2 1                *
|     1 2               *                |     21                *
|     1 2               *                |     21                *
|     1 2               *                |     21                *
|     1 2               *                |     21                *
-    1 2                *                -    21                 *
-  1 2                  *                -  21                   *
-  2                    *                -21                     *
-                       *                -                       *
|                       *                |                       *
|                       *                |1                      *
|  2                    *                | 2 1                   *
|1   2                  *                |   2 1                 *
|  1  2                 *                |    21                 *
|    12                 *                |    21                 *
|    12                 *                |    21                 *
|    12                 *                |    21                 *
-   12                  *                -   21                  *
-1  2                   *                - 21                    *
-1                      *                -1                      *
-                       *                -                       *
|                       *                |                       *
|                       *                |2                      *
|  2                    *                | 21                    *
|  12                   *                |    21                 *
|   12                  *                |    2 1                *
|   12                  *                |    2 1                *
|   12                  *                |    2 1                *
|   12                  *                |    2 1                *
-  12                   *                -    2 1                *
-2                      *                -    2  1               *
                                               2   1             *
                                                2
```

Figure I-3. Sample output of Program A (queue history comparison).

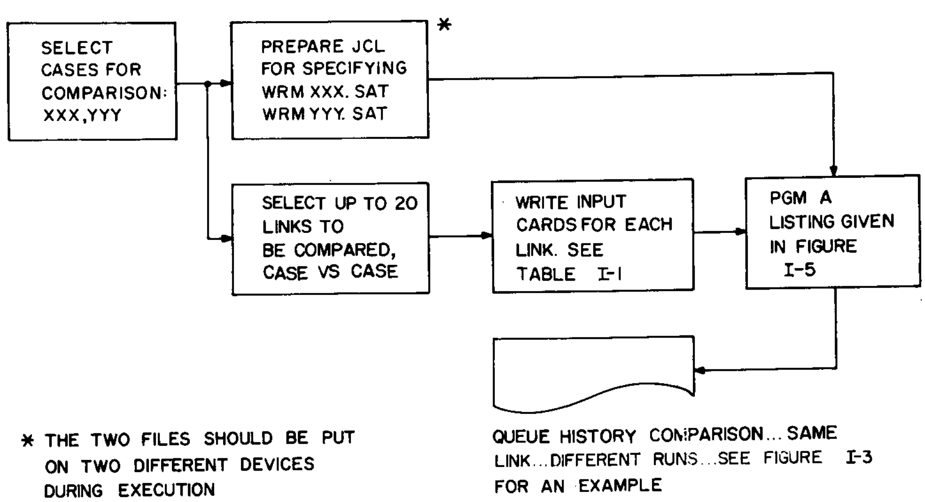


Figure I-4. Use of Program A for queue history comparison.

TABLE I-1

INPUT FORMAT FOR PROGRAM A

INPUT	DESCRIPTION	FORMAT
CASE 1	First Case: XXX	I5
CASE 2	Second Case: YYY	I5
CYCLE	Cycle Length (Seconds)	I5
-	-	5X
DATE	Date of Analysis or other Note	3A4
-	-	8X
STEP	Steps to be shown; Must be Multiple of 5* and Submultiple of Cycle Length; Give in Seconds	I2
-	-	8X
NIN	Origin Node of Link ("From")	I3
NØUT	Destination Node of Link ("To")	I3
MAX	Number of Vehicles that can fit in Single Lane; Program Assures 30 or less in its Dimensions**, 20 ft/vehicle generally	I3

* Assumes Periodic Output Statistics Given Every 5 Seconds.

** Output will show asterisk just <u>beyond</u> maximum extent if less than 30, just <u>at</u> maximum extent if extent <u>is</u> 30.

PROGRAM A

```
304  314   60    SEPT 27 1974        5        033023 23  2
```

PROGRAM B

```
304  314   60    SEPT 27 1974        5        033023 23  2
```

PROGRAM C

```
306      60     5
1    023024   600  111
2    033023  1800  111
3    820021  1800  010
5    021022  1800  001
5    022023   600  110
6    024025   600  100
6    025026   600  010
4    802012  1800  000
4    012022  1800  100
4    022032   600  010
4    843033  1800  010
4    023013   600  010
4    804014  1800  000
4    014024   600  100
4    024034   600  010
4    845035  1800  000
4    035025   600  001
4    025015   600  010
```

NOTE: THESE INPUTS CORRESPOND TO THE SAMPLE OUTPUTS IN FIGURES I-3, I-7, AND I-11

Figure I-6. Sample inputs for Programs A, B, and C.

```
      C        PGM A...SATNET...LANE COMPARISONS (QUEUE)
      C
      C
            INTEGER HDA(30),HDB(30),SA(30),SB(30),STAR,LAB(5),BLANK,PLUS
            DATA HDA/10*0,10*1,10*2/
            DO 100 I=1,10
            HDB(I)=I-1
            HDB(I+10)=I-1
        100 HDB(I+20)=I-1
            DATA BLANK/1H /
            DATA LAB/1H1,1H2,1H3,1H4,1H5/
            DATA STAR/1H*/
            DATA PLUS/1H+/
            INTEGER SIGA,SIGB,SIGCOD(10)
            DATA SIGCOD/1H/,1H-,1H|,1H*,1H*,1H>,1H*,1H*,1H*,1H*/
            INTEGER CASE1,CASE2,CYCLE,DATE(3),STEP,QA(5),QB(5),NIN,NOUT,MAX
            DO 3000 NCARD = 1,20
      C        DO LOOP PERMITS COMPARISON FOR UP TO 20 LINKS BETWEEN A PAIR
      C        OF UTCS RUNS
            READ(5,10,END=4000) CASE1,CASE2,CYCLE,DATE,STEP,NIN,NOUT,MAX
         10 FORMAT(3I5,5X,3A4,8X,I2,8X,3I3)
            WRITE(6,15)
         15 FORMAT('1')
            WRITE(6,20) CYCLE,DATE,STEP,NIN,NOUT,CASE1,CASE2
         20 FORMAT(T30,'COMPARISON OF QUEUES....TWO CASES'//T30,'CYCLE LENGTH=
           *',I5,2X,'SECONDS',T70,'COMPARISON RUN',2X,3A4//T30,'STEPS OF',I5,2
           *X,'SECONDS'//T30,'LINK(',I3,',',I3,') IS UNDER STUDY'//T30,'CASE 1
           * IS RUN',I5,T70,'CASE 2 IS RUN',I5//)
            MAXPLS=MAX+1
            IF(MAXPLS.GT.30) MAXPLS=30
            NX=CYCLE/STEP
            NW=NX*STEP
            IF(NW.EQ.CYCLE) GO TO 22
            WRITE(6,21)
         21 FORMAT(T40,'TERMINATION...STEP IS NOT A SUBMULTIPLE OF CYCLE')
            STOP
         22 CONTINUE
            READ(3,23) DUMA,DUMB
         23 FORMAT(A4/A4)
            READ(4,23) DUMA,DUMB
            DO 30 I=1,50000
            READ(3,25) NA,NB
         25 FORMAT(7X,2I3)
            IF(NA.EQ.NIN.AND.NB.EQ.NOUT) GO TO 31
         30 CONTINUE
         31 BACKSPACE 3
            DO 40 I=1,50000
            READ(4,25) NA,NB
            IF(NA.EQ.NIN.AND.NB.EQ.NOUT) GO TO 41
         40 CONTINUE
         41 BACKSPACE 4
            WRITE(6,50) HDA,HDA
         50 FORMAT(T30,30I1,T70,30I1)

            WRITE(6,50) HDB,HDB
            DO 60 I=1,30
            SA(I)=BLANK
         60 SB(I)=BLANK
            NS=STEP/5
            DO 1000 IX=1,50000
            DO 69 I=1,NS
         69 READ(3,70) NA,NB,QA,SIGA
         70 FORMAT(7X,2I3,19X,5I4,26X,I1)
            IF(NA.NE.NIN.OR.NB.NE.NOUT) GO TO 2000
            DO 71 I=1,NS
         71 READ(4,70) NA,NB,QB,SIGB
            IF(NA.NE.NIN.OR.NB.NE.NOUT) GO TO 2000

            SIGA=SIGA+1
            SIGB=SIGB+1
            NA=QA(1)+1
            NB=QA(2)+1
            NC=QA(5)+1
            IF(NA.GT.30) NA=30
            IF(NB.GT.30) NB=30
            IF(NC.GT.30) NC=30
            SA(NA)=LAB(1)
            SA(NB)=LAB(2)
            SA(NC)=LAB(5)
            SA(MAXPLS)=STAR
            ND=QB(1)+1
            NE=QB(2)+1
            NF=QB(5)+1
            IF(ND.GT.30) ND=30
            IF(NE.GT.30) NE=30
            IF(NF.GT.30) NF=30
            SB(ND)=LAB(1)
            SB(NE)=LAB(2)
            SB(NF)=LAB(5)
            SB(MAXPLS)=STAR
            SA(1)=SIGCOD(SIGA)
            SB(1)=SIGCOD(SIGB)
         80 CONTINUE
            WRITE(6,90) SA,SB
         90 FORMAT(T30,30A1,T70,30A1)
            SA(NA)=BLANK
            SA(NB)=BLANK
            SA(NC)=BLANK
            SB(ND)=BLANK
            SB(NE)=BLANK
            SB(NF)=BLANK
       1000 CONTINUE
       2000 CONTINUE
            REWIND 3
            REWIND 4

       3000 CONTINUE
       4000 CONTINUE
            STOP
            END
```

Figure I-5. Program A listing.

```
COMPARISON OF OCCUPANCY....TWO CASES
CYCLE LENGTH=   60  SECONDS                    COMPARISON RUN  SEPT 27 1974
STEPS OF     5  SECONDS
LINK( 33, 23) IS UNDER STUDY
LANES=  2.
CASE 1 IS RUN  304                             CASE 2 IS RUN  314
```

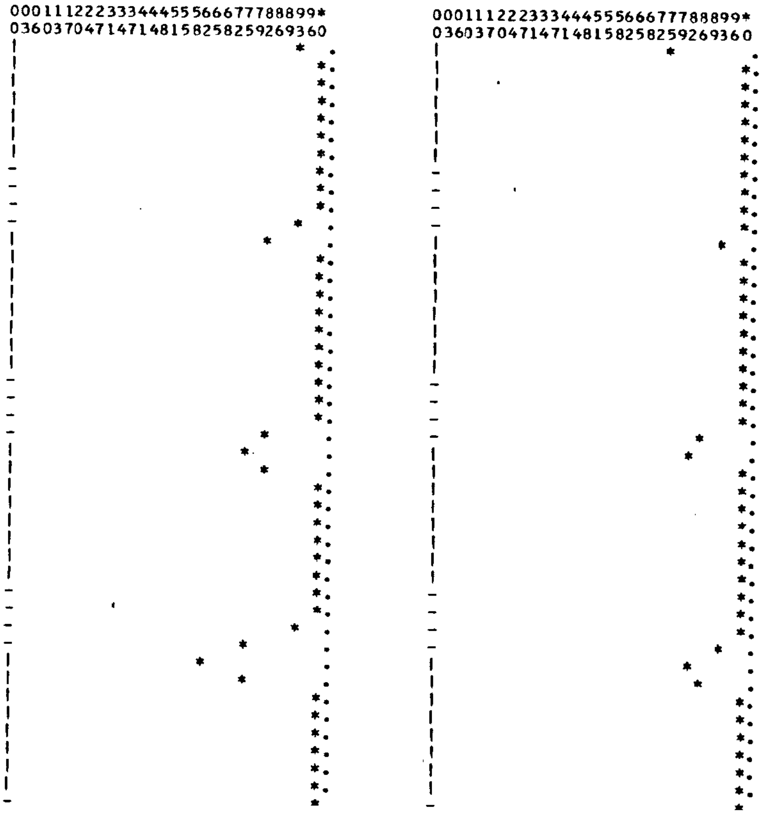


Figure I-7. Sample output of Program B (occupancy history comparison).

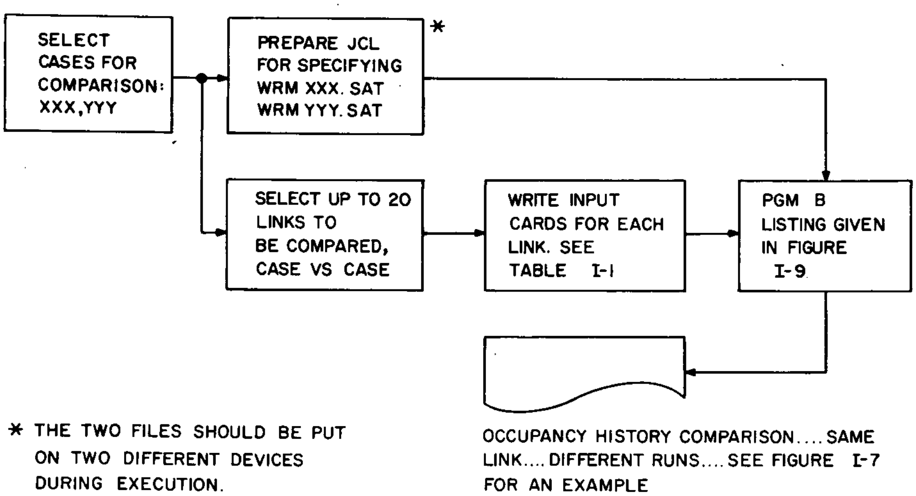


Figure I-8. Use of Program B for occupancy history comparison.

```
FORTRAN IV G LEVEL  21                  MAIN              DATE = 75064          11/56/00            PAGE 0001

      C          PGM B...SATNET...LANE COMPARISONS (OCCUPANCY)
      C
      C
            REAL LANES,OCCA,OCCB,XMAX,ANS(30),MAXOCC
            INTEGER HDA(30),HDB(30),SA(30),SB(30),STAR,LAB(5),BLANK,PLUS
            DO 100 I=1,30
            ANS(I)=((FLOAT(I)-1.0)/29.0)*100.0
            HDA(I)=ANS(I)/10.0
        100 HDB(I)=ANS(I)-HDA(I)*10
            DATA BLANK/1H /
            DATA LAB/1H1,1H2,1H3,1H4,1H5/
            DATA STAR/1H*/
            DATA PLUS/1H+/
            INTEGER SIGA,SIGB,SIGCOD(10)
            DATA SIGCOD/1H/,1H-,1H|,1H*,1H>,1H*,1H*,1H*,1H*/
            INTEGER CASE1,CASE2,CYCLE,DATE(3),STEP,QA(5),QB(5),NIN,NOUT,MAX
            DO 3000 NCARD=1,20
      C
      C          DO LOOP PERMITS COMPARISON FOR UP TO 20 LINKS
      C          BETWEEN A PAIR OF UTCS RUNS
      C
            READ(5,10,END=4000) CASE1,CASE2,CYCLE,DATE,STEP,NIN,NOUT,
           *MAX,LANES
         10 FORMAT(3I5,5X,3A4,8X,I2,8X,3I3,2X,F1.0)
            WRITE(6,15)
         15 FORMAT('1')
            WRITE(6,20) CYCLE,DATE,STEP,NIN,NOUT,LANES,CASE1,CASE2
         20 FORMAT(T30,'COMPARISON OF OCCUPANCY....TWO CASES'//T30,
           *'CYCLE LENGTH=',I5,2X,'SECONDS',T70,'COMPARISON RUN',2X,
           *3A4//T30,'STEPS OF',I5,2X,'SECONDS'//T30,
           *'LINK(',I3,',',I3,') IS UNDER STUDY'//T30,'LANES=',F4.0//
           *T30,'CASE 1 IS RUN',I5,T70,'CASE 2 IS RUN',I5//)
            XMAX=MAX
            NX=CYCLE/STEP
            NW=NX*STEP
            MAXOCC=XMAX*LANES
            IF(NW.EQ.CYCLE) GO TO 22
            WRITE(6,21)
         21 FORMAT(T40,'TERMINATION...STEP IS NOT A SUBMULTIPLE OF CYCLE')
            STOP
         22 CONTINUE
            READ(3,23) DUMA,DUMB
         23 FORMAT(A4/A4)
            READ(4,23) DUMA,DUMB
            DO 30 I=1,50000
            READ(3,25) NA,NB
         25 FORMAT(7X,2I3)
            IF(NA.EQ.NIN.AND.NB.EQ.NOUT) GO TO 31
         30 CONTINUE
         31 BACKSPACE 3
            DO 40 I=1,50000
            READ(4,25) NA,NB
            IF(NA.EQ.NIN.AND.NB.EQ.NOUT) GO TO 41

FORTRAN IV G LEVEL  21                  MAIN              DATE = 75064          11/56/00            PAGE 0002

         40 CONTINUE
         41 BACKSPACE 4
            WRITE(6,50) HDA,HDA
         50 FORMAT(T30,30I1,T70,30I1)
            WRITE(6,50) HDB,HDB
            DO 60 I=1,30
            SA(I)=BLANK
         60 SB(I)=BLANK
            NS=STEP/5
            DO 1000 IX=1,50000
            DO 69 I=1,NS
         69 READ(3,70) NA,NB,OCCA,SIGA
         70 FORMAT(7X,2I3,F3.0,62X,I1)
            IF(NA.NE.NIN.OR.NB.NE.NOUT) GO TO 2000
            DO 71 I=1,NS
         71 READ(4,70) NA,NB,OCCB,SIGB
            IF(NA.NE.NIN.OR.NB.NE.NOUT) GO TO 2000


            SIGA=SIGA+1
            SIGB=SIGB+1
            NA=(OCCA/MAXOCC)*100.0
            IF(NA.GT.30) NA=30
            IF(NA.EQ.0) NA = 1
            SA(NA)=STAR
            ND=(OCCB/MAXOCC)*100.0
            IF(ND.GT.30) ND=30
            IF(ND.EQ.0) ND = 1
            SB(ND)=STAR
            SA(1)=SIGCOD(SIGA)
            SB(1)=SIGCOD(SIGB)
         80 CONTINUE
            WRITE(6,90) SA,SB
         90 FORMAT(T30,30A1,'.',T70,30A1,'.')
            SA(NA)=BLANK
            SB(ND)=BLANK
       1000 CONTINUE
       2000 CONTINUE
            REWIND 3
            REWIND 4
       3000 CONTINUE
       4000 CONTINUE
            STOP
            END
```

Figure I-9. Program B listing.

TABLE I-2

INPUT FORMAT FOR PROGRAM B

INPUT	DESCRIPTION	FORMAT
CASE 1	First Case: XXX	I5
CASE 2	Second Case: YYY	I5
CYCLE	Cycle Length (Seconds)	I5
-	-	5X
DATE	Date of Analysis or Other Note	3A4
-	- -	8X
STEP	Steps to be shown; Must be Multiple of 5* and Submultiple of Cycle Length; Give in Seconds	I2
-	-	8X
NIN	Origin Node of Link ("From")	I3
NØUT	Destination Node of Link ("To")	I3
MAX	Number of Vehicles that can fit in Single Lane; Program Assures 30 or less in its Dimensions**; 20 ft/vehicle generally	IØ

* Assumes Periodic Output Statistics Given Every 5 Seconds.

** Output will show asterisk just beyond maximum extent if less than 30, just at maximum extent if extent is 30.

PROGRAMS C AND D—LINK GROUPINGS (MOVEMENT GROUPINGS)

It is sometimes of interest to accumulate statistics on the basis of certain links or segments of links. For instance, one may wish to know the total delay, total stops, and productivity of a grouping, as shown in Figure I-10, and these statistics may be desired on an interval-by-interval basis (not cumulative) for an interval aggregation to be specified.

Program C was written for these purposes; a typical output is shown in Figure I-11. Note that six groups must be specified (see Table I-3). No link may be counted in more than one group, nor may different parts (i.e., movements) of a link appear in different groups.

It may also be desired to reformat the indicated output so that each group may be studied individually. This is accomplished by Program D, after a SORT is executed on the output of Program C. An example is given in Figure I-12.

Figure I-13 summarizes the use of Programs C and D. Table I-3 describes the input format. Figures I-14 and I-15 contain listings of Programs C and D, respectively.

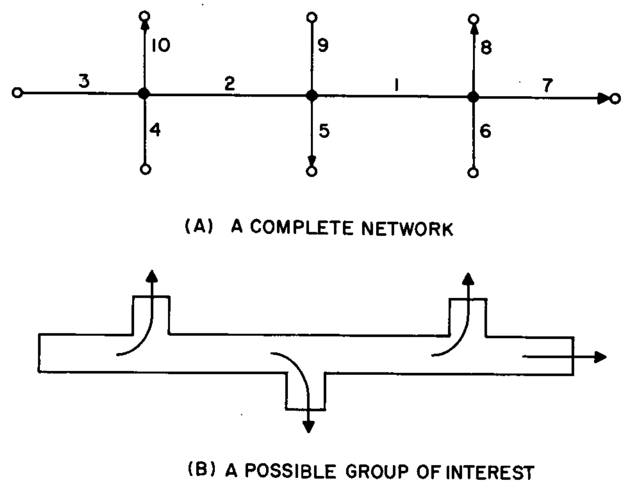


Figure I-10. Input for Program C.

INTERVAL	GROUP	SYSTEM DISCHARGE	TOTAL DELAY	TOTAL TRAVTIME	TOTAL STOPS	CYCLE FAILURE	STATISTICS PER DISCHAR VEHICLE DELAY	TRAVTIME	STOPS
1	1	30.	647.	1435.	20.	0.	21.56	47.83	0.67
1	2	31.	648.	1974.	12.	0.	20.89	63.68	0.39
1	3	40.	0.	0.	0.	0.	0.0	0.0	0.0
1	4	142.	2614.	6572.	51.	2.	18.41	46.28	0.36
1	5	36.	784.	3268.	19.	0.	21.77	90.77	0.53
1	6	36.	907.	2299.	29.	0.	25.21	63.87	0.81

Figure I-11. Sample output of Program C (certain statistics for link groups).

INTERVAL	GROUP	SYSTEM DISCHARGE	TOTAL DELAY	TOTAL TRAVTIME	TOTAL STOPS	CYCLE FAILURE	STATISTICS PER DISCHAR VEHICLE DELAY	TRAVTIME	STOPS
1	1	30.	647.	1435.	20.	0.	21.6	47.8	0.7
2	1	28.	850.	1712.	34.	0.	30.3	61.1	1.2
3	1	30.	1383.	2506.	31.	1.	46.1	83.5	1.0
4	1	30.	1890.	3297.	30.	1.	63.0	109.9	1.0
5	1	28.	2357.	4110.	28.	1.	84.2	146.8	1.0
6	1	24.	2526.	4175.	24.	1.	105.3	174.0	1.0
7	1	25.	2403.	3833.	25.	1.	96.1	153.3	1.0
8	1	24.	2505.	4362.	24.	1.	104.4	181.8	1.0
9	1	19.	2972.	4421.	19.	1.	156.4	232.7	1.0
10	1	21.	2581.	4304.	21.	1.	122.9	204.9	1.0
1	2	31.	648.	1974.	12.	0.	20.9	63.7	0.4
2	2	18.	699.	2076.	13.	0.	38.8	115.3	0.7
3	2	17.	507.	1914.	12.	0.	29.8	112.6	0.7
4	2	18.	350.	1615.	7.	0.	19.4	89.7	0.4
5	2	18.	332.	1131.	12.	1.	18.4	62.8	0.7
6	2	18.	1545.	2928.	28.	1.	85.8	162.7	1.6
7	2	16.	2677.	4825.	19.	1.	167.3	301.6	1.2
8	2	17.	3563.	5668.	17.	1.	209.6	333.4	1.0

Figure I-12. Sample output of Program D.

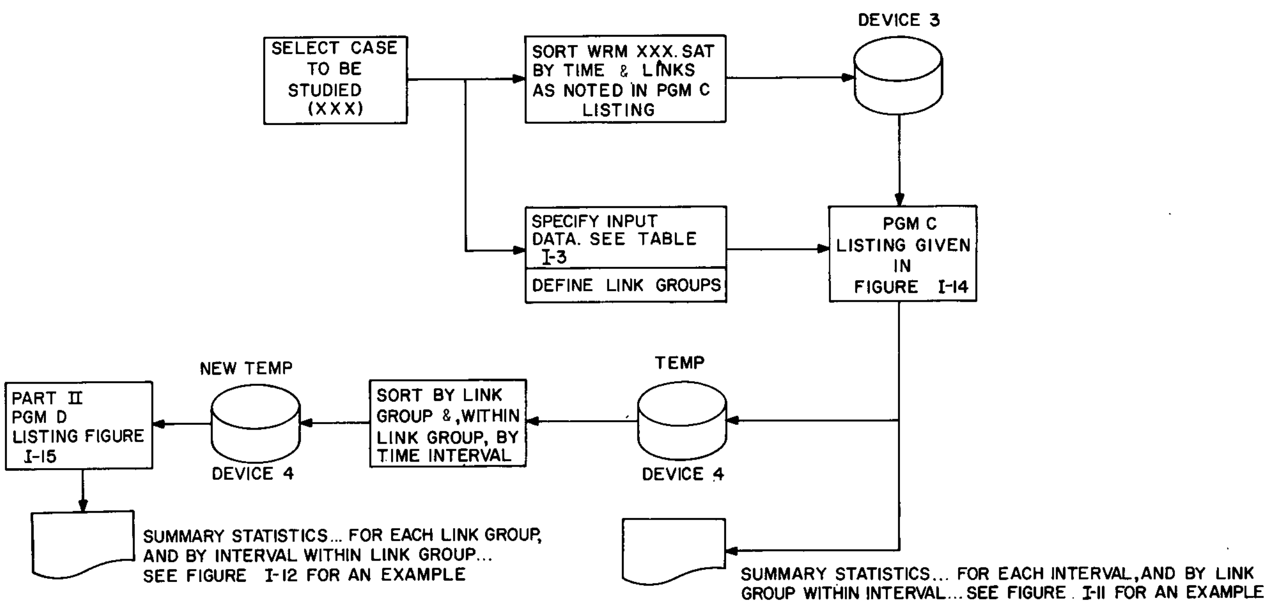


Figure I-13. Use of Program C for summary statistics.

```
C          PGM C...SATNET
C
C          SUMMARIZES CERTAIN STATISTICS BY GROUPS OF LINKS
C
C          DEVICE 3 IS WRMXXX.SAT SORTED ON A TEMP AS (2,6,A,8,6,A)
C
C          DEVICE 4 IS A TEMP FILE
C
C
      INTEGER CASE,INTER,STEP,LIST(6,50),X(6),MOVE(3),GRP,LINK,LEN
      REAL OUT(6,8),PREV(6,5),IN(8),NOW(6,5)
C     DO 9000 IJK = 1, 30
      READ(5,100) CASE,INTER,STEP
  100 FORMAT(I3,2X,I5,3X,I2)
      WRITE(6,110) CASE, INTER, STEP
  110 FORMAT('1',T5,'CASE',I5,2X,'TO FOLLOW, PGM 12 OUTPUT',5X,'PERIODS
     1OF',I5,2X,'SECONDS, ORIGINAL DATA IN STEPS OF',I4,2X,'SECONDS'//T5
     2,100('*')//)
      NA = INTER/STEP
      NC = 0
      DO 200 I=1, 50
C
C          THIS DO LOOP CONSTRAINS THE NUMBER OF LINKS IN THE SYSTEM
C          UNDER STUDY TO BE 50 OR LESS
C
      READ(5,120,END=300) GRP,LINK,LEN,MOVE
  120 FORMAT(I1,3X,I6,I5,2X,3I1)
      NC = NC + 1
      LIST(1,NC) = LINK
      LIST(2,NC) = GRP
      LIST(3,NC) = LEN
      DO 130 K=1,3
      KA=K+3
  130 LIST(KA,NC)=MOVE(K)
  200 CONTINUE
  300 WRITE(6,310)((LIST(I,J),I=1,6),J=1,NC)
  310 FORMAT(10X,I10,I5,I10,2X,I1,2X,I1,2X,I1)
      WRITE(6,320) NC
  320 FORMAT('0',10X,I2,2X,'TOTAL LINKS'//)
      NE = NC - 1
      DO 350 ND=2, NC
      NF=ND-1
      DO 340 J=ND,NC
      IF(LIST(1,J).GT.LIST(1,NF)) GO TO 340
      DO 330 I=1,6
      X(I)=LIST(I,NF)
      LIST(I,NF)=LIST(I,J)
  330 LIST(I,J)=X(I)
  340 CONTINUE
  350 CONTINUE
      WRITE(6,310)((LIST(I,J),I=1,6),J=1,NC)
      WRITE(6,320) NC
C
C               *******
C         CODE * ALL * LINKS THAT ARE IN THE NETWORK AND ON 5 SEC TAPE
C               *******
C
      DO 410 I=1,6
      DO 400 J=1,8
  400 OUT(I,J)=0.0
      DO 410 J=1,5
      NOW(I,J)=0.0
  410 PREV(I,J)=0.0
      NQ = 0
      DO 2000 IABC = 1,50000
      NSTEP = (NA-1)*NC
      IF(NSTEP.EQ.0) GO TO 425
      READ(3,420,END=2100) (DUM,I=1,NSTEP)
  420 FORMAT(A4)
  425 CONTINUE
      DO 500 I=1,NC
      READ(3,430,END=2100) NCHK,IN
  430 FORMAT(7X,I6,3X,4F4.0,20X,F6.1,3X,F3.0,5X,F5.1,F3.0)
      IF(LIST(1,I).NE.NCHK) WRITE(6,440)
  440 FORMAT(T40,'WARNING....OUT OF SYNC'/)
      J = LIST(2,I)
      DO 450 K=2,4
  450 NOW(J,1) = NOW(J,1) + IN(K)*LIST(K+2,I)
      NOW(J,2) = NOW(J,2) + IN(5) * IN(1)
      IF(IN(7).NE.0.0)
     *NOW(J,3) = NOW(J,3) + FLOAT(LIST(3,I)) * IN(1) / IN(7)
      NOW(J,4) = NOW(J,4) + IN(8)
      NOW(J,5) = NOW(J,5) + IN(6)
  500 CONTINUE
      DO 600 I=1,6
      DO 600 J=1,5
      OUT(I,J) = NOW(I,J) - PREV(I,J)
      PREV(I,J) = NOW(I,J)
  600 NOW(I,J) = 0.0
      DO 700 I=1,6
      XDEN = OUT(I,1)
      IF(XDEN.EQ.0.0) GO TO 700
      DO 690 K = 6,8
  690 OUT(I,K) = OUT(I,K-4)/XDEN
  700 CONTINUE
      NQ = NQ + 1
      WRITE(6,800)
  800 FORMAT('1',T2,'INTERVAL',T12,'GROUP',T22,'SYSTEM',T32,'TOTAL',
     *T42,'TOTAL',T52,'TOTAL',T72,'STATISTICS PER DISCHAR VEHICLE',
     *T62,'CYCLE'/T22,'DISCHARGE',T32,'DELAY',T42,'TRAVTIME',T52,
     *'STOPS',T62,'FAILURE',T72,'DELAY',T82,'TRAVTIME',T92,'STOPS'//)
      WRITE(6,810)((NQ,J,(OUT(J,I),I=1,8)),J=1,6)
  810 FORMAT(T2,I3,I10,5F10.0,3F10.2/)
      IF(NQ.EQ.1) WRITE(4,815) NC
  815 FORMAT(50X,I3)
      WRITE(4,820)((NQ,J,(OUT(J,I),I=1,8)),J=1,6)
  820 FORMAT(2I2,5F10.0,3F6.2)
 2000 CONTINUE
 2100 CONTINUE
 9000 CONTINUE
      REWIND 4
      STOP
      END
```

Figure I-14. Program C listing.

TABLE I-3

INPUT FORMAT FOR PROGRAM C

	INPUT		FORMAT
FIRST CARD	CASE	Case Being Studied	I3
	-	-	2X
	INTER	Interval to be aggregated (Seconds)	I5
	-	-	3X
	STEP	Original Interval (Step) that File has; Usually 5 seconds	I2
ALL OTHERS	NOTE:	There must be six groups and there must be one card per link. A link cannot be divided into two groups. Two separate analyses are needed.	
	GRP	The Group to which the link is Assigned; Pick From 1 to 6	I1
	-	-	3X
	LINK	The "From - To" Code (Nodes), expressed as a six-digit integer	I6
	-	-	2X
	MOVE	Zero-One Code for three movements: Left, thru, Right; Are they to be included in link characteristics (i. e., Group Characteristics) or Not? "1" if "Yes"; "0" if "No"	3I1

```
      C        PGM D...SATNET
      C
      C        DEVICE 4 IS TEMPXXX CREATED FROM PGM C SORTED AS (3,2,A,1,2,A)
      C
      C
            REAL OUT(8)
            READ(4,815) NC
            DO 1000 J=1,7
            WRITE(6,800)
            DO 900 K=1,NC
            READ(4,820,END=2000) NQ,J,OUT
        900 WRITE(6,810) NQ,J,OUT
       1000 CONTINUE
        800 FORMAT('1',T2,'INTERVAL',T12,'GROUP',T22,'SYSTEM',T32,'TOTAL',
           *T42,'TOTAL',T52,'TOTAL',T72,'STATISTICS PER DISCHAR VEHICLE',
           *T62,'CYCLE'/T22,'DISCHARGE',T32,'DELAY',T42,'TRAVTIME',T52,
           *'STOPS',T62,'FAILURE',T72,'DELAY',T82,'TRAVTIME',T92,'STOPS'//)
        810 FORMAT(T2,I3,I10,5F10.0,3F10.1/)
        815 FORMAT(50X,I3)
        820 FORMAT(2I2,5F10.0,3F6.1)
       2000 STOP
            END
```

Figure I-15. Program D listing.

APPENDIX J

GUIDELINES FOR THE TREATMENT OF TRAFFIC CONGESTION ON STREET NETWORKS

INTRODUCTION

Many communities now experience severe traffic congestion on the street systems of central business districts or other high-activity centers. Their development trends, typically toward increasing land-use density in such areas, may lead to worsened conditions in the future. Thus, the application of traffic engineering techniques becomes increasingly important. Unfortunately, the operations and control techniques that function effectively with lesser traffic volumes deteriorate when saturation exists for any length of time.

The information and guidelines contained herein are intended to aid the practicing traffic engineer in effectively addressing the operational problems of congested networks, by reminding him of available options, by uncovering some subtleties that can be overlooked, and by presenting him with quantitative insight into the relative benefits of the various options. In this way, an appreciation can be obtained of when various options are effective and/or necessitated, of how much impact can be expected, and what combinations are most effective.

First and foremost, these guidelines are intended to aid the traffic engineer in making operational improvements to facilitate traffic movement in an existing situation. It is important to recognize that they do not address the benefits of land use management to optimize the demand for travel, nor do they address the rescheduling of demand via staggered work hours and other measures, nor do they (extensively) investigate reorganization of the traffic stream via increased transit usage. More basically, they do not address the policy issue of whether vehicular traffic movement should be facilitated.

There is no intent to minimize those issues that are not addressed. Indeed, it is the land use pattern that shapes the traffic distribution and is the most basic and most important root cause of traffic congestion. It is strongly recommended that the traffic engineer and the transportation and urban planners stress this basic reality, and plan accordingly. Such actions and/or changes, however, tend to have benefits in a different time frame than is of concern in these guidelines.

Likewise, the problems of demand management and of transit operations are usually the concern of others, not those who are involved with the operational concerns of these guidelines. These techniques have their own intrinsic merit and whether or not significant improvements can be achieved within the framework of these guidelines, recourse to such other measures in transportation systems management will frequently be warranted. It is, however, beyond the scope of these guidelines to assess or even to address the effectiveness of such measures.

STRUCTURE OF THE GUIDELINES

These guidelines for the treatment of traffic congestion on street networks were proposed under Project 3-18(2) of the National Cooperative Highway Research Program (NCHRP). They were developed by means of analytical studies, simulation tests, and field studies. The field studies stressed traffic characteristics under congested flow, and not tests of specific policies.

The guidelines are directed for use by the practicing traffic engineer in coping with congestion and saturation. It has not been possible to develop unequivocal statements of when particular techniques or combinations of techniques are better than others. However, certain categorical statements can be made, and a logical analysis framework specified.

Because terminology may differ among various individuals, and, indeed, a review of the literature points to this, it was considered useful to start with the inclusion of a section on precise definitions of the terminology associated with traffic congestion/saturation to avoid the likelihood of confusion or misinterpretation of these terms and characteristics as they are used throughout the guidelines. The next section provides an exposition of the key elements in the recommended "plan of attack" in treating problems of congestion and saturation. This is followed by a section that identifies the range of solutions that are available. The remainder of the manual is devoted to details on the enumerated candidate treatments categorized under the following headings: "Minimal Response Signal Remedies—Intersection," "Minimal Response Signal Remedies—System," "Highly Responsive Signal Control," "Enforcement and Prohibition," "Turn Bays and Other Nonsignal Remedies," "Major Lane Arrangement," and "Disruptions to the Traffic." It is important to note that these latter sections should be used only within the framework of the earlier discussion.

TERMS AND CHARACTERISTICS

Three groups of terms must be defined:

1. Congestion-related terms (congestion, stable saturation, unstable saturation, and oversaturation).
2. Terms relating to the types of oversaturation, as categorized by cause.
3. Characteristics distinguishing the productivity of an intersection and the perception of congestion.

Throughout, it will be recognized that the problem focuses on an intersection that has become critical. A *critical intersection* (CI) is an intersection whose capacity is the limiting value of a segment of roadway or of an entire system. There may be more than one critical intersection in a system. It is most common that there is one critical intersection locally that has affected several blocks in one or more directions, giving the appearance of areawide congestion.

Congestion-Related Definitions

Oversaturation represents a condition that occurs when queues of vehicles fill entire blocks and interfere with the performance of adjacent upstream intersections.

A review of the literature, sparse as it is, indicates an imprecision in the use and definition of the terms congestion, saturation, or oversaturation. It was not unusual to find all of these terms used to describe the same level of traffic operations. More precise definitions are developed herein to allow for more exact description of the various traffic performance states. The specific definitions used have been constructed around a queue formation mechanism and, as such, they relate to the extent and growth of queues. It is recommended that they be adopted in all uses and classification of congestion.

The following traffic performance definitions are all described in terms of one approach to a signal. As such, they refer to the capacity of one signal's green time for the approach under consideration. The various regions are shown in Table J-1.

The term *uncongested operations* refers to a situation where there is no significant queue formation. Traffic performance may range from very low demand per cycle to conditions where the demand is a substantial fraction of the capacity value. Short queues may occasionally form toward the upper end of this performance range but do not last for any length of time because the capacity of the signal to process traffic exceeds the demands being placed on it.

Congested operations characterizes the entire range of operations which may be experienced when traffic demand approaches and/or exceeds the capacity of the signal. As such, the term describes conditions as diverse as those where demand only slightly exceeds capacity for a time, and minimal queues form, to those where demand for service is so great as to cause segments of the system to jam because of extensive queue development. Within the context of the problem, it is necessary to segment the realm of congested operations into two major categories: saturated and oversaturated operations.

Saturated operations describes that range of congestion wherein queues form, but their adverse effects on the traffic in terms of delay and/or stops are local. *Local effects* in this context means that traffic performance is only affected at the intersection at which the queue occurs, and that no other intersection's performance is affected by this queue.

Saturated operations has been further subcategorized into stable and unstable ranges. *Stable saturation* is considered to exist when a queue has formed and is not growing, and delay effects are local. *Unstable saturation* is considered to exist when a queue exists and is growing, and delay effects are still local.

Oversaturated operations is characterized as a situation wherein a queue exists, and it has grown to the point where the upstream intersection's performance is adversely affected. The upstream intersection's performance is considered to be adversely affected when its performance characteristics are determined by the performance of the downstream intersection from which the queue has developed, and not by its own physical and operational limitations.

Types of Oversaturation as Categorized by Cause

Intersection oversaturation may occur under varying conditions that can be classified into two groups: (1) nonrepetitive and (2) repetitive (the common peak period situations where demand exceeds capacity for periods of time). The nonrepetitive cases of oversaturation may be classified by their cause, which is usually apparent: accident, stalled vehicle, illegal parking, etc. In some situations, illegal parking is so common that this classification is somewhat artificial.

The repetitive, demand-based oversaturation situation may be classified by five basic types so as to better identity and reference the cause:

1. *Type I*—The critical intersection has a smaller green/cycle ratio than does the upstream intersection.

TABLE J-1

TRAFFIC DESCRIPTIONS

UNCONGESTED	CONGESTED		
	SATURATED		OVERSATURATED
	STABLE	UNSTABLE	
NO QUEUE FORMATION	QUEUE FORMATION, BUT NOT GROWING DELAY EFFECTS LOCAL	QUEUE FORMATION AND GROWING, DELAY EFFECTS STILL LOCAL; (A TRANSIENT STATE MAY BE ONLY OF SHORT DURATION)	QUEUE FORMATION AND GROWING TO POINT WHERE UPSTREAM INTERSECTION PERFORMANCE ADVERSELY AFFECTED

2. *Type II*—The critical intersection and the upstream intersection have the same green/cycle ratio. However, the capacity of the critical intersection is less than that of the upstream intersection because of factors other than the green/cycle ratio (such as turning movements and/or physical conditions).

3. *Type III*—Heavy turn-in movements from the upstream cross street fill up the entire link or a significant part of it during a red phase on the arterial and cause spillback on the arterial.

4. *Type IV*—This type of oversaturation of a critical intersection results from the signal offset between the critical intersection and its upstream intersection.

5. *Type V*—This is defined as being a combination of two or more of Types I through IV oversaturation.

Productivity vs. Quality of Performance

Data were collected (*J-11*) to investigate the impact of congestion, and the relative ability of various measures of effectiveness to depict this impact. From this work, it was determined that queue-based measures were the most meaningful indicators of congestion. For this reason, the previous definitions were written in terms of queue formation.

As part of this work, it was established that the quality of performance at an intersection, as measured by intersection crossing times (i.e., speed), degrades substantially before the productivity of the intersection, as measured by its throughput, does. Drivers respond to downstream queue size by going slower when crossing the intersection. Thus, they perceive the congestion. Moreover, they are more closely spaced than they would otherwise be. Still, the productivity of the intersection is not necessarily harmed.

From these studies, one may establish a rule-of-thumb: when downstream queues come within 200 to 250 ft of the upstream intersections, the upstream drivers will sense that congestion exists. This is quite apart from productivity considerations and even from delay as experienced in the downstream link. It is desirable to avoid this change in performance.

For completeness, it should be noted that the foregoing is based on the queue that is present at the initiation of upstream green. Attempts were made to correlate to average queue over the cycle. For vehicles stopped at initiation of green, it was an adequate substitute. Although instantaneous queue (that is, queue that is instantly formed) is preferred as a measure of effectiveness (MOE), average queue can be used as a proxy.

PLAN OF ATTACK

This section summarizes the over-all recommendations for treating the problem of congestion and saturation. There is a logical set of steps to take, which constitutes an over-all "plan of attack." Briefly, these steps are:

1. Address the root causes of congestion—first, foremost, and continually.

2. Update and, if necessary, improve the signalization.

3. Provide more space by use of turn bays and parking restrictions.

4. Consider both prohibitions and enforcement realistically—Is an effort futile? Will it only transfer the problem?

5. Take other available steps, such as right-turn-on-red, recognizing that the benefits will generally not be as significant as either signalization or more space.

6. Develop site-specific evaluations where there are conflicting goals, such as providing local parking vs. moving traffic, when the decision is ambiguous.

The framework for addressing a particular problem includes both preliminary identification of the cause of the problem, and categorization of the treatments available. Later sections in this manual provide details on the enumerated candidate solutions. It is vital, however, that the engineer use the later sections only after preparing, a preliminary opinion on the underlying cause, within the framework of the range of solutions available—using, in particular, Tables J-2, J-4, J-5, and J-6.

Address Root Causes

First, it must be recognized that the congestion/saturation problem should be attacked at the root causes; for example:

- The choice of alternate routes.
- The concentration of movement implied in some land use distribution.
- The use of on-street facilities for goods loading/unloading.
- The multiplicity of access/egress points, and of standing queues on the rights-of-way.
- The concentration of work trips in a short period.
- Unrestricted hours of goods activity.
- Numbers of vehicles, particularly low-occupancy private automobiles.
- Inefficient curb space management: parking, moving lanes, etc.
- Lack of grade separation in areas of extreme intensity.
- The improper design, location, and integration of access/egress points.

The engineer should continually educate other specialists and the public in this need. However, in the time frame of a local, site-specific problem that he must address, these guidelines must often suffice.

Improve Signalization

Discussions with traffic engineers, and surveys, have revealed that a systematic consideration of signalization for congestion and saturation is rarely done. Yet, it is difficult to overstate how often the basic problem is poor signalization. Once the signalization is improved through reasonably short cycle lengths, proper offsets (including queue clearance), and proper splits many problems disappear. Sometimes, of course, there is just too much traffic. At such times, equity offsets (setting designed for equitable treatment of competing flows in oversaturated conditions) to aid cross flows and different splits to manage the spread of congestion are appropriate—if other options cannot be called on. Study of representative traffic patterns lends strong credibility to the result that minimal response (i.e., preplanned) signal policies generally suffice.

Provide More Space

If a problem cannot be remedied by signalization, the next major set of actions is summarized in two words: more space. Left-turn bays and, where appropriate, right-turn bays can aid individual movements as well as remove impediments to the through flows.

Without question, additional lanes are a benefit. However, this tends to be an arterial-long solution. Two-way left-turn lanes offer special advantages, particularly along strip development sites.

One-way systems, arterials with unbalanced lanes, and reversible lanes offer advantages, but also represent either major implementation problems or site-specific treatments. One-way systems require studies quite beyond congestion, although that may be the prime motivator for such a study. Unbalanced lanes require certain volume patterns to be of use.

Consider Prohibition and Enforcement

Before instituting any prohibition and/or enforcement program, the traffic engineer must decide the following:

1. Can it be enforced strictly enough to realize most or all of the projected benefit? (Curb parking prohibition to provide a moving lane is an example.)
2. Will it simply transfer or even accentuate the over-all problem? (The impact on other streets from circulation of vehicles that would otherwise be double parking is a prime example.)

Only then can he consider that there is a potential benefit.

Some solutions that are available can have either a net benefit or a net disbenefit, depending on the site and the situation. Right-turn-on-red (RTOR) is such a case. If it allows vehicles to "escape" from a congested arterial, it is quite suitable. If, however, it allows vehicles to "steal" available space on such an arterial, it is inappropriate.

The question of such prohibitions as turning prohibitions arises; these can only be used if alternate routes exist.

Consider Solutions In Terms of Economics

Very often, application of these guidelines will clarify the issue and identify a solution. In some cases, the final decision will rest on conflicting desires that might be usefully viewed in economic terms. Is the removal of 5 parking spaces worth the delay savings to the traffic stream? Are off-street goods facilities justified economically? Are pedestrian phases justified in terms of total person-minutes saved? What is a proper allocation of curb space?

If necessary, the engineer can develop such an analysis for his individual case. More general treatment of these situations is recommended for future research. Some work on curb space management for goods facilities has already been done (*J-1*).

RANGE OF SOLUTIONS AVAILABLE

The fact that these guidelines are being consulted because of substantial traffic congestion virtually assures that one is dealing with signalized intersections. However, it does not follow that one has only signal remedies at hand. Indeed, the possible treatments may be broadly classified as signal (timing and coordination) and nonsignal.

Within the signal classification, there are two major subclassifications: minimal response (i.e., preplanned) and responsive signal control. It is not at all well established that highly responsive control is justified over preplanned signal control, particularly for the heavier flow range. This is an indication that is being reinforced by trends in major computer-based research projects. Not only is traffic flow rather regular from day-to-day, but the practical limit on detectorization (due to economics) does not allow sufficient information to be extracted so that a highly responsive plan can truly be implemented. Even if it were, it is not certain that it would be cost effective.

Within the nonsignal classification, there are also two major subclassifications: regulatory and operations. Regulatory consists of enforcement and of prohibitions. Operations, as used herein, consists of all other traffic measures.

Within this section, the following topics are addressed:

1. What improvement is to be sought?
2. What solutions are available?
3. What should be done initially?

Knowledge on these topics, combined with the "plan of attack" already presented, represent the essence of the recommended framework for attacking the problem of congestion and saturation.

Improvements Desired

Table J-2 lists a set of the most common improvements that a traffic engineer may wish to make, faced with a traffic congestion problem. These improvements are stated in the broad terms usually encountered as goals or objectives. A closer inspection of the items on this list, however, establishes that, depending on the traffic levels (relative to capacity), some of these are truly relevant, and others are merely consequences of the primary items. The list is not operative; that is, it does not say how to accomplish anything. Depending on the observed appearance of the problem, there is a logical predisposition to how much of the traffic stream can be helped. Of those that can be helped, there are two groups: the principals involved in the traffic situation and contributing to it, and those affected by the traffic situation caused by the principals.

Table J-3 summarizes the traffic stream flow components and their nature, in terms of an identified critical intersection. The engineer's task is twofold: reduce or eliminate the traffic congestion problem to the maximal extent possible; and, while doing this, provide as much benefit as possible to groups adversely affected, but without harming the over-all remedy.

As can be seen in Table J-3, the turning movements at the critical intersection itself can be classified in both

TABLE J-2

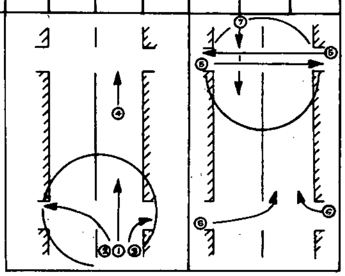
LIST OF IMPROVEMENTS DESIRED

- **Reduce Geographic Spread of Congestion**
- **Reduce Rate of Spread of Congestion**
- **Increase Throughput**
- **Reduce Delay**
- **Reduce Stops**
- **Improve Regularity of Service**

TABLE J-3

FLOW COMPONENTS AND THEIR CLASSIFICATION

	MAIN STREAM(S)				CROSSING STREAM(S)		
	①	②	③	④	⑤	⑥	⑦
PRIMARY NATURE OF FLOW	THRU	LEFT	RIGHT	DOWN STREAM	CROSSING	TURN-INS	OPPOSING DIRECTION
CONTRIBUTING TO THE CRITICAL SITUATION	*	*	*	–	–	*	–
IMPACTED BY THE CRITICAL SITUATION	–	*	*	*	*	–	*



NOTE: CRITICAL INTERSECTION CIRCLED FOR EMPHASIS

categories. The treatment of these movements depends on the particular problem manifested and on the relative magnitude of the movement in contributing to the criticality. To illustrate, a right-turn bay may provide some increase in throughput capability (thus providing some alleviation of the criticality), but its principal impact may be significant savings in delay that would otherwise be encountered by right turners. The bay may be justified strictly on the basis of aiding this one flow component.

It was noted earlier that some of the items in Table J-2 are actually secondary to other items in that table, for given flow levels. The primary objectives that should be sought by the engineer are summarized in Table J-4. These are dependent on the traffic level; first and foremost, the engineer must realize that, at the more extreme flow levels, it is essential to "avoid spillback." This is the explicit objective. The mathematical niceties of minimum stops and/or minimum delay collapse in the face of intersections blocked by vehicles.

Some comments on Table J-4 are in order. First, the primary objective to which the engineer should address himself does change depending on the flow level. Secondly, the major index of performance, or measure of effectiveness (MOE), also changes. However, both sets are well correlated to queue extent measures, so that queue and/or occupancy patterns, particularly during red and at the onset of green, are good indicators across the entire range of conditions.

It is appropriate to also call attention to the last item in Table J-2—"improve the regularity of service." During simple congestion, the variance as well as the mean of the delay per cycle increases as the demand approaches capacity. (Many readers are probably acquainted with the exploding nature of the average queue as the demand rate approaches the service rate. The same sort of queuing analysis also reveals that the variance—the spread of values—also explodes as the demand rate approaches the service rate.) Thus, the individual driver will be exposed to greater variability in his individual experience from day-to-day. Improvements that minimize the mean delay will also enhance the regularity of the delay suffered.

Likewise, on the higher flow levels, the actual blockage of intersections occurs probabilistically. If the potential for blockage is increased (by filling the downstream link), a variable phenomenon is introduced as to whether there will actually be blockage at a fixed time after the potential occurs. Avoiding the potential thus contributes to enhancing the regularity that the driver perceives.

There are times when a basic solution has been achieved and some possible benefits can be realized by additional improvements addressed to specific subgroups. Right-turn bays are such a case because they frequently have little impact on such a measure as average delay of all vehicles, but they have truly substantial benefits for a smaller segment of the traffic stream—the right turners.

TABLE J-4

OBJECTIVES TO BE SOUGHT DEPEND ON TRAFFIC CONDITION

TRAFFIC CONDITION	Congestion →	Saturation/Oversaturation
OBJECTIVE TO BE DESIRED	{Minimize Delay / Minimize Stops} →	{Avoid Spillback / Provide Equitable Service}
IMPACT OF USING OR TRYING TO ACHIEVE ABOVE OBJECTIVE (i. e., RESULTS)	• Intersection-specific congestion, not area-wide • Throughput Adequate • Service Regularity Enhanced • Enhancements to Special Groups (such as right turners) Possible	• Reduce Spatial Extent and Rate of Spread of Congestion • Potential for occurrence of area-wide Congestion reduced, and service regularity thus enhanced • Throughput not Impeded • Special Treatment of User Groups Possible
MAJOR INDEX OF PERFORMANCE	Delay, Stops	Queue Extent and Occupancy

Solutions Available and What Should Be Done

Table J-5 summarizes the range of solutions available, identified by the classification set forth at the beginning of this section.

It must immediately be stated that there is no simple statement of an ordered list of recommended solutions, in decreasing order of preference. There are, however, indications of how much of one solution must be done to have an equivalent impact of so much of another solution. These considerations are incorporated in the detailed coverage given the solutions in later sections of this manual. The engineer must use this knowledge in conjunction with local conditions and practices to reach a decision. There are also indications of how best to use two solutions in conjunction with each other. A simple set of initial steps that can be followed as an elimination checklist is given in Table J-6.

Before continuing, a word of caution is necessary. The engineer must reach a preliminary judgment of the underlying cause of the problem; at the same time, he must assure himself that the solution is not trivial. Extensive queues may lead one into the depths of these guidelines too quickly: such problems can arise because of poor offsets, outdated splits, and excessive cycle lengths. As a first step, therefore, the engineer should prepare a preliminary opinion on the underlying cause. Given a preliminary opinion, there are a number of possible solutions that one is either tempted or inclined to consider. The remainder of these guidelines is addressed to the candidate solutions, the considerations involved, and the relative merits.

MINIMAL RESPONSE SIGNAL REMEDIES—INTERSECTION

This section addresses only remedies that may be effected by changes in signal timing at the intersection level. The specific role and effect of offsets are addressed under the heading "Minimal Response Signal Remedies—System" in at later section.

Cycle Length

One of the most prevalent misconceptions in the traffic engineering community is that the capacity of an intersection increases substantially as the cycle length is increased. This belief has been a debated question in the past (*J-2*), and studies (*J-3, J-4*) have provided data to support such questioning.

The lack of substantial capacity increases with increasing cycle length is rooted in at least three factors: (1) the loss time per cycle is not that severe because of both use of the amber and start-up delays lower than often assumed; (2) the inefficient use of longer greens because of increasing headways; and (3) the inability to provide a demand to fill long green times. The last item is just another manifestation of the storage problem, addressed later.

Figure J-1 illustrates cycle length as a function of critical lane demand (total) for a range of values of loss time per cycle, LPC (loss time per cycle being defined as sum of all amber not used as green and all start-up delays). For a two-phase operation, LPC $\simeq$ 6 sec is reasonable. The horizontal axes show the sum of green required for all critical lanes, and the corresponding sum

TABLE J-5

RANGE OF SOLUTIONS AVAILABLE

CONTROL MEASURE	TYPE	WHERE APPLIED	CHANGE POSSIBLE
SIGNAL	MINIMAL RESPONSE	INTERSECTION	CYCLE LENGTH
			BLOCK LENGTH
			SPLIT
			EXTRA PHASES
		SYSTEM	PREPROCESSING TRAFFIC
			PHASE ARRANGEMENT TO AID PROGRESSION
			EQUITY OFFSETS
			SPLITS TO APPORTION STORAGE AVAILABLE
			DOWNSTREAM EFFECTS
	HIGHLY RESPONSIVE	INTERSECTION	EXISTING POLICIES
			A MAXIMUM-QUEUE POLICY
			EXISTING HARDWARE
		SYSTEM	SPREAD FROM INTERSECTION
NON-SIGNAL	REGULATORY	ENFORCEMENT	AREA-WIDE
			DOUBLE-PARKING
			OTHER
		PROHIBITIONS	PARKING
			TURNING
			OTHER
	OPERATIONS	TURNING	LEFT BAYS
			RIGHT BAYS
			DUAL TURN LANES
			TWO-WAY LEFT TURN LANES
			RTOR
		LANE ARRANGEMENT (MAJOR)	ONE-WAY STREETS
			UNBALANCED
			REVERSIBLE
		DISRUPTION	PEDESTRIAM IMPACT
			BUS STOPS
			PARKING/UNPARKING
			DOUBLE-PARKING
			MID-BLOCK

of equivalent critical lane volumes for the average headways. The headway values were obtained from the literature (*J-4, J-5*), and are consistent with values found in this study.

It should be noted that Figure J-1 has two messages: (1) relatively short cycle lengths suffice for rather high volumes (under 60 sec for up to 1400 passenger cars per hour, pcph, for the sum of the critical lanes, with 2.29-sec headways assumed); and (2) substantial increases in cycle length provide little corresponding increase in capacity (at C = 80 sec, 1450 pcph; at C = 120 sec, 1500 pcph). In the illustration, a 50 percent increase in cycle length above 80 sec provides only a 3.4 percent increase in capacity. One must ask if there is not a better way to gain more increase.

By viewing Figure J-1 so that cycle length is the (horizontal) axis and flow as the vertical (see insert in Fig. J-1), one can emphasize the slow growth of intersection capacity with cycle length. It is important to note that any cycle length implemented must allow for adequate pedestrian crossing times.

Block Length and Storage

Some Basics and Upstream Block Length

It has been illustrated that cycle length may not be as powerful a capacity-improver as one might think. Nor has there been any evidence that there is a positive good to short cycle lengths in some cases.

If an approach has a flow of f_1 pcph per lane (pcphpl) and a cycle length C, note that there are f_1 $(C/3600)$ vehicles per lane per cycle. A minimum condition is therefore that this platoon does not exceed the available link storage, because, if it did, the potential for spillback

TABLE J-6

AN INITIAL CLASSIFICATION AND ELIMINATION CHECKLIST

Apparent Problem	Initial Steps
Area-Wide Congestion	• Identify the Critical Intersection (CI). • It is unlikely that there are two CI's. If there are, it is likely that each can be considered with its own area of influence. • Do not erroneously identify the intersection downstream of the CI as the CI. • Classify the type of oversaturation (Types I through V, as defined)
Spillback in a Link	• Determine whether a simple split adjustment is sufficient. • Determine whether the cycle length is too long for the block length and flow. • Determine whether the offset is poor for the primary and secondary flow mix. • Identify any special blockages in the link (double parking, queues for garages, car washes, etc.).
Single Intersection	• Isolate primary symptom.
→ Single Approach	• Check split and offset as above. • Check burden caused by turns.
→ Two Approaches, same Right-of-Way	• Check same as one approach, but with emphasis on interference with each other.
→ Two or More Approaches More than One Right-of-Way	• Consider methods for increasing net capacity. • Consider methods for minimizing spatial extent of possible oversaturation and area-wide congestion.

would exist. (It is assumed that the entire platoon is stopped, and saturation is approached.) Thus,

$$f_1\left(\frac{C}{3600}\right)\zeta \leq \mathcal{L} \qquad \text{(J-1)}$$

where:

ζ = vehicle storage length, 20 ft is used in some examples herein; to use other distances (23 ft is often used), simply substitute the desired value for ζ; in preestablished figures herein (e.g. Figs. J-3 and J-4), simply use $\mathcal{L}$ reduced by the factor $(20/\zeta)$ when $\zeta \neq 20$ is desired);

$\mathcal{L}$ = link storage, ft;

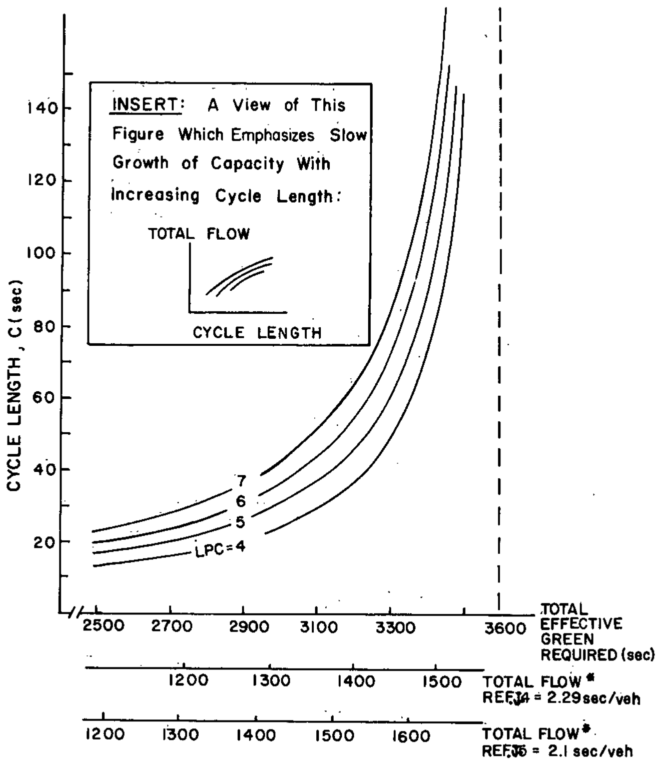
C = cycle length, sec; and

f_1 = flow, pcphpl.

It should be observed that $\mathcal{L}$ need not be the physical length of the link; if a policy decision is made that the stored vehicles should come no closer than within 200 ft of the upstream intersection, $\mathcal{L}$ is 200 ft less than the physical length. Such a policy decision is in accord with the avoidance of the perception of congestion, and not just the avoidance of spillback.

Figure J-2 depicts the block length/cycle length congestion situation. Note that, if it could be assumed that the platoon could be moved through the intersection without stopping, the congestion situation would not, by definition, exist.

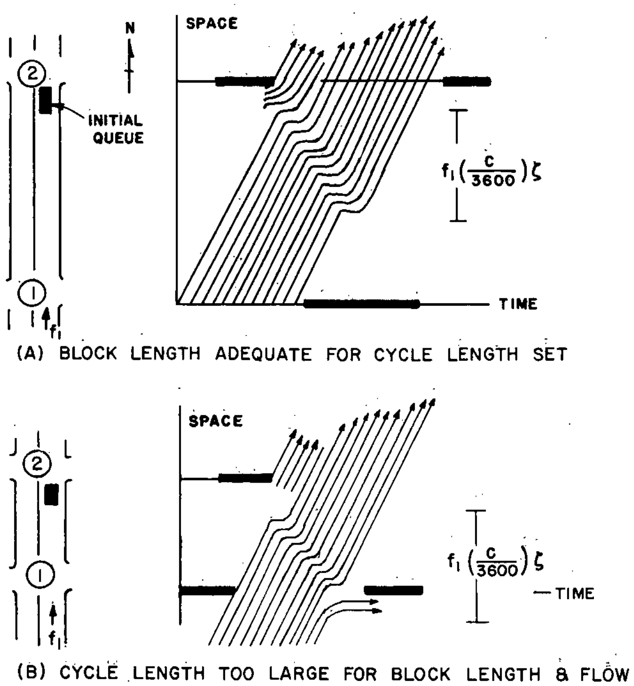
In Figure J-2 (A) a normal delay to an input platoon is depicted. The vehicles stored at the downstream intersection (the CI) may have been turn-ins or may have been internally generated. However, if the upstream in-



* SUM OF CRITICAL LANE VOLUMES IN PASS. CAR EQUIVALENTS.

NOTE: $LPC_{REF.J5} \simeq LPC_{REF.J4} + N\,(3.7 - 0.75)$ FOR N PHASES

Figure J-1. Cycle length required for various flows.



NOTE: In Part B, the stoppage presents some vehicles from discharging Intersection 1. It also causes conditions that could lead to a blockage of Intersection 1 (e.g., If there were any delays in moving through Link 1-2, the cross stream traffic at Introduction 1 would be quickly affected).

Figure J-2. Block length problem.

tersection were located as in Figure J-2(B), the platoon would have been stopped as it tried to cross the intersection. The signal would have turned red before all $f_1(C/3600)$ vehicles had a chance to enter the link. The input flow would be decreased, the output would be decreased, and the apparent congestion would show up on the next upstream link. Moreover, there is an excellent chance that some drivers will be physically stopped in the intersection proper when the signal changes. This spillback will block the east-west flow, causing delay to it.

In order to avoid this situation, it is necessary to assure that excessively long platoons are avoided. Eq. J-1 may be rewritten as a constraint on cycle length:

$$C \le \left(\frac{3600}{f_1}\right)\frac{\mathcal{L}}{\zeta} \qquad \text{(J-2)}$$

Figure J-3 combines the results of Eq. J-2 and Figure J-1. It is apparent that there are two contrary forces at work: as the total critical lane flow (all approaches) increases, cycle-length increase brings some benefit; at the same time, flow increases on any one approach decrease the maximum cycle length permitted.

The example given in Figure J-3 indicates that, for the conditions shown to be existing in the illustration, the only practical standard cycle length is 80 sec. If it were boosted to 120 sec, to increase intersection capacity, it would cause upstream congestion and spillback.

As noted, the distance $\mathcal{L}$ may be less than the physical length. It should be sufficiently less so as to allow a "cushion" (200 ft as noted, but at least 100 ft). Should it also allow for queues that will form during the (upstream) red—such queues as may be caused by turn-ins and by internal generation? Proper definition of the flows avoids the need for this special consideration.

By defining the flows as the net critical lane demand on the CI, the definition of the flows would include any such vehicle generation. In the example of Figure J-3, the flows shown are the net demand for throughput at the CI, and the definition is as recommended; the storage computation is thus simplified and is conservative, although not extremely so.

Figure J-3 may also be used as shown in Figure J-4. Suppose you wish to know what the maximum critical lane flows are for a given minimum cycle length, C, and set of $\mathcal{L}$ values. For the example given in Figure J-3, the $C = 77$ sec would seem to indicate that the sum of critical lane volumes cannot exceed 1410 pcph, the westbound critical lane can be no more than 700 pcph, and the northbound critical lane can be no more than 940 pcph.

Referring to Figure J-4, for another expression of the storage problem, assume it is known that $f_1 = 825$ pcph (critical lane), that $f_2/f_1 = 0.7$, and that $\mathcal{L}_1 = 300$ ft. Then the cycle length must be between 55 and 65 sec to avoid storage problems on approach 1 and to meet the capacity required. At the same time, note that $f_2 = 0.7$ (825) = 577 pcph and $f_1/f_2 = 1.43$. Inspection of Figure J-4 indicates that there is no storage problem for approach 2 with $55 \le C \le 65$ unless the $\mathcal{L}_2$ is exceptionally short.

When the split is fixed, with the relative flows as indicated, f_B (the competing flow) need not actually be 577

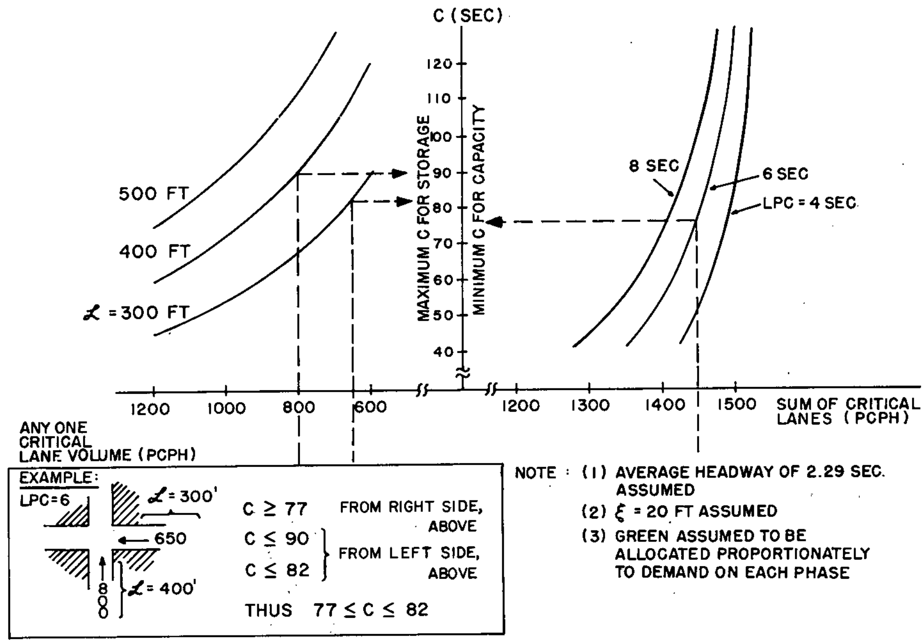


Figure J-3. Storage-induced bounds on cycle length.

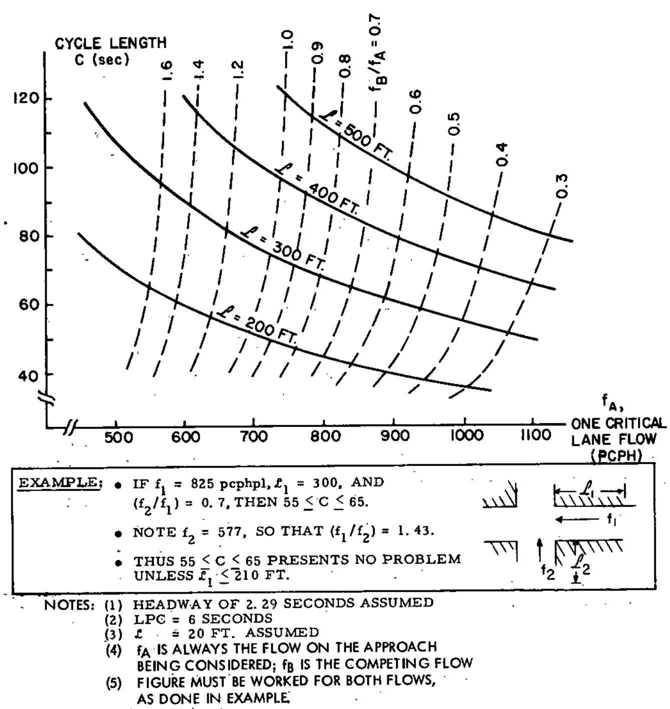


Figure J-4. Another expression of storage problem.

pcph (critical lane) in the foregoing example. If a decision is made that the split of effective green is in the ratio of 0.7 to 1.0 (Side Street B to Main Street A), the same computation would hold: if $f_A = 825$ pcph, $55 \leq C \leq 65$. Moreover, if f_A jumps to 875 pcphpl with $\mathcal{L}_A = 200$ and $f_B/f_A = 0.7$, there is no adequate cycle length. There will be congestion on Main Street A, and possibly spillback (if $\mathcal{L}$ is the physical length). Only if the split is changed to a ratio of 0.6 to 1.0 can the problem be avoided. Increasing the cycle length in the face of such congestion is the wrong solution.

A final note on the link(s) upstream of the CI—Figures J-3 and J-4 were prepared with LPC = 6 and headway = 2.29 sec assumed. These values can be changed not only by statistical fluctuations from cycle-to-cycle but also by site-specific conditions leading to persistent inefficiencies (e.g., grades on the approach). An increase of LPC from 6 to 8 sec, for instance, can cause an upward shift of all the dashed curves in Figure J-4 by 33 percent. A change in the average headway could cause a lateral shift of the same curves. Clearly, the figures are good for ascertaining probable effects and determining relative value. Because of possible site-specific variations in LPC, headway, or average vehicle length, they will not precisely predict results to narrow resolution.

Downstream Block Length

The links downstream of a CI can be made to become congested and can spill back into the CI if one acts on the assumption that the problem dissipates after the CI.

In general, the downstream intersection can operate at a greater green/cycle ratio than the CI itself. Still, the storage/spillback problem exists when there are a sufficient number of turn-ins from the CI into the downstream link, so that the main through platoon must stop, even for a short time.

It is therefore necessary to apply the cycle-length determination previously discussed to the downstream links. If the downstream intersections are to have the same cycle length as the CI or one-half the CI (the first is to be preferred), these downstream links can control the cycle length at the CI.

If a decision is made that the cycle length at the CI is to be larger than the downstream cycle lengths (perhaps because the CI is to have multiple phases), it may be shown that the queue extent to be stored can reach a maximum of over twice the single-cycle discharge of the CI into the link.

Note from Figure J-5 that the queue and occupancy fluctuate in a complex way over a "common cycle" or "least common denominator" when the two cycle lengths are not equal. In the field, this may appear irrational and/or may be obscured by fluctuations in demand.

To check on whether a storage problem exists in the downstream link, the following steps are recommended:

1. Use Eq. J-2 (or the left side of Figure J-3) with the downstream f and $\mathcal{L}$. The cycle length thus determined is a constraint on the CI.

2. For some selected C_D (cycle lengths at the downstream), it is possible that the combination of C_{CI} and C_D will give rise to queue extents even larger than one full CI discharge. This can be avoided by assuring that

$$C_{CI} \geq 2 \frac{(1 - g_D)}{(1 - g_{CI})} C_D \qquad \text{(J-3)}$$

where the subscripts D and CI refer to downstream and critical intersection, respectively; and g is the green/cycle ratio for the phases of interest. When there are heavy turn-ins from the CI into the downstream link, g_{CI} can be increased to reflect their contribution.

It should be noted that Eq. J-3 is only practical when it is originally intended to have C_{CI} much higher than the original system (of which C_D is a part). If this is not the intention, one must review the reasons why C_{CI} and C_D are to be different in the first place. Some engineers believe that the fluctuations caused by such differences, as exhibited in Figure J-5, are undesirable. In any case, a "worst case" can be considered by redoing Step 1 with one-half of the $\mathcal{L}$ originally intended, thus allowing for double the single-cycle CI discharge.

Splits

For congested flow, the standard rule of proportioning available effective green in the ratio of the critical lane flows will not suffice. It should be appreciated that as the demand approaches capacity, greater queues will be experienced, as will greater delays and greater fluctuations (variance) delay per cycle.

For unstable saturation and for oversaturation, a different concept should prevail. There are situations in which the CI simply cannot handle the total demand put upon it, and the question arises as to whether the same sense of equitable treatment must still hold?

It is recommended that the split be apportioned so that the rate of growth of congestion in both (or all) directions be equalized, with both directions exceeding their respective links or defined system boundaries at the same time. One may show that the split should be

$$g_A = \frac{(f_A/s) + (Q_A/Q_B)\ (K - f_B/s)}{(1 + Q_A/Q_B)} \qquad \text{(J-4)}$$

where:

$Q_A = (S_A - f_A\ C)$, and likewise for Q_B;
$S_A =$ available storage per lane, vehicles;
$f_A =$ per lane demand on phase A, pcph;
$s =$ per lane saturation discharge, pcph;
$K = g_A + g_B = 1 - (\text{LPC}/C)$; and
LPC = loss time per cycle.

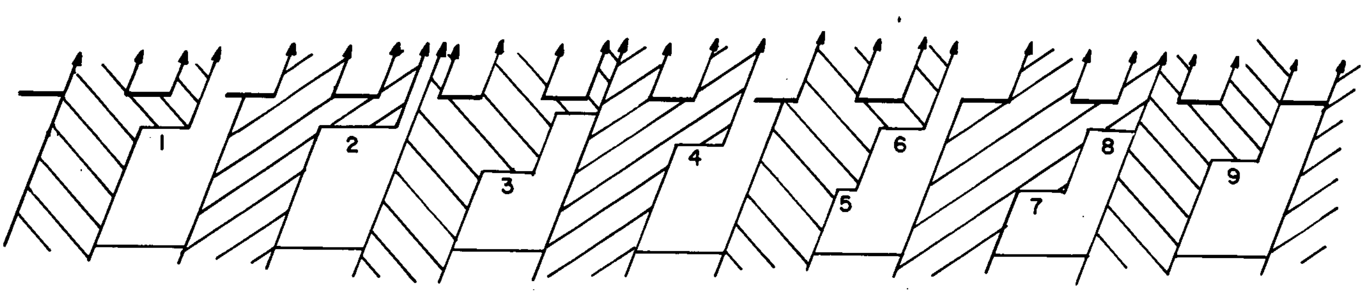
Eq. J-4 is formulated for equalizing the growth in the immediate upstream link (i.e., the time at which the next upstream intersections are first exposed to great spillback potential). If the demand will persist sufficiently long, much greater spread is inevitable.

Eq. J-4 may also be used for a system equalization. Suppose the engineer judges that, if congestion must spread, he wishes it to spread no further than a total distance of L_A on approach A and a total distance of L_B on approach B. The version of Eq. J-4 that results for system spread is:

$$g_A = \frac{(f_A/s) + (S_A/S_B)\ (K - f_B/s)}{1 + S_A/S_B} \qquad \text{(J-5)}$$

Under this condition, the direction of spread will be controlled, but oversaturation will impact (possibly) several intersections. Offset policies, addressed later, may alleviate the impact, assuming that the spread is inevitable.

Figure J-6 illustrates the foregoing concepts of split determination. The conventional wisdom dictates that



NOTE: CI HAS 120 SECOND CYCLE, 60 SECOND GREEN : (G/C) = 0.50
DOWNSTREAM HAS 70 SECOND CYCLE, 40 SECOND GREEN: (G/C) = 0.57

OBSERVE: QUEUE EXTENT FLUCTUATES OVER "COMMON CYCLE" OF 7 (120) = 840 SECONDS.

NUMBERS: LOCATIONS OF LAST-STOPPED VEHICLE ARE NUMBERED.

Figure J-5. Queue extent downstream of CI.

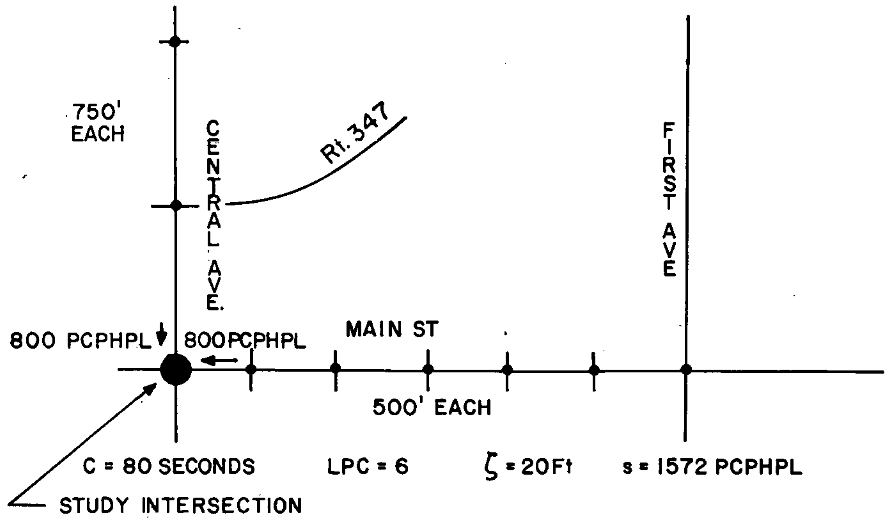


Figure J-6. Illustration of split determination.

$g_A = g_B$ because the competing flows are equal. However, use of Eq. J-4 provides a different solution. First, some preliminaries; denote Main Street as "A" and Central Avenue as "B." Also, $K = 1 - 6/80 = 0.925$; $Q_A = (500/20) - 800(80/3600) = 7.22$; and $Q_B = (750/20) - 800(80/3600) = 19.72$. Thus,

$$g_A = \frac{\left((800/1572) + (7.22/19.72)\left(0.925 - \frac{800}{1572}\right)\right)}{(1 + 7.22/19.72)} = 0.485$$

$$g_B = 0.925 - 0.485 = 0.440$$

rather than the $g_A = g_B = 0.462$, which would result from $g_A = g_B$. The "leftovers" per cycle in each lane are as follows:

$$\{\text{ARRIVALS/LANE}\} - \{\text{DEPARTURES/LANE}\}$$
$$= 800\left(\frac{80}{3600}\right) - 1572\ (0.485)\ \frac{80}{3600} = 0.84$$

and the excess space available to be used up, in terms of vehicles, is

$$\left\{\begin{matrix}\text{TOTAL} \\ \text{SPACE} \\ \text{PER LANE}\end{matrix}\right\} - \{\text{ARRIVALS/LANE}\}$$
$$= \frac{500}{20} - 800\left(\frac{80}{3600}\right) = 7.2$$

because the space is needed each cycle for the plateau that must inevitably stop. Thus, it will be $N \simeq 7.2/0.84 = 8.5$ cycles before the link flows over. This is $9(80) = 720$ sec $= 12$ min.

If an extended peak is anticipated, and one wishes to avoid affecting the major arterials Route 347 and First Avenue, it would be appropriate to use Eq. J-5 with $L_A = 6(500) = 3000$ and $L_B = 750$. Then, with an equivalent vehicle spacing of 20 ft assumed,

$$S_A = 150$$
$$S_B = 37.5$$

and

$$g_A = \frac{(800/1572) + \left(\frac{150}{37.5}\right)\left(0.925 - \frac{800}{1572}\right)}{\left(1 + \frac{150}{37.5}\right)} = 0.435$$

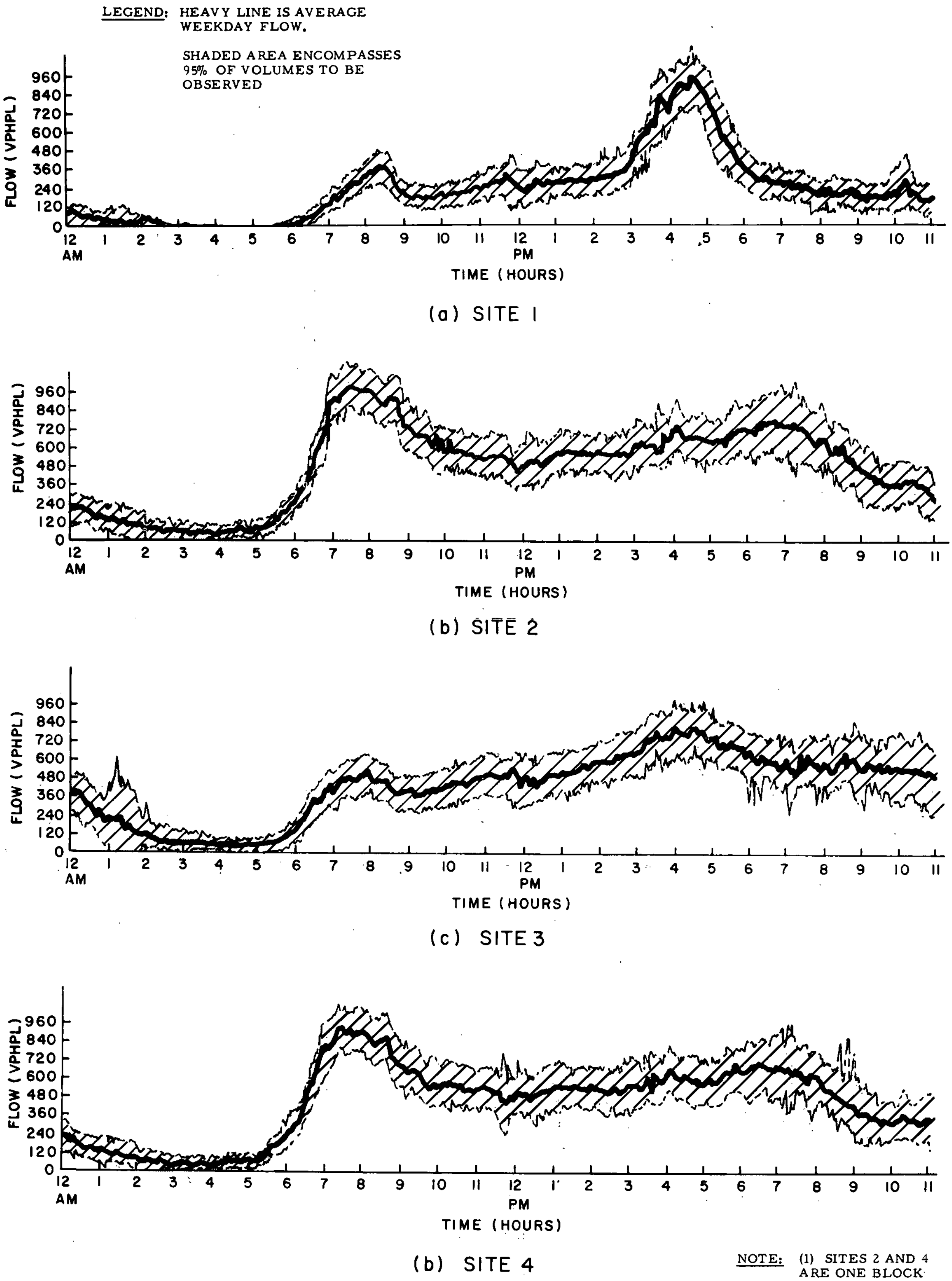
$$g_B = 0.925 - 0.435 = 0.490.$$

Note that the split is virtually the reverse of the single-block consideration. Greater flow imbalances and more extreme relative lengths on the two competing directions will cause the resulting split to deviate more sharply from the simple proportion $g_A/g_B = f_A/f_B$.

If an intersection has a pronounced pattern of major shifts in demand amongst its various approaches, the engineer may wish to frequently change the split so as to match the changing demand. This can be done with actuated equipment, but, with limited detector set-backs (due to urban signal spacing) and heavy demand on all phases, the engineer may wish to preplan the split changes. These split changes can be pretimed by time of day, and last as little as 10 to 15 min.

As part of the study resulting in these guidelines, data were acquired on daily traffic fluctuations through the courtesy of the Metropolitan Toronto Department of Roads and Traffic. The average weekday pattern at four sites, two of which (sites 2 and 4) were on the same street, in the same direction, is depicted in Figure J-7. Also shown is the band within which values are expected to fall 95 percent of the time. Another view of this variation for one of the sites—the distribution of times at which certain volume levels are first reached—is shown in Figure J-8.

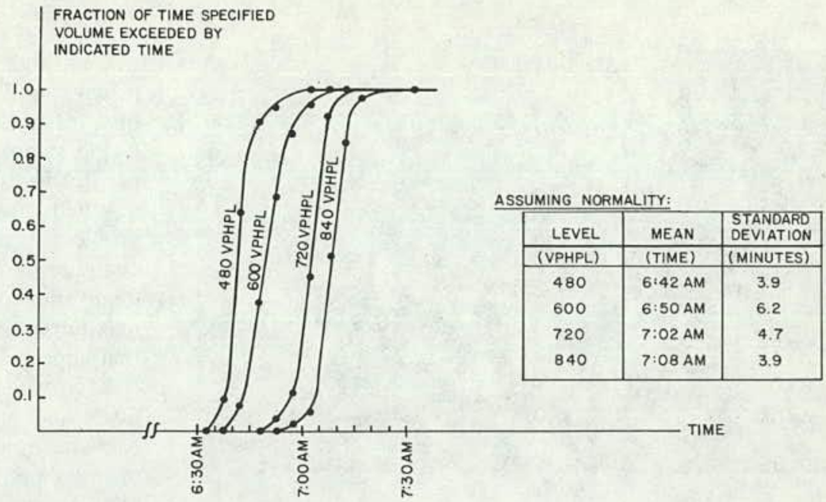
It is clear that, if the engineer wishes to implement a particular split assignment for a period of n min starting at a particular time $t - t_0$, it is only sensible that the split match the demand as planned. If the demand can start up to say 5 min earlier than $t = t_0$ or up to (say) 5 min later than $t = t_0$, it makes little sense to preplan a split in the order of $t = 10$ min: it will often be mismatched to the traffic demand. If, however, the same split is suitable for



NOTE: (1) SITES 2 AND 4 ARE ONE BLOCK APART ON SAME STREET, IN SAME DIRECTION

(2) ALL SITES ARE TWO MOVING LANES.

Figure J-7. Average weekday pattern and variation of individual flows.



LEVEL (VPHPL)	MEAN (TIME)	STANDARD DEVIATION (MINUTES)
480	6:42 AM	3.9
600	6:50 AM	6.2
720	7:02 AM	4.7
840	7:08 AM	3.9

Figure J-8. Times at which certain volume levels reached (Site 2).

the traffic over $n = 30$ min, even if the first part of the demand is improperly matched, the split is suitable for much of the traffic.

When the engineer does find it suitable to implement a number of different splits, it is important to remember that the pattern may shift seasonably or in other systematic ways. The engineer must realistically assess whether he can actually monitor and update his very detailed signalization periodically.

Extra Phases

As a rule, multiple phases should be avoided, particularly because they generally require an increase in the over-all cycle length. Other options should be considered, such as: turn bays, to remove those awaiting gaps from the through traffic stream or in the pedestrian flow—but without a distinct (protected) phase; shorter cycle lengths, to provide more end-of-cycle turning opportunities per hour; parking restrictions near the intersection, particularly on single-lane approaches, to provide "by-pass" opportunities for through traffic that would otherwise be impeded by turners; leading and/or lagging greens, to aid turning vehicles; and turn prohibitions at specific intersections, if acceptable alternate routes to desired destinations exist. Figures J-9 and J-10 illustrate some of the problems encountered that would possibly be solved by applying these options. For example, if one uses a rule of two end-of-cycle left turners per cycle (*J-6*), a 60-sec cycle provides for 120 left turners per hour per approach, whereas a 120-sec cycle provides for only 60 left turners per hour.

Still, there are cases when multiple phasing is necessary. It may be a local policy to provide protected left turns, as is the case in Dallas, Texas. It may be that the individual left turn volumes exceed 120 vph, and turn prohibitions are not practical.

Even when the left-turn volumes are less than 120 vph, it may be shown that there are conditions under which a left-turn phase can be added without increasing the cycle length required. (The equation on which Fig. J-1 is based can be manipulated with the left turns first expanded by the addition of a left-turn factor and then unexpanded to demonstrate this; in the latter case, the LPC is increased accordingly.) In general, if the total of the critical lane volumes, including the unexpanded left turners (i.e., without any left-turn factor added), exceeds 1400 pcph, an additional phase can be added when individual left-turn volumes exceed 60 pcph. When the per lane opposing flow exceeds 400 pcphpl, this limit drops to 1200 pcph total.

The foregoing rule-of-thumb is a statement that (for instance) a 2-phase timing without left-turn bays can sometimes be replaced by a 3-phase timing with left-turn bays, without necessarily increasing the traffic-based cycle length (Eq. J-1). A solution that would provide the left-turn bays for "out of the way" storage of the turners without a protected movement (i.e., the 3rd phase) would allow still lower cycle lengths.

MINIMAL RESPONSE SIGNAL REMEDIES—SYSTEM

This section addresses remedies that can be effected by means of signal coordination. It is assumed that the

Figure J-9. Turn bays and shorter cycles can sometimes avoid multiple phases.

(A) TURNER IMPEDING THE THRU FLOW

(B) THRU VEHICLES BY-PASSING TURNER ON SINGLE-LANE APPROACH

Figure J-10. Effect of left turners on single-lane approach.

engineer has considered those signal timing issues that relate primarily to the intersection—cycle length and split—and that he has also considered any possible impact of block length when computing cycle lengths. To solve a specific problem, he should always consider signal remedies relating to the system, as well as nonsignal options. He should never progress from front to back of these guidelines, stopping with the "first solution that looks as if it will work."

In this section, three consecutive fall-back positions will be identified for signal coordination: (1) effective progression of traffic while the opportunity exists, (2) progression to avoid spillback, and (3) signal offsets, such as to enhance equity, given that spillback is inevitable and/or historically occurs.

Progression of Traffic Using Offsets

With light flows on a one-way street, a satisfactory progression is easy to attain; the offsets are set so that a forward progression exists, with the signals turning green just as a moving platoon reaches them. In other words, the progression speed is equal to the vehicle speed.

When a two-way street or a network is considered, the problem is more complex. A maximum bandwidth or a computer routine for minimum delay (and/or stops) might be employed. (A maximum bandwidth solution maximizes some combination of the "windows of green" that appear on the two directions of an arterial.)

Even when the street of interest is only one-way, heavy flows and/or turn-ins cause queues that make a simple forward progression unsuitable. During congested flow, the offsets should be adjusted to allow observed or historically known queues to be discharged before additional vehicles arrive. This not only decreases the number of stops, it also shortens the spatial extent of the stopped vehicles. Figure J-11 illustrates the foregoing situation.

Because of queues that exist in individual links, the offset should be

$$t_{off} = \left\{ \frac{L}{V} - \frac{Q}{R} \right\} \tag{J-6}$$

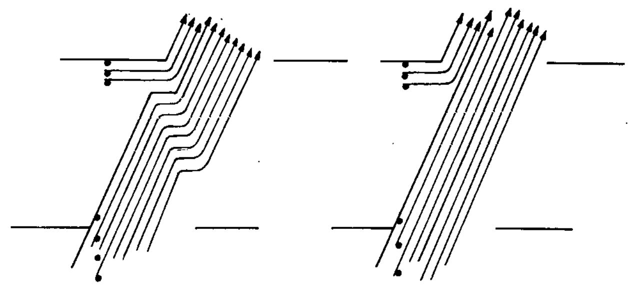
where:

t_{off} = desired offset, sec;
L = signal spacing, ft;
V = design speed, fps;
Q = estimated or known queue, veh/lane; and
R = discharge rate, veh/sec/lane, and is typically in the range of 1 veh/2.0 sec to 1 veh/2.3 sec/lane.

Note that as the anticipated queue grows (i.e., as congestion becomes more severe), the offset is initiated earlier and earlier. The progressions that result conform to the standard recognized forms: simultaneous progression and backward or reverse progression. However, the engineer should not think of these as three uniquely different types of progression: they each result naturally from a common equation (Eq. J-6), as congestion becomes more severe.

To indicate the common root of the three progression types, consider the common situation when $V = 45$ fps (i.e., 30 mph) and $R = 1$ veh/2.3 sec/lane. Let t_0 be the zero-queue forward progression. When the anticipated queue at green initiation occupies 20 percent of the block length, a simultaneous progression results. (The 20 percent figure is a useful rule-of-thumb which results from solving Eq. J-6 for Q, with $t_{off} = 0$ and typical values, including average vehicle length.) When the queue occupies 40 percent of the block length, a reverse progression, which is a "flip" of the original progression, results and the offset is $-t_0$.

Figure J-12 shows the results of switching from a standard forward progression to a simultaneous progression on a representative one-way arterial with saturation flow. In line with the previous reasoning, the simultaneous pattern was more appropriate than the simple forward progression. Indeed, this accounts for the empirical observations that simultaneous progressions alleviate congestion.



(A) SIMPLE PROGRESSION

(B) OFFSET CHANGED IN ORDER TO DECREASE STOPS AND SPATIAL EXTENT

Figure J-11. Offset changed because of standing queue.

Changing the arterial offsets also changes the relation of these signals to other signals (the cross-street traffic can be affected). If the cross-street offsets cannot be adjusted, perhaps because too many other arterials would be impacted in the engineer's opinion, there can be substantial adverse impact. However, if the main arterial has sufficient congestion so that spillback occurs, the cross-street traffic will be adversely impacted without the switch to simultaneous.

The network of Figure J-13 may also be considered. Two cases will be illustrated: small turning movements and major turning movements. These correspond to the defined saturation Types I and III, respectively.

In the first case (minor turns), the simultaneous offset plan produced less average delay per vehicle than did the original simple forward progression: 92.3 sec/veh vs. 117.4 sec/veh for the system, and 59.1 vs. 72.6 for link 1. Moreover, the extent of oversaturation is reduced; with the original plan, link 1 became oversaturated, and then links 2, 3, and 4 successively became oversaturated. With the simultaneous plan, the oversaturation does not spread and cause a system lock-up, as could happen. Figure J-14 illustrates the queue build-up in the network.

Note that proper offset is the companion of short cycle length in avoiding spillback-inducing queues. This was shown in Figure J-11 and in the simulation results of Figure J-12. More is to be said later on the avoidance of cross-direction spread of congestion and oversaturation, such as exhibited in Figure J-14.

The second case (major turns) is illustrated in Figure J-15. The problem is not so simple as the cross traffic being assured the opportunity to cross; a large number of the vehicles wish to enter the heavily used link (link 1). If they cannot, they will wait in their own link (link 2), causing stored vehicles to grow therein. This will happen to all links in succession, for heavy turning exists on all links in this illustration. Note, however, that the simultaneous offset did reduce queue extents and forestall the onset of oversaturation.

The engineer should not be misled by the fact that reference is made only to "simultaneous" settings in these examples. Rather, he should apply Eq. J-6 whenever he recognizes a queue-extent congestion problem (saturation Type IV). As a practical reality, however, queues occupying about 20 percent of the block length need simultaneous offsets. Moreover, the engineer may be pressed for time, with other concurrent duties, and may wish to "try out" a solution. Simultaneous settings are a good starting point, when saturation Type IV is suspected.

One must also recognize that an adjustment of the offsets causes a change in initial queue on which the adjustment was based. This process could be iterative, particularly when heavy flows, internal generation, and/or major turnins exist. However, approximations frequently suffice. Figure J-16 depicts the improvement in one representative link attained with queue averaging. The average queue over all time (entire cycle lengths) was observed under the existing setting, and the offset was changed to clear this average queue (Eq. J-6). Such adjustment of offset via queue averaging provides for reduced queue extent, with offsets in each link tailored to the queue development therein.

Phase Arrangement to Aid Progression

The previous treatment of offset addresses the problems of reducing congestion and of avoiding spillback, the latter in concert with reduced cycle length. It is implicitly assumed that one identifiable critical direction exists; congestion leading to oversaturation on both directions of a two-way street cannot be well treated. However, this is a rare occurrence.

It sometimes happens that the engineer is faced with a mandate to have protected left turns, and therefore has multiphase signals. If the engineer must give preference to one direction because of congestion/saturation problems and wishes to do as well as possible for the other direction, he has a special opportunity that should be noted—namely, phase arrangement. By selective use of the phase arrangement at successive intersections, the engineer can enhance the "window" or progression presented to the reverse-direction traffic flow.

Consider a situation in which the northbound through traffic is to be given a simple progression. The restrictions are as follows: 80-sec cycle length; main street green not to exceed 45 sec; through movements require 30 sec; turns require 10 sec; vehicle speed of 45 fps; block lengths of 1000 ft. The reverse direction is to be helped as much as possible within this context.

Figure J-17 illustrates the four possible phase arrangements. In general, the solutions will be characterized by the following:

1. The pattern will start with an extreme shift of the southbound through movement (i.e., option 2 or option 3 of Figure J-17).

2. The pattern will alternate between these two options along the arterial.

3. The excess green will be allocated to the through movements because of their greater importance (i.e., 45 sec available minus 10 sec turn minimum = 35 sec through), except for the entry control points; these points—the southmost northbound through green and northmost southbound through green—would not receive the excess green if it would advance the initiation times at these points.

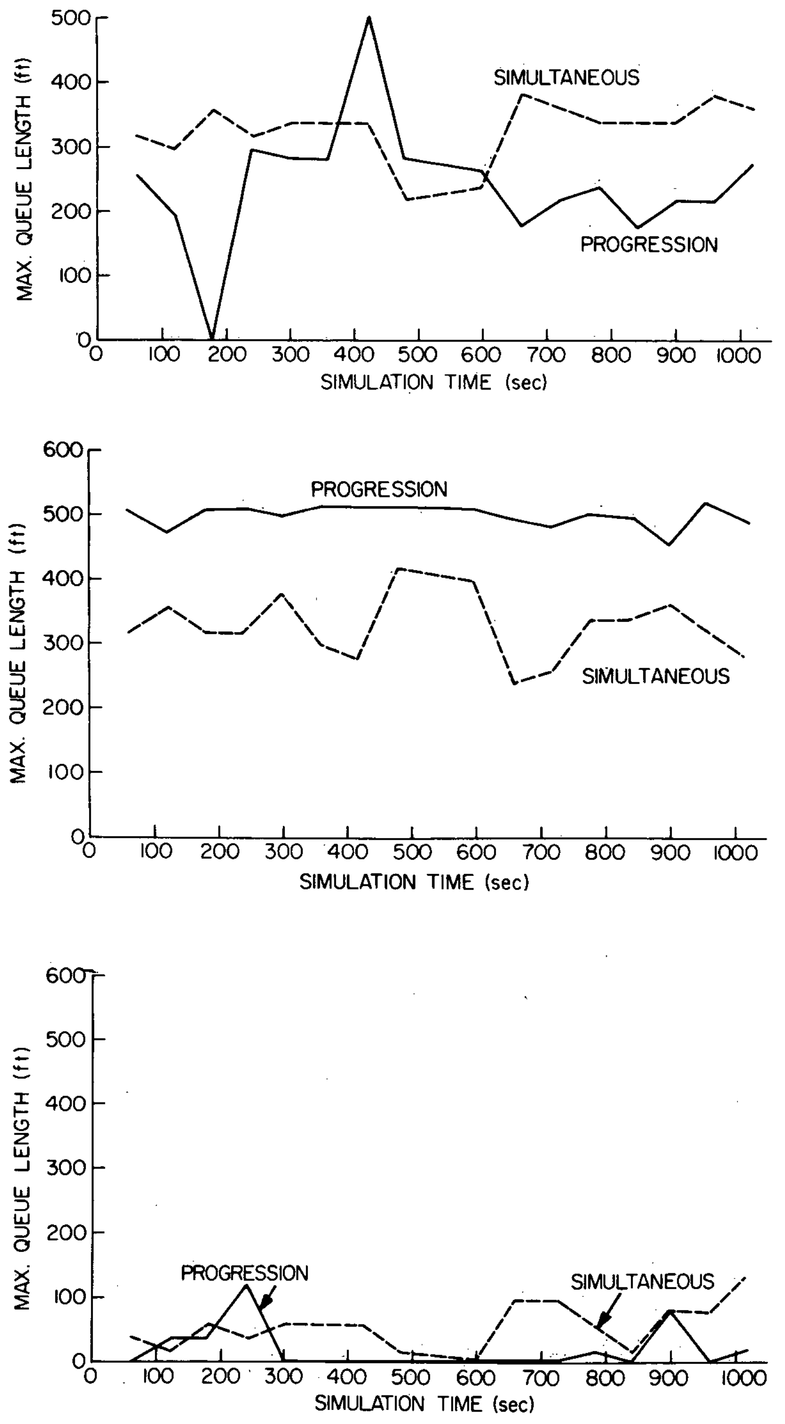


Figure J-12. Maximum queue length under progression and simulation offsets.

There are two possible extreme starting solutions (options 2 and 3). It is therefore necessary to select between them; a time-space diagram trace of the impact on a southbound platoon is recommended. Figure J-18 illustrates the better solution for the problem cited.

If the sum of the desired offsets happens to equal the cycle length in all links, a simple solution can be attained: combinations of options 1 and 4, or specification of option 1 or option 4.

Equity Offsets

It is unfortunately true that it is not always possible to avoid spillback, for there may be too many vehicles attempting to enter a particular link. With the extreme traffic congestion, it is not uncommon to see vehicles storing themselves in the intersection, to the detriment of the cross traffic. Figure J-19 illustrates such a situation.

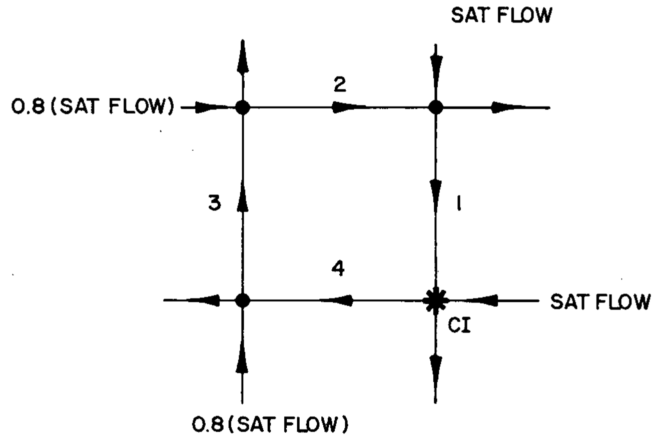


Figure J-13. A sample network.

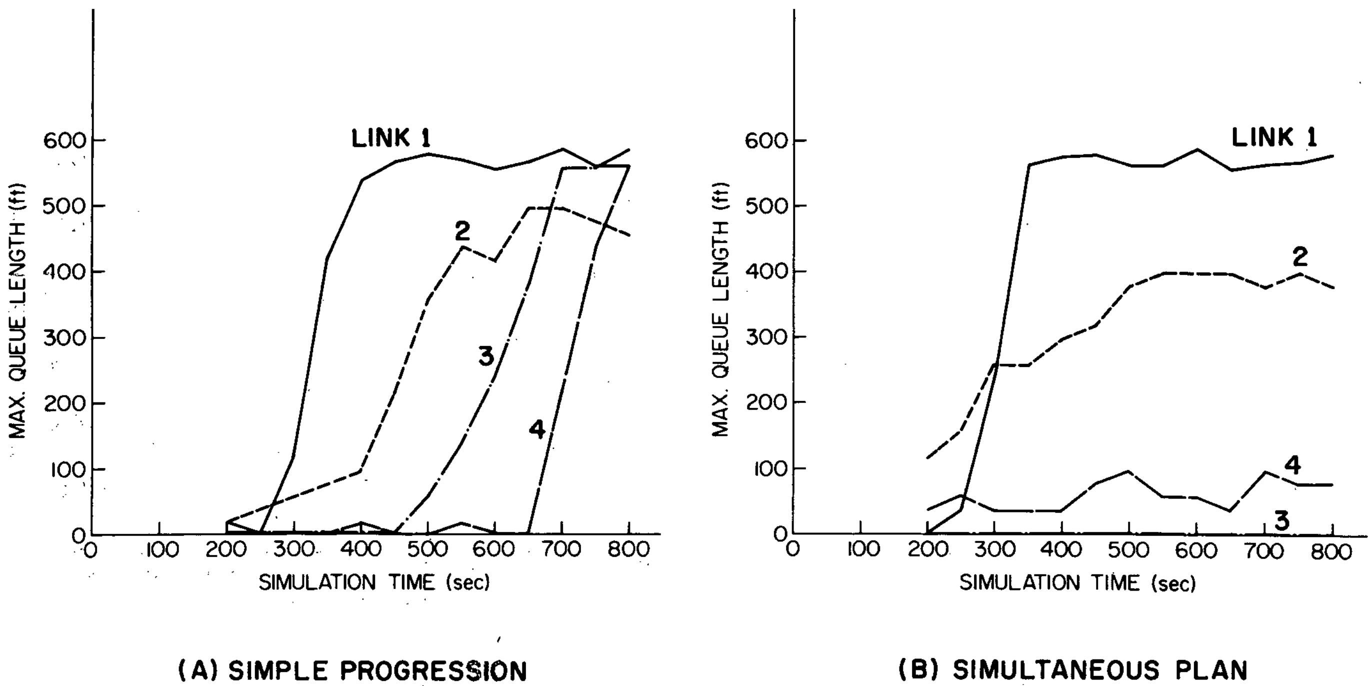


Figure J-14. Simulation of simple progression and simulation offset plans for Figure J-13 network (minor turns).

One common solution to this spillback problem is to place a skilled traffic control officer at this site to prevent such events. Another approach is an intensive ticketing program for such offenders. The former approach is not only historically more effective, but it is also the one demanded by a public afflicted with spillback.

A possible alternative solution exists by changing the basic concept of what the offset is supposed to accomplish. However, this solution should not be implemented until one is certain that a better offset cannot alleviate the problem, as addressed under subsection "Progression of Traffic Using Offsets." The treatment to be presented now is only for that period after the best possible offset has failed because of the sheer volume demanding access to the link.

The treatment, shown in Figure J-20, is as follows: allow the oversaturated direction to have green until the vehicles blocking the intersection just begin to move; switch green to the cross traffic; allow the cross stream to move until it has had an equitable input into the oversaturated link, or at least into the intersection.

The offset that accomplishes this function is defined herein as equity offset because of the objective of equitable treatment of the competing flows. It is not determined by Eq. J-6. The upstream red should begin L/V_{ACC} sec after downstream green initiation, where V_{ACC} is the acceleration wave speed in fps. Assume g_1 as the green time at the upstream intersection (percent of cycle) and g_{CI} as the green time at the critical intersection. Thus

$$t_{off} = g_1 C - \frac{L}{V_{ACC}} \qquad \text{(J-7)}$$

where C is the cycle length in seconds. Typically, $V_{ACC} \simeq 16$ fps.

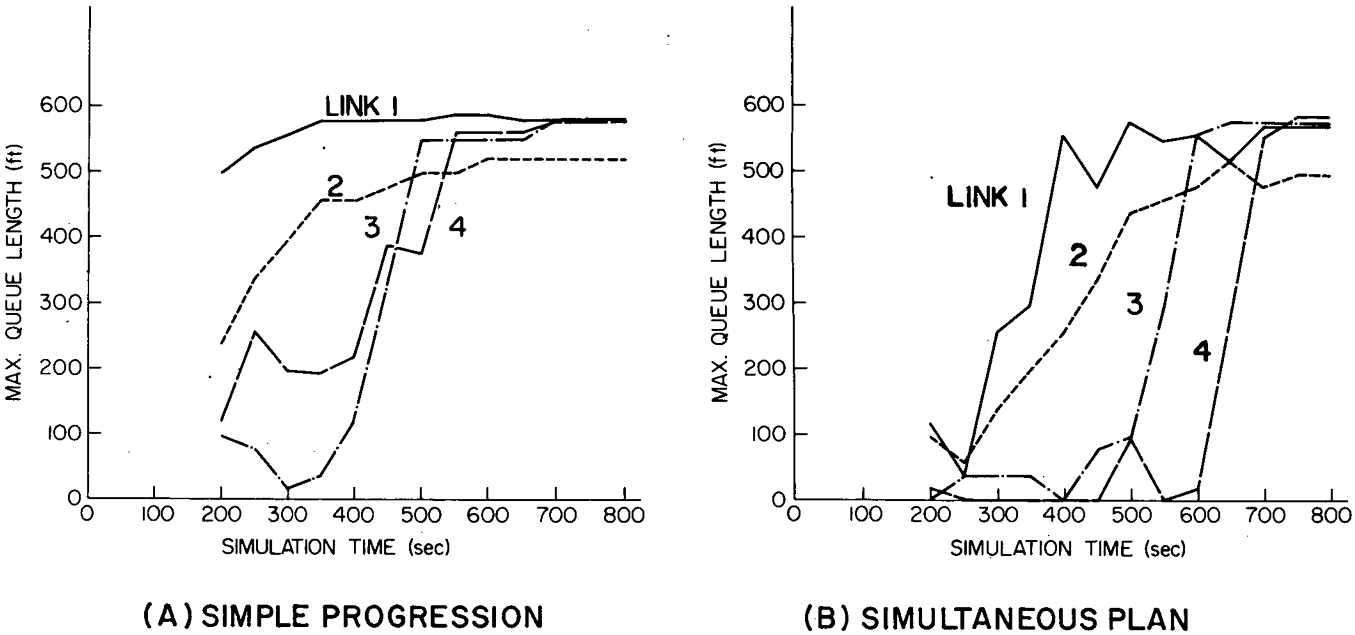


Figure J-15. Simulation of simple progression and simulation offset plans for Figure J-13 network (major turns).

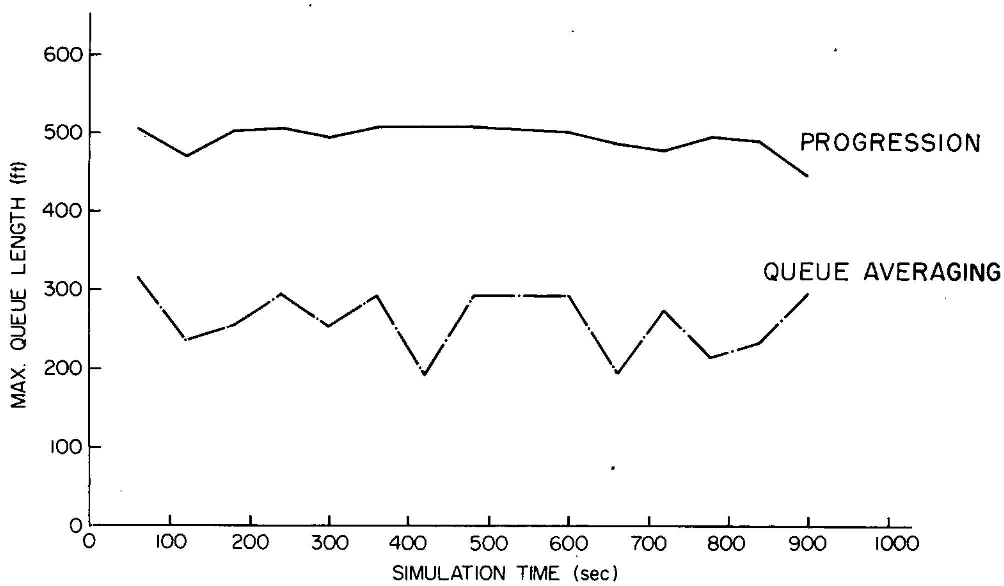


Figure J-16. Simple progression vs. queue averaging.

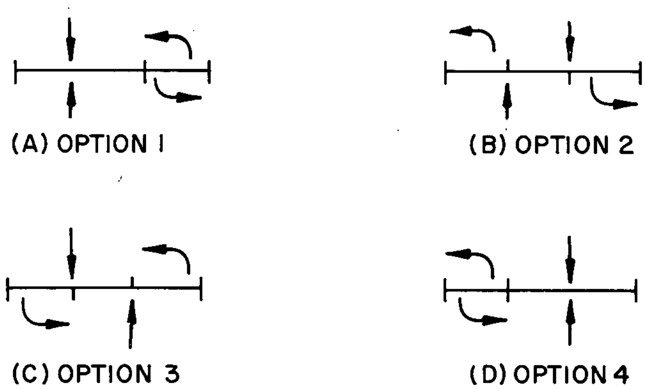


Figure J-17. Options for phase arrangement.

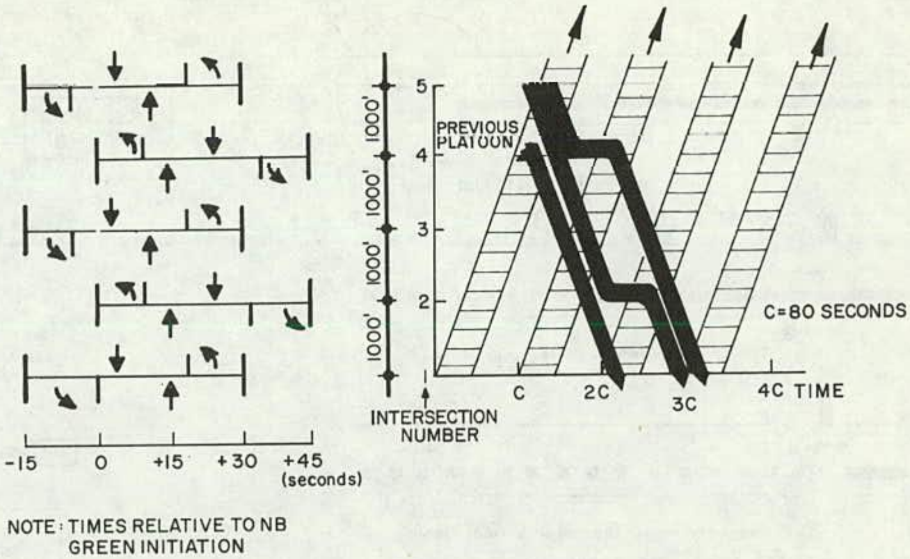


Figure J-18. Solution to phase sequence illustration.

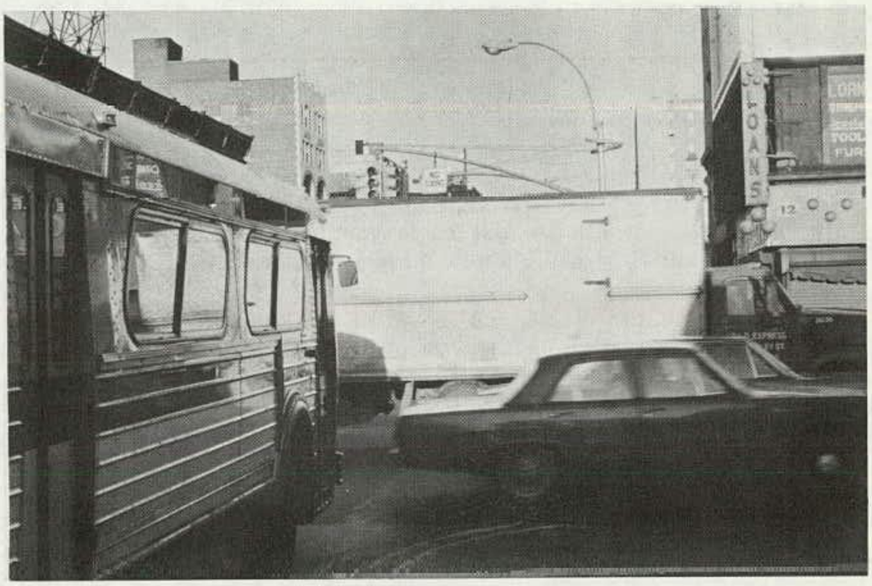


Figure J-19. Intersection blockage.

There are actually two cases of common interest. In the first case, there are negligible turn-ins from the cross traffic. The split at the upstream intersection can be as determined in subsection on "Splits" under section heading "Minimal Response Signal Remedies—Intersection."

In the second case, there are substantial turn-ins. In general, the cross-street should be allowed just enough green to put its "fair share" of vehicles in the oversaturated link. If the turn-ins "steal" all the storage space, the arterial through vehicles cannot enter and the oversaturation will propagate rapidly up the arterial.

It is counterproductive to prepare equations or nomographs to cover all approach width combinations and volume/turn combinations. The following principles should govern:

1. The cross-street green should not be so long as to allow turn-ins to take disproportionate amounts of the storage in the oversaturated link.

2. It should be recognized that if the turning demand accumulates, it will interfere with the through movement on the cross street. This is the case in Figure J-15. If possible, storage should be provided.

3. If the turns are indeed significant, and storage and capacity exist, establishment of turn lanes with separate signalization should be considered. In this way, the through movement could be continued.

4. In allowing such cross through movements, it should be recalled that the through arterial movement needs only as much green as it can effectively use at the critical intersection. Any additional green is in fact wasted, and is

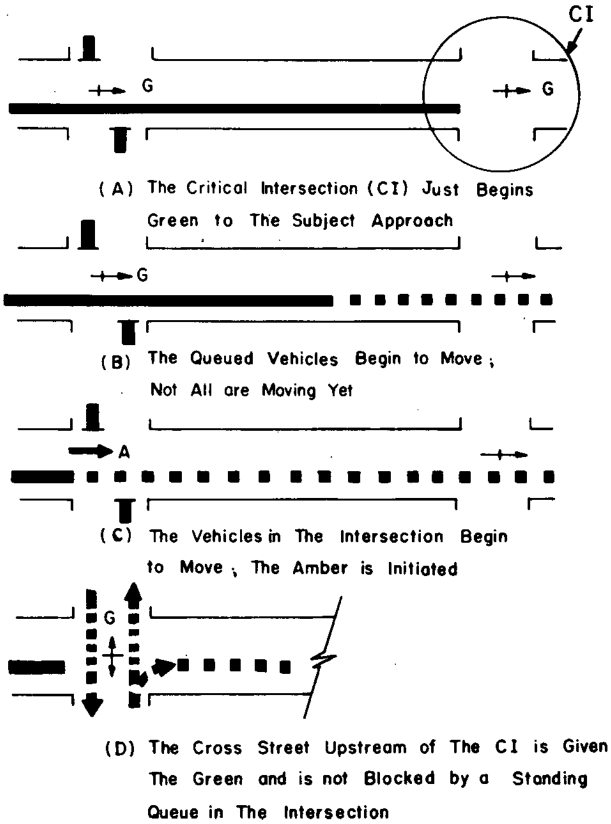


Figure J-20. Illustration of equity offsets.

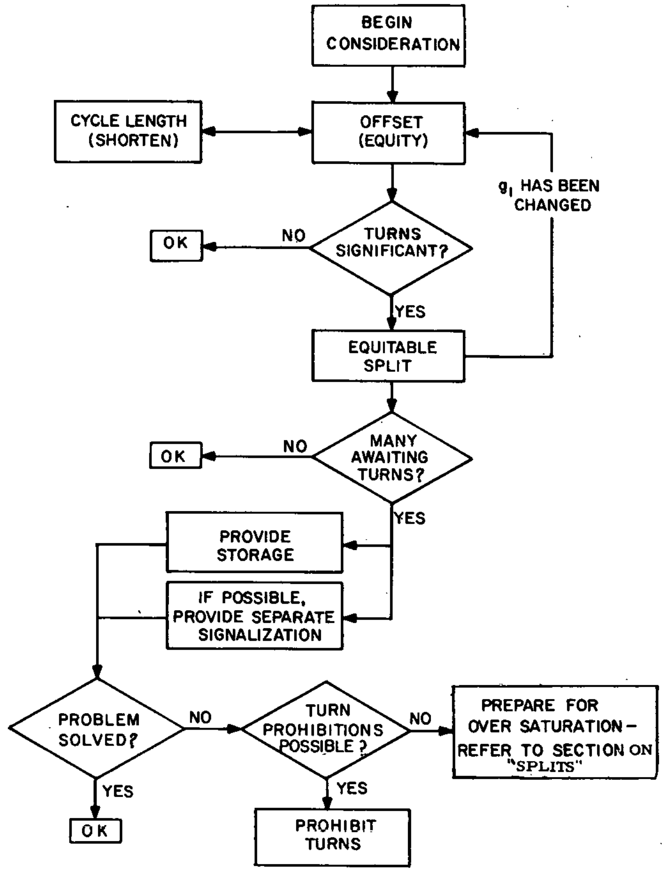


Figure J-21. Flow diagram of hierarchy of solutions for over-saturated link that cannot be remedied.

only allocated to the arterial to keep it from the cross turning movement (if signalization by movement does not exist).

5. If necessary, turn prohibitions should be considered, so that the cross-stream through movement is not severely impacted by queued turning vehicles.

6. If none of the foregoing is feasible, refer to the discussion under "Splits" on the regulation of the spread of (unavoidable) congestion from the upstream intersection.

Note that a hierarchy of solutions has been portrayed, as diagrammed in Figure J-21. More hierarchies of this sort are presented in later sections, where nonsignal options are considered.

Recall that it is assumed that the oversaturation is unavoidable in the one link. Only then is equity offset implemented. The hierarchy of Figure J-21 applies to other (cross) links. It does not apply to the original (unavoidable oversaturated) link.

The original link must have unavoidable saturation: neither any signal nor any available nonsignal remedy could have helped it. Only then is this link abandoned and the best possible is done for other traffic.

To illustrate the impact of offsets, Figure J-22 depicts an arterial on which the volume assures oversaturation, at least of links 2 and 3. There are no turns. The equity offset for link 2 is computed as

$$t_{off,\,2} = (0.6)\,60 - \frac{600}{16} = -1.5 \text{ sec}$$

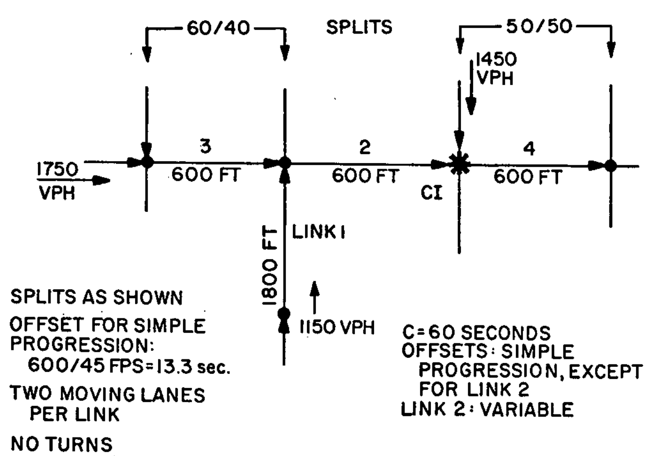


Figure J-22. Illustration of impact of offsets.

so that a simultaneous offset will provide an equity offset function.

Figures J-23 through J-26 illustrate the UTCS-1 simulator output for the per lane queue and the percent occupancy of each of the four links of interest. The only offset that is varied is the one in link 2.

It is readily apparent from Figure J-23 that the equity offset (OFFSET = 0) is quite important to the cross-street traffic (link 1). Just as clearly, no offset can help link 2 (Fig. J-24) because there is just too much demand. Link 3 (Fig. J-25) suffers somewhat from the equity offset, because not as many vehicles can force their way out of the link by occupying the release intersection. Link 4, the downstream link, is not affected by the offset changes, as one would expect; this is shown in Figure J-26.

Apportioning Splits and Downstream Effects

Splits to apportion available storage and downstream block length considerations are covered in a previous section, "Minimal Response Signal Remedies—Intersection," under the subheadings "Splits" and "Downstream Block Length." Offsets are discussed in this section under the subheading "Progression of Traffic Using Offsets." The engineer must always remember that the downstream link is part of his system, and can, in fact, control what can be done at a critical intersection.

HIGHLY RESPONSIVE SIGNAL CONTROL

The emphasis in these guidelines is on minimal response signal policies and on nonsignal remedies used in conjunction with such signal policies. Highly responsive signal control policies are not given prime emphasis for three reasons:

1. The regularity in the traffic demand from day-to-day typically supports the utility and viability of minimal response policies.
2. The extensive detectorization and computer-oriented control required with such policies is beyond the resources, both budgetary and staff orientation, of most jurisdictions.
3. There is major work underway, or relatively recently completed, on other projects addressing this topic. (The work in Washington, D.C. (three generations of the Urban Traffic Control System—UTCS) and that in Toronto, Canada, are examples.)

It should be noted that standard actuated equipment is quite validly classified as a "highly responsive/intersection" type of control. This is true, and properly set actuated equipment can indeed be used to reduce congestion at a site. The timing of such equipment, including proper detector location, is covered in detail in Ref. (*J-7*) and the references cited therein.

It must be further noted, however, that as congestion tends to saturation and oversaturation, cycle length and offset impose a necessary discipline on traffic. Thus, these guidelines, with their consideration of highly responsive signal policies, become more applicable.

This section briefly outlines the relative advantages and disadvantages of some congestion-tailored responsive policies that appear in the literature. It also introduces an additional responsive policy that was developed in this research program.

Existing Policies—Intersection

Three relevant control concepts were selected from a review of the literature: (1) smooth flow concept, (2) minimum delay via split switching (Gazis), and (3) queue proportionality in real-time (Longley). These policies are reported in Refs. (*J-8, J-9, J-10*). A fourth policy, termed "queue-actuated control," was developed and is reported in Ref. (*J-11*).

The smooth flow concept has as its primary objective the real-time adjustment of the offsets as written in Eq. J-8:

$$t_{off} = \left\{ \frac{L}{V} - \frac{Q}{R} \right\} \qquad \text{(J-8)}$$

The Gazis' minimum delay policy finds an optimum time to switch the split, and the optimal split values, with the objective of minimum total delay at an intersection. It is not truly an on-line highly responsive policy, but rather an optimization theory based on advanced mathematics and detailed knowledge of the demand functions. The Longley policy is not unlike a real-time, rapid adaptive version of the split policy stated under subsection on "Splits."

Table J-7 summarizes some of the important disadvantages of each of these policies and of the queue-actuated policy. The queue-actuated policy switches from phase-to-phase when certain (determined) maximum queue extents are reached, with the objective of minimizing average delay while not exceeding certain limits of queue extent. It does not depend on queue measurement as such (i.e., number of vehicles), but rather on attaining a certain threshold value (i.e., queue of 12—yes or no).

Extensive simulations (*J-11*) indicate that of the three concepts first cited, Longley's policy is generally the most effective. Queue-actuated control is the most effective at high volume levels and in environments with heavy turns. Because of its deliberate formation of queues, queue-actuated control also has the advantage of greater productivity.

Figure J-27 shows the average delay experienced at an isolated intersection under three different control policies. The *degree of saturation* is as defined by Webster "the ratio of the actual flow to the maximum flow which can be passed through the intersection from one arm."

Queue-actuated control can be considered as a viable control policy in situations where minimal response policies (that is, minimal response policies in conjunction with nonsignal remedies) would, or do, fail. However, this does not mean that minimal response policies do not perform better in many situations. Indeed, minimal response policies are better than queue-actuated control on a system basis.

Existing Hardware—Intersection

As mentioned earlier, in the introductory remarks to this section, standard actuated equipment is validly classified as a highly responsive intersection type of control that can be used to reduce congestion. Details on existing

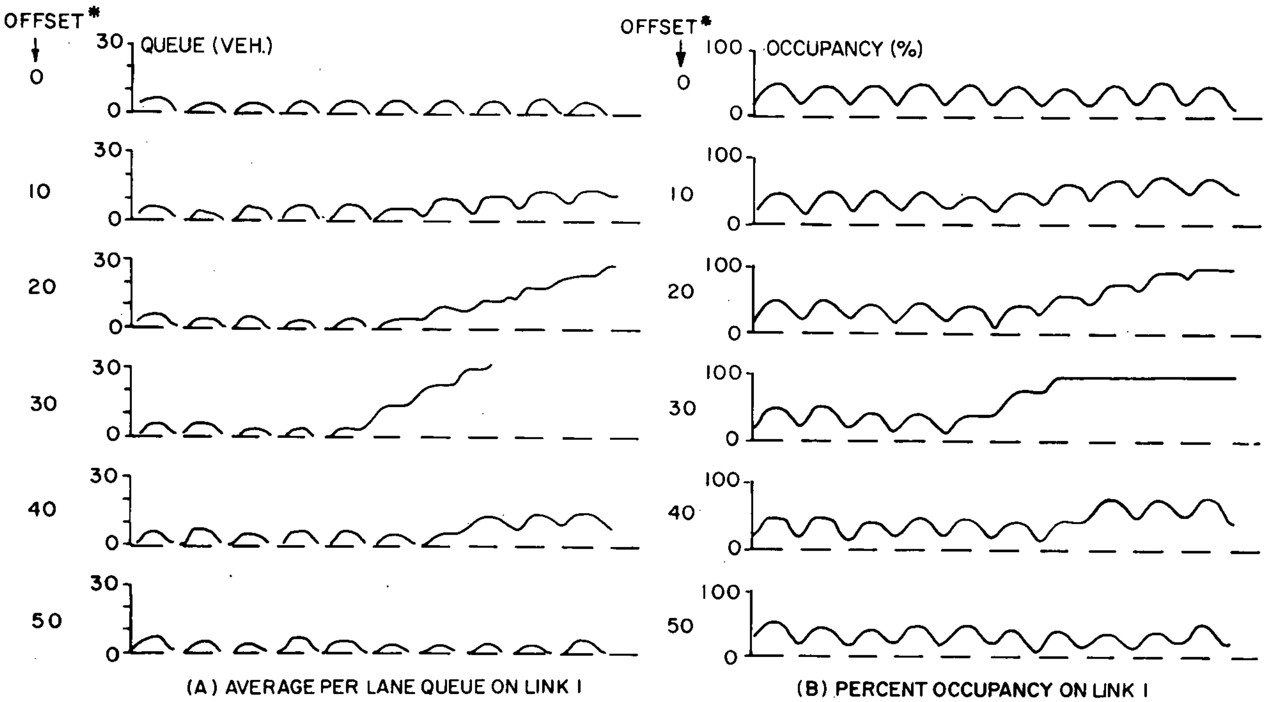


Figure J-23. Average queue and occupancy on Link 1.

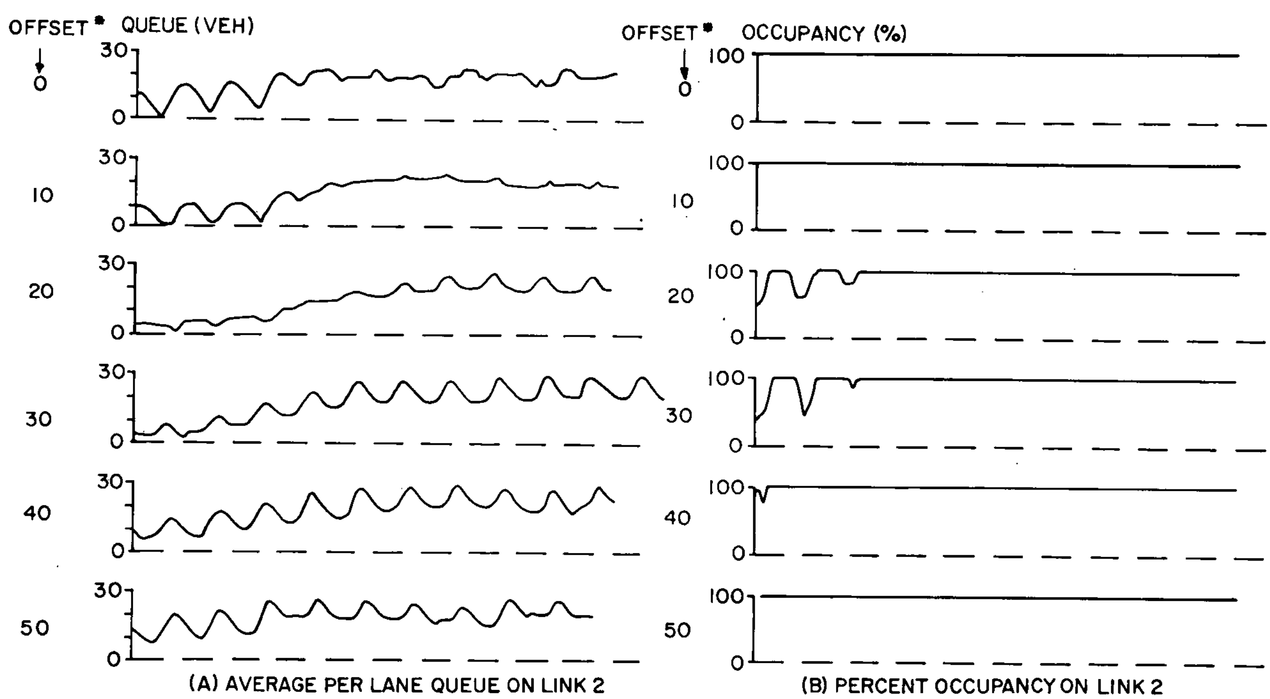


Figure J-24. Average queue and occupancy on Link 2.

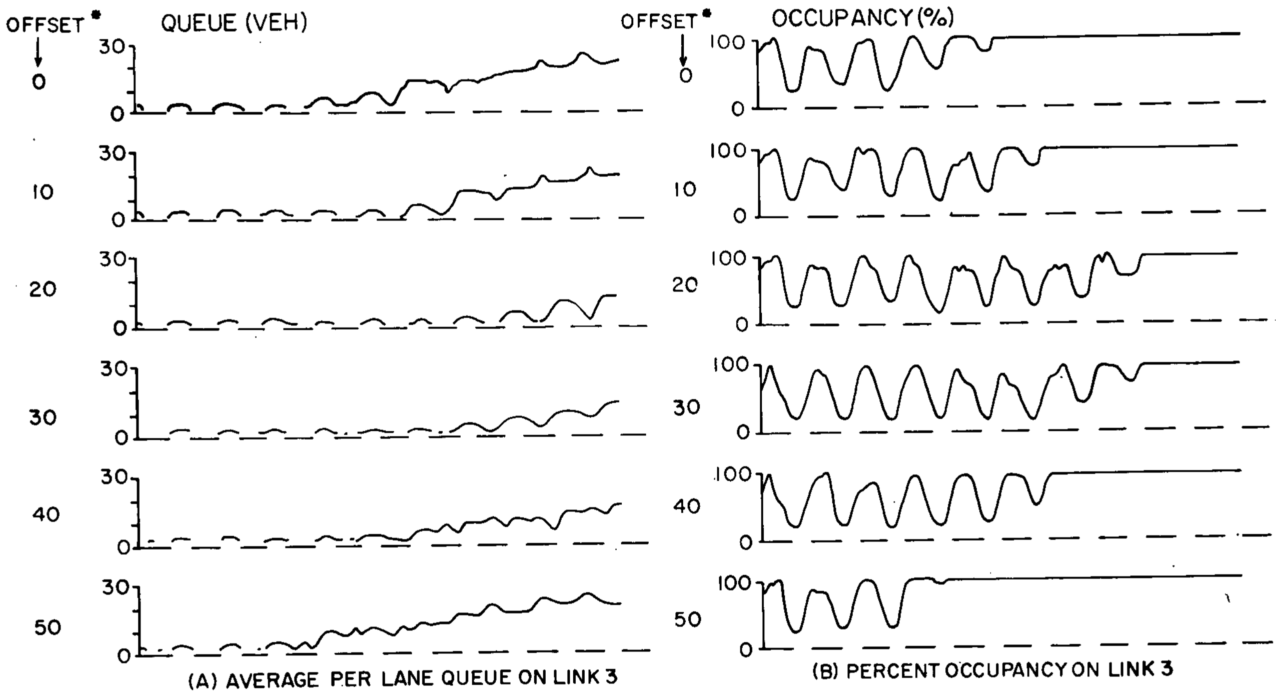


Figure J-25. Average queue and occupancy on Link 3.

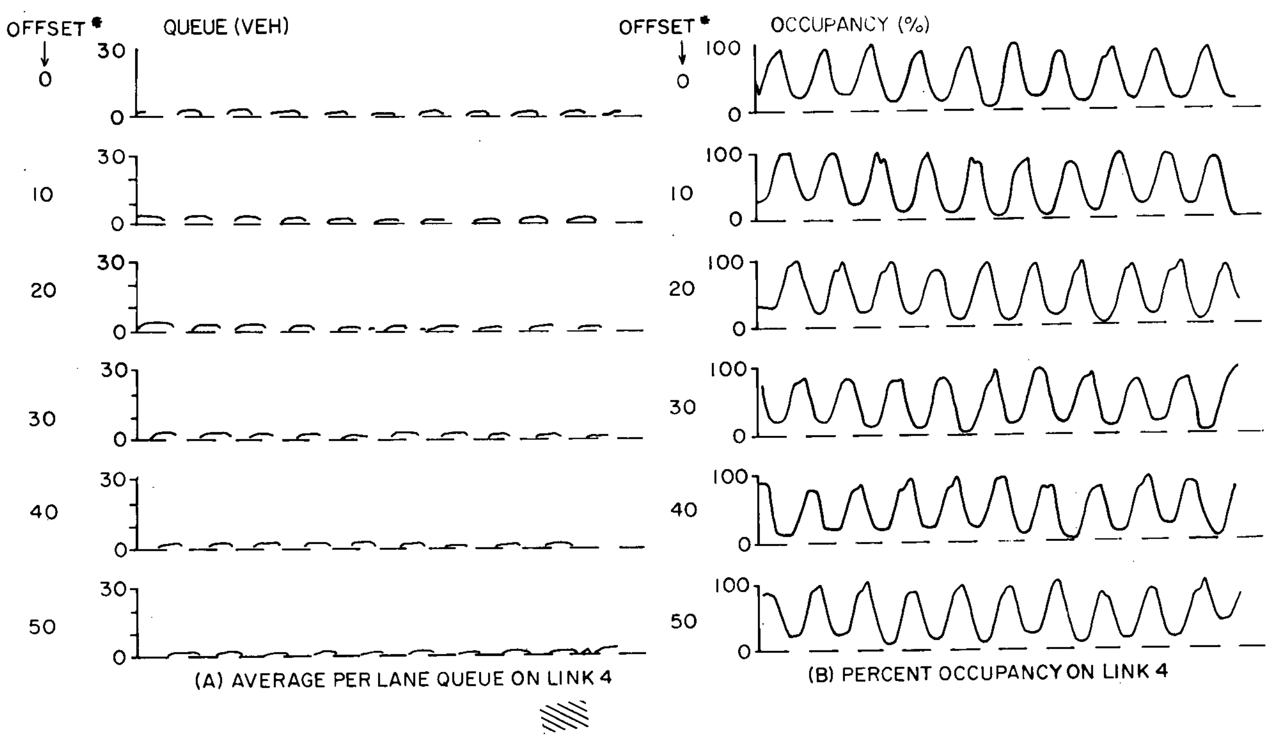


Figure J-26. Average queue and occupancy on Link 4.

TABLE J-7

DISADVANTAGES OF SPECIFIC HIGHLY RESPONSIVE POLICIES

SMOOTH FLOW CONCEPT • Measurement of Queues Required • On-Line Computations
GAZIS' MINIMUM DELAY • Demand Function Description Needed • Queue Extent Can Be Excessive • Applicable Only to Convex Portion of Demand Curve
LONGLEY'S POLICY • Measurement of Queues Required • On-Line Computations
QUEUE-ACTUATED CONTROL • Detectors Located According to Competing Demand (Historic) • On-Line Computations

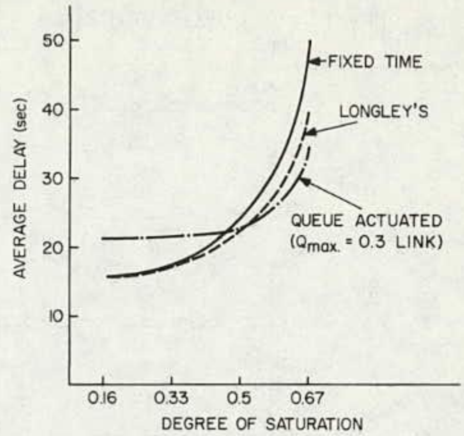


Figure J-27. Delay comparison of different control measures: fixed time, Longleys' control, and queue-actuated control at an isolated intersection.

hardware are well covered in Ref. (*J-7*) and will not be considered further here.

System Considerations

Highly responsive control of congested systems is, in fact, the control of congestion as it spreads from a critical intersection, contaminating other intersections. Minimal-response, preprogrammed control is often quite adequate and appropriate, and detailed treatment of system control in a highly responsive mode is beyond the scope of these guidelines. Should it be deemed appropriate, and if the engineer has been, or is, considering a computer-based system, he is referenced to the UTCS work in Washington, D.C., and to the ASCOT work developed for NCHRP (*J-16*).

NONSIGNAL CONTROLS—ENFORCEMENT AND PROHIBITION

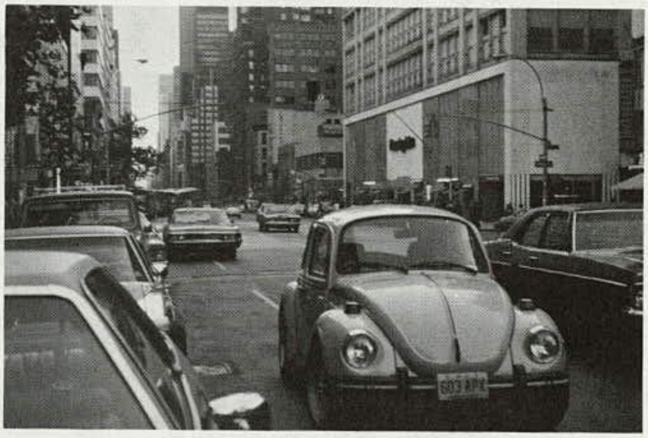
Proper enforcement can well be the most powerful nonsignal control treatment available, although it is not usually within the duties of the traffic engineering staff.

At the same time, the engineer should have some understanding of how much benefit he will receive if a strict enforcement program clears the observed infraction pattern. Double-parking (see Fig. J-28) is clearly an impedance to traffic, but does it, in fact, affect the street's throughput or delay pattern?

Although "prohibitions" can be classified as a regulatory measure, they can also be viewed very much as an operations measure. It is indicated as both in Table J-5.

The material herein on prohibitions should only be used after assessing the benefits that can be realized with turn bays, with and without separate phasing. Prohibitions are generally a means of achieving other treatments (for example, providing an additional lane via parking prohibition), and not themselves treatments.

(A) TRUCK AFFECTING TRAFFIC



(B) MULTIPLE DOUBLE PARKERS

Figure J-28. Illustration of double parking situations.

Enforcement

Issues and Considerations

Viewed strictly in terms of congested operation, the issue of enforcement reduces to one question: Is the specific enforcement being considered worth the effort? The answer to this question is not simple, and must be considered in light of the following:

1. The benefit that would be realized if the infractions (such as double parkers) could be cleared effectively.
2. The effort that would be necessary to attain the desired or necessary level of success, in terms of both cost and manpower.
3. The resultant adverse impact that would be experienced by particular groups or the community at large (such as establishments losing shoppers, etc.).

As an example, suppose the street in Figure J-29 needed more capacity, according to a preliminary analysis. Should the truck-loading zone be removed to provide this capacity? The answer requires a complex analysis. The trucks will have to deliver their goods, and—if not permitted in such a zone—will double park. It is known (*J-11*) that they will almost always park within 100 ft of their destination, regardless of regulations. How much of an impact will that behavior cause?

The engineer must consider options. Can removal of some of the zone provide a sufficient discharge capacity? Are other options from Table J-5 available to solve the problem?

Enforcement and reasonableness of that which is being enforced are interissued. Further, it must be recognized that certain solutions only hide or transfer the problem or are doomed to defeat. For example, if the truck zone were created to solve a double-parking problem (which is not now apparent, because it is solved), its removal will not provide extra capacity—it will simply bring forth parking violations. Likewise, actively keeping potential double parkers on the move may simply add to the traffic volume as the vehicles circulate looking for a parking spot.

Figure J-30 illustrates the nature and extent of double parking in one, 4-lane one-way link. The figure can be explained by noting that a number of vehicles would stand for some time, leaving whenever a law enforcement agent or emergency vehicle passed through the block, but returning shortly. The average duration of double parking (including standees) was 12.8 min.

Figure J-31 presents the results of a simulation of the impact of double parkers in a 600-ft 3-lane link. The same curve could be used for three moving lanes without an additional (i.e., a fourth) lane for parking. The figure was computed for various V/C ratios of the moving traffic, various double-parking positions, and a double-parking duration of 12 min.

Examination of this figure reveals that double parkers have a substantial impact for volumes for which V/C exceeds 0.75, and the impact is least severe when the double parker is at the center of the link. It is most severe when he is at the discharge end of the link. Moreover, not only does the impact grow substantially with V/C ratio,

Figure J-29. Side street that may need additional capacity.

but the impact also persists well beyond the double-parking duration at the higher V/C ratios.

Note also that the curves in Figure J-31 can be used to estimate the impact of adverse uses of a curb lane from which parking was removed so as to increase capacity. For longer durations of double parking, the values in Figure J-31 can be proportionately adjusted so that they can serve as an approximation (the values in Figure I.31 can be multiplied by (15/12) 100% if, for instance, a 15-min duration is being considered). A later section in these guidelines, under the heading of "Disruptions to the Traffic," addresses bus stops and the impact of buses stopping in lanes that could otherwise be through lanes.

Evaluation of Options

Two of the most frequently persistent violations that aggregate the congestion/oversaturation problems are intersection blockage and parking regulation violations. Equity offsets, discussed earlier, represent a means of circumventing the first, avoiding or delaying the need for on-scene traffic control officers.

The discussion of the previous subsection "Issues and Considerations" indicates the magnitude of impact that the second problem, parking violations, can have. Perhaps this will point out a level of unacceptability to the individual engineer, which will make his decision clear: strict enforcement.

On the other hand, local realities may dictate that strict enforcement is not practical. These realities may include the fact that many of the vehicles are goods vehicles essential to local vitality, or that the vehicles would only circulate, adding to the apparent traffic flow. (Note that one vehicle circling a city block once every 4 min can add 15 vph to the count on each of the four streets in 1 hr.)

Recognizing these facts, one must remedy the estimated impacts by other treatments contained herein. If neither decision is clear-cut (i.e., strict enforcement or compensation via other treatments), the engineer may wish to formulate a cost-benefit or cost-effectiveness analysis to aid him. Delays, such as those indicated in Figure J-31, can be translated into costs.

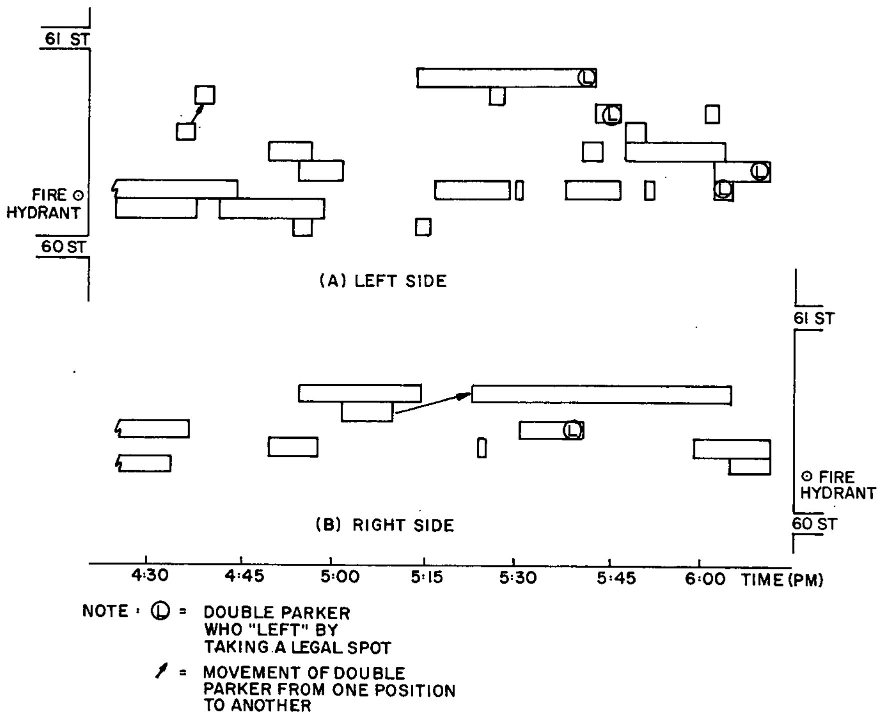


Figure J-30. Double parking during the peak on one, 4-lane one-way link in a downtown area.

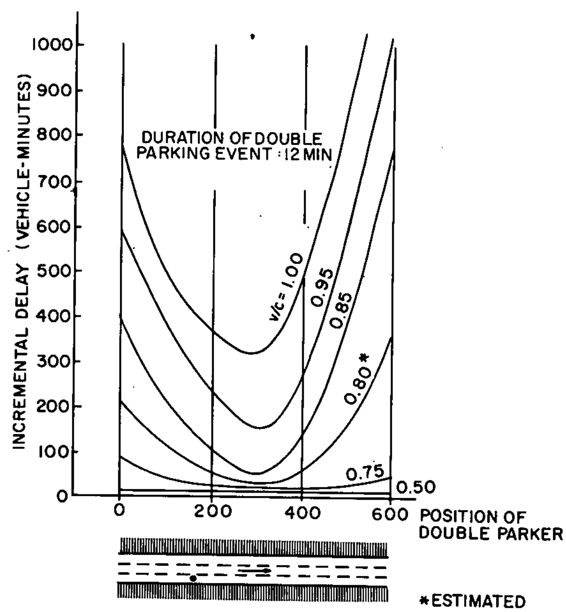


Figure J-31. Illustration of delay caused by double parker.

Prohibitions

Parking

Parking prohibitions would be enforced to provide (when combined with restriping) an additional lane for the length of an arterial, adding to its capacity; an additional lane, or part thereof, for special turn lanes or bays; and an additional lane reserved specifically for transit vehicles. The subject of turn bays is addressed later under "Turn Bays and Other Nonsignal Remedies." Additional lanes added specifically for transit vehicles are, of course, useful in the relief of congestion to the extent that they reduce the burden on existing lanes. Criteria for such lanes, or for the dedication of existing lanes for such purposes, are a matter of policy planning and/or transportation planning beyond the scope of these guidelines.

Nothing in the literature, in the data collected for calibration of known simulators, or in the field work established that any added moving (curb) lane would have a lower capacity than other lanes. Indeed, limited indications are that they are indistinguishable from a capacity standpoint. However, local perceptions of the proper use of curb lanes (i.e., "curbs are for parking . . . for dropping off passengers . . . for standing") lead to adverse uses that will nullify much of the benefits otherwise gained (refer to subsection on "Issues and Considerations").

Turning

Many intersections can have their performance greatly enhanced, and their critical nature removed, by turning prohibitions. However, this does not mean that total system travel time will necessarily be reduced. Moreover, use of turning prohibitions assumes that an option exists for an alternate route.

A number of the critical intersections that would be greatly aided by turning prohibitions are critical intersections only because vehicles must turn there. Often, the engineer's options are limited to getting the "turners" out of everyone else's way. This matter is discussed in greater detail under "Turn Bays and Other Signal Remedies."

To illustrate the impact that prohibiting left turns can have, note that the left-turn equivalence factor is commonly between 2 and 3 (i.e., a left turner "looks like" 2 to 3 through passenger cars). With 10 percent left turners, prohibition of left turns can amount to a 10 percent to 20 percent increase in throughput on the approach; for example:

- X vehicles before, 10 percent of which are left turners, have an equivalent volume of $1.1X$ to $1.2X$.
- Assuming equivalent vehicles equalize by lane, this is a critical lane volume of $(1.1X/N)$ to $(1.2X/N)$, across N lanes.
- Y vehicles after (with prohibition), where $Y \leq X$, yield $(Y/N) \leq (X/N)$.

If actual vehicles equalize by lane, the percent increase would be even more dramatic.

This computation illustrates that left-turn prohibition is one of the most beneficial treatments available. Unfortunately, it is also one that is not possible, in many cases, because of lack of alternate routes.

Before considering the implementation of a turning prohibition, the engineer should have eliminated all those options regarding turn bays (see subsection under "Turn Bays and Other Nonsignal Remedies") and extra phases (see also subsection under "Minimal Response Signal Remedies—Intersection"). If still necessary, the engineer must consider the load that will be put on other links and intersections by the prohibition under consideration. (It may, of course, turn out that moderate-volume alternate routes are in fact readily available. In this rare event, the engineer should accept his good fortune and implement the prohibition.)

Right-turn prohibitions will rarely aid in increasing approach throughput, except when the link into which they are turning is so congested, or the vehicle-pedestrian conflicts so intense, that they are in effect storing themselves while awaiting access (refer to Figure J-15 and related text). In the interests of equity, particularly when right turns are so heavy as to "steal" all available storage, prohibitions may be enacted to allow other flows proper access to links. Reference should be made to the contents of the subsection on "Equity Offsets" before resorting to such a prohibition.

Other Prohibitions

The impacts that prescribe "no parking" will also tend to prescribe "no standing." The benefits that arise from passenger pick-up/drop-off however, particularly near public transit sites, may require even more reflection and consideration before implementing "no stopping" regulations.

The engineer must also consider (to the maximal extent his duties, responsibilities, and powers allow) other uses that can adversely impact his overworked street system—prime among these are driveway placement, particularly when it involves activities that develop queues on public rights-of-way; car washes with entrances on major streets (see Fig. J-32); loading docks; and drive-in windows.

TURN BAYS AND OTHER NONSIGNAL REMEDIES

Turn Bays

The separation of turning vehicles creates one of the best opportunities for alleviating congestion and avoiding oversaturation at a critical intersection via nonsignal treatments.

For right turners, the separation may frequently be obtained by only partial removal of parking. For left turners, the separation may require striping a left-turn bay, causing the removal of a lane of parking in one or more links so as to provide lane continuity.

What should be considered before considering the establishment of turn bays? Certainly, the preliminary work outlined in the sections prior to the discussion of the individual signal solution should have been done. Signal timing should have been revised, if necessary, including cycle length, split, and offset. The engineer, before considering this section, should also have already identified a probable objective:

1. *Decreasing effective green,* because the other phase at the CI needs it.
2. *Increasing productivity,* because oversaturation is growing and no additional effective green can be obtained from other phases.
3. *Providing a "fringe benefit"* to right turners, so that they are not delayed by downstream congestion which they will not enter.

Having identified an objective, the engineer should then make a decision with the objective, and the appropriate evaluation measures used, in mind.

Right-Turn Isolation

The creation of a right-turn bay allows an increase in productivity on the approach or, if it is desired, a decrease in the effective green allocated to the phase. It also

Figure J-32. Lane lost by standing queues on right-of-way.

permits a decrease in the local delay, with the right turners realizing most of the delay savings. The increase in productivity, expressed as a percentage, can be comparable to the right-turn percentage.

The length of the turn bay should not be much shorter than the queue that typically forms. In this way, maximum presorting can occur. Thus, in practical terms, short cycle lengths aid this objective, because released platoons are smaller.

On the other hand, the following example will illustrate that much of the benefit is achieved by the existence of a bay of moderate size. Short cycle lengths, nonetheless, aid presorting and should be used as a companion measure.

Figure J-33 shows the impact of right-turn bays (note that left-turn bays on a one-way street are considered the same as right-turn bays). The example illustrates a 2-lane arterial with input flows capable of oversaturating it, with oversaturation initiating at the indicated CI. Three cases are studied: (1) no turn bays, (2) turn bays of length 5 vehicles, and (3) turn bays of length 9 vehicles. Figure J-34 depicts the delay per vehicle under the three conditions. The standard deviation measure is the standard deviation of the values in successive time periods of 150 sec each. Thus, it is a measure of the growth of delay with time; a large variance indicates that the measure is tending to catastrophic failure (i.e., system oversaturation).

Figure J-35 shows the spillback history (seconds of spillback in a 30-min simulation) along the arterial upstream of the critical intersection.

It is interesting that the turn bays, which serve some benefit to the system (witness Fig. J-34), do not necessarily provide a remedy. The volumes assure that there will be oversaturation eventually. In link 2, where the CI implies longer queues, it is only the longest bay that has any impact. Link 3 has more discharge capacity than link 2, and did not benefit as much from the turn bays. (Recall from the very definition of a CI that link 2

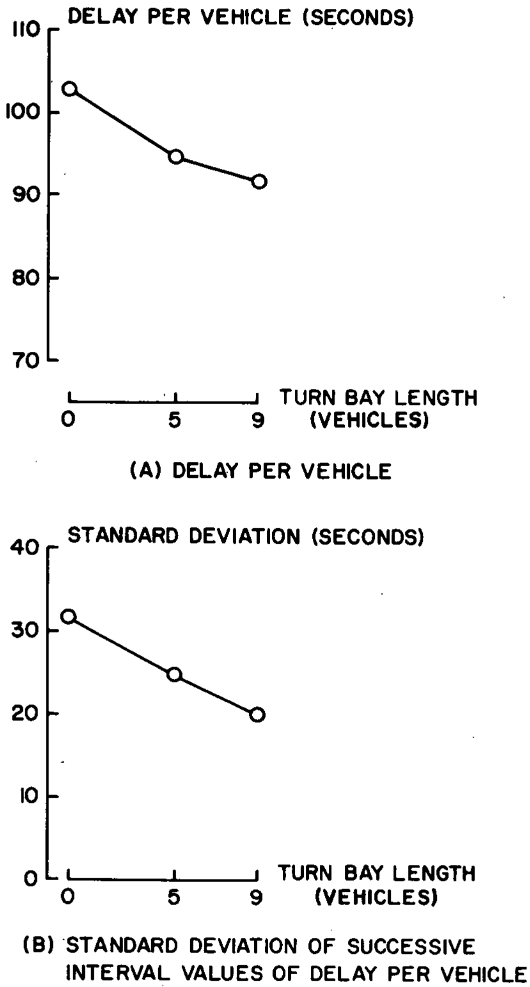


Figure J-34. System results for turn bays.

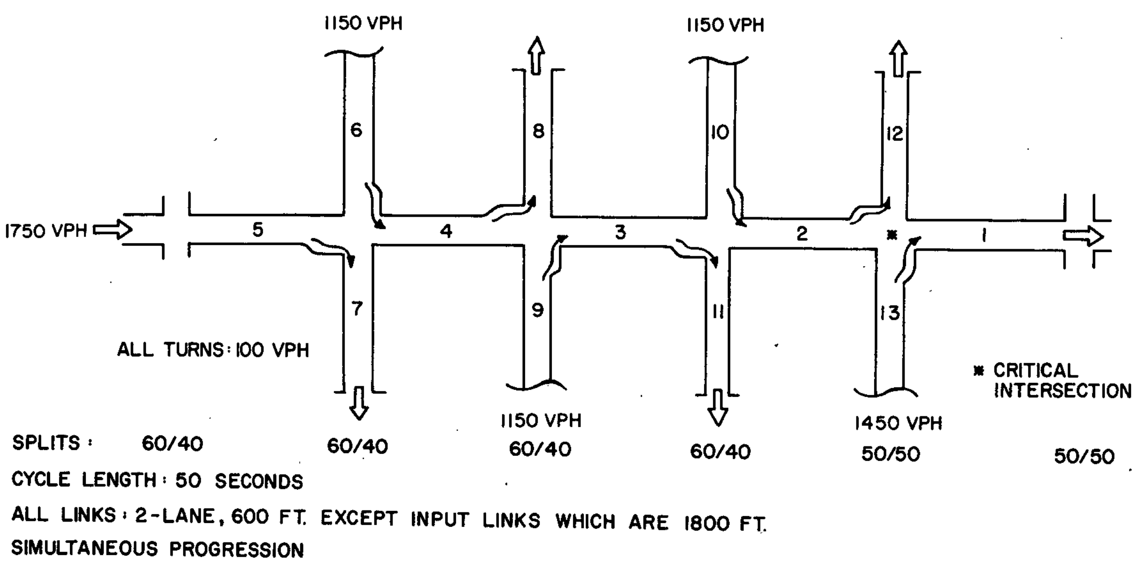


Figure J-33. Illustration with right-turn bays.

could not previously handle the demand being put on it. That is, its discharge point—the intersection—could not.) Link 4, being further away from the CI, had shorter queues and thus benefited more than link 3.

Comparison of Turn Bay and Signal Treatments

It is worthwhile to compare benefits of signal and nonsignal treatments. As was shown earlier, the simultaneous offsets result in the clearance of a queue of about 20 percent of the link lengths. The question arises whether some other offset would be more favorable? Would it have as much effect as the turn bays? Or would it have more effect than the turn bays?

Because the system does oversaturate, equity offsets might be appropriate. For the arterial links upstream of the CI,

$$t_{off} = \left\{ g_1 C - \frac{L}{V_{ACC}} \right\} = 0.6(50) - \frac{600}{16}$$

$$t_{off} = -7.5 \text{ sec}$$

Thus, $t_{off} = 0$ sec is giving the side street the green 7.5 sec before any vehicles that are blocking the intersections have had a chance to move (see Fig. J-36, parts A and B).

A further advance of the offset (initiated earlier does no harm as long as not so many through vehicles are allowed to cross so as to cause a blockage when they do stop (see Fig. J-36, part C; this feature can also be noted in Figs. J-23 through J-26).

In consideration of the foregoing, a "backward progression" of —15 sec offset is selected. It is approximately the reverse of the simple forward progression for this arterial. As such, it allows for about 40 percent queue clearance at 30-mph vehicle speed (i.e., clearance of a queue about 40 percent of the link length).

Figure J-37 shows the spillback history with the backward progression in force. Two results emerge: (1) the signal treatment is more effective than the nonsignal remedy (compare Figs. J-37 and J-35), and (2) longer bays may actually do harm. In this particular case, this is so because turners are being given preferential access to the arterial, contrary to the intent of the oversaturation settings represented by the (approximately) equity offsets. The engineer must take care not to lose sight of his objectives—in this case, the cross-traffic enhancement via equity offset and the turners' enhancement via long-turn bays.

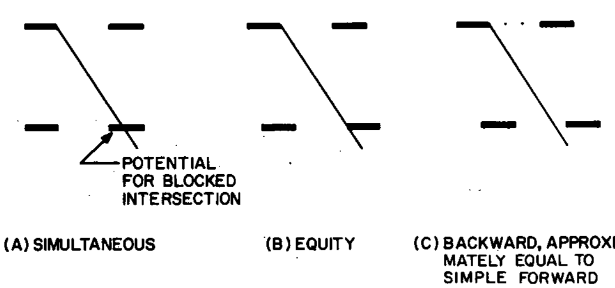


Figure J-35. Spillback in right-turn illustration.

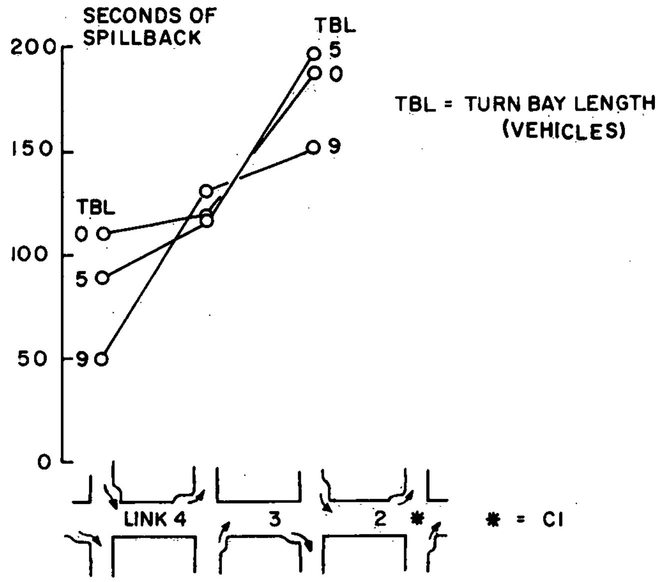


Figure J-36. Comparison of candidate offsets.

The engineer must also take cognizance of one more important matter: Has he unwittingly shifted the CI? The turn-bay treatment has increased the discharge capacity for link 2; more vehicles are able to enter link 1. Because links 1 and 2 have comparable green times, the potential for moving the CI exists. The engineer must not forget to check the downstream impact.

Figure J-38 reflects the results of the turn bays (system-wide) on the downstream link. Note that a backward progression offset, which allows for clearance of a standing queue up to 40 percent of the block length, helps. But the danger of moving the CI exists and must be considered.

In the previous examples, all turns were at the rate of 100 vph. Figure J-39 depicts the delay per vehicle for a range of turning volumes and bay lengths. The fact that more of the volume is turning causes the queue pattern to shift in a way that benefits the delay. This can be

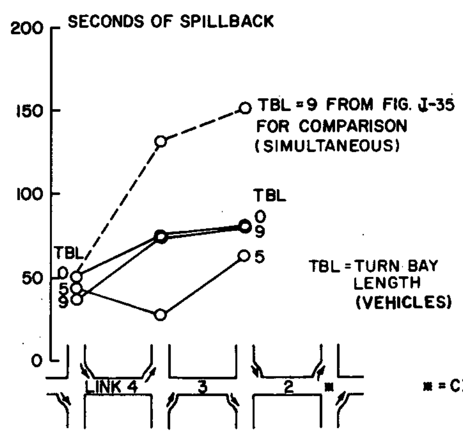


Figure J-37. Spillback in right-turn illustration using backward progression.

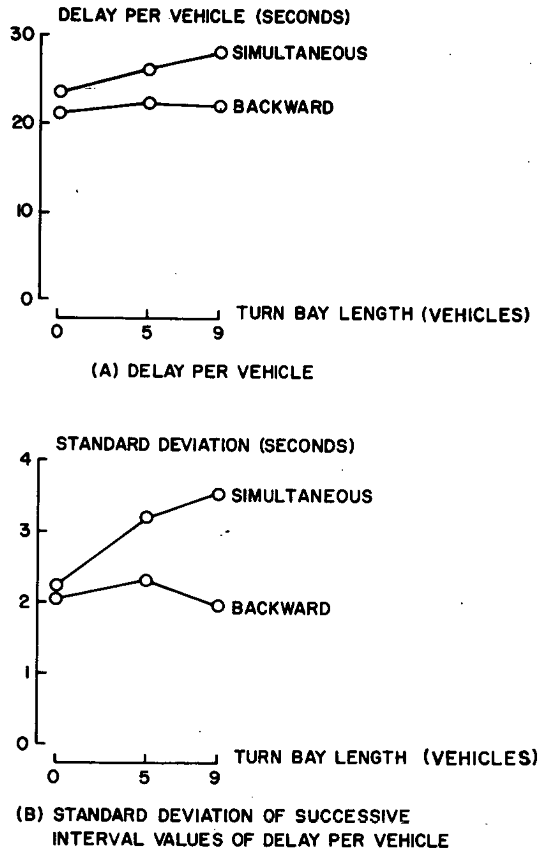


Figure J-38. Results for Link 1 of system with turn bays.

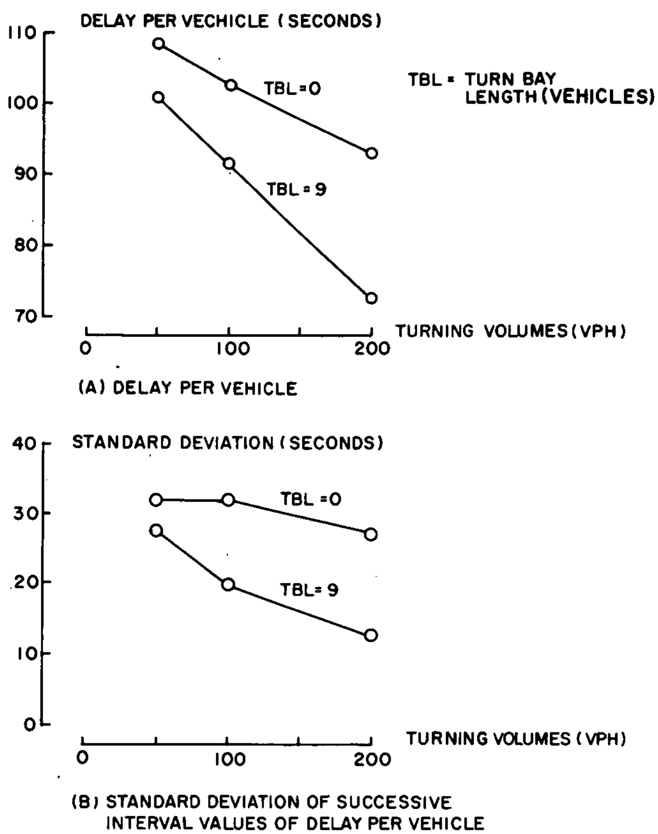


Figure J-39. System results for different turning volumes and bay lengths.

seen on the TBL = 0 lines. However, the advantage of the TBL = 9 is not simply parallel to that of TBL = 0. Indeed, as is to be expected, the percentage improvement grows with volume: 50 vph = 6.7 percent, 100 vph = 10.8 percent, and 200 vph = 21.8 percent. As can be shown with some simple analytics (*11*), the productivity of the intersection can increase by the percentage of turners when a sufficiently long bay is added. Because the turners occupy a different slot, the queue extent decreases by the percentage of turners. Because the delay is determined by the queue extent, the delay therefore decreases.

It is desired that the queue extent be not much longer than the turn-bay length for maximum advantage. In the case illustrated, 1750 vph/2 lanes on the arterial and $C = 50$ sec assure 24 veh/cycle in two lanes. This means 12 veh/lane or less (when the turners enter the bay), for which TBL = 9, and is as desired.

Left-Turn Isolation

It was mentioned earlier that prohibition of left turns can increase the productivity of an intersection by from 1 to 2 times the percentage of turners. Actually, four situations, which may be considered as the prototype cases, prevail, as depicted in Figure J-40: (1) the base condition (no turn bay); (2) the bay which vehicles enter from a moving stream, leaving gaps in it; (3) the bay which vehicles presort themselves into, allowing the through vehicles to "close ranks"; and (4) prohibition. Estimates of the per lane discharge that must be accommodated are based on the equations shown in Figure J-40, where N is the number of vehicles, p_T is the percent left turners, F is the equivalency factor for left turners, and L is the number of lanes.

The estimated increase in productivity, expressed as a percentage, is summarized in Table J-8. The use of the productivity increase can also be considered to allow a decrease in effective green.

The turn-bay length should be about twice the expected number of left turners per cycle if the "bay entered, gap left" situation is of interest. The turn-bay length should be almost as long as the expected queue if the "bay entered, vehicles move up" situation is of interest. The former anticipates diversion of the vehicles from a moving stream. The latter anticipates "sorting out" from a stopped platoon.

The engineer will rarely have the space to meet the latter requirements. Still, actual field situations will be a mix of these two prototype situations, for the early part of the queue will behave like "bay entered, vehicles move up."

Again, note that shorter cycle lengths work in conjunction with a nonsignal remedy: short cycle lengths mean smaller platoons, which require shorter turn bays.

Figure J-41 presents a decision checklist to be used in considering a left-turn-based problem. The final decision must be evaluated with due consideration of the following:

1. Is an alternate route available for the left turners? How much does it adversely impact them? Can the alternate route afford to be impacted by the additional flow?
2. How many parking spaces must be removed to aid the flow, via a turn bay or even an additional lane? What

is their economic value?

3. What delay is being suffered now?

Frequently, the decisions can be reached by systematically examining the options as outlined in Figure J-41, keeping the foregoing issues in mind. Sometimes, a benefit-cost or cost-utility decision (beyond the scope of these guidelines) may be required for a "close decision" or highly sensitive issue.

Figure J-42 illustrates the impact of a turn prohibition at one site, and of adding a lane for left turners (i.e., a long turn bay). This is intended simply as an example of the increases in productivity cited previously. At the specific site in question, left-turn prohibition is not feasible, but provision of such a turn lane is.

Right-Turn-On-Red (RTOR)

It should be noted at the offset that RTOR can also apply to left-turn-on-red from a one-way street to a one-way street and is so used herein.

At the time these guidelines were written a major project was underway on RTOR, sponsored by the Federal Highway Administration (*J-12*). The engineer may wish to refer to that work for a treatment of RTOR and all its component issues. The following comments relate only to congestion effects.

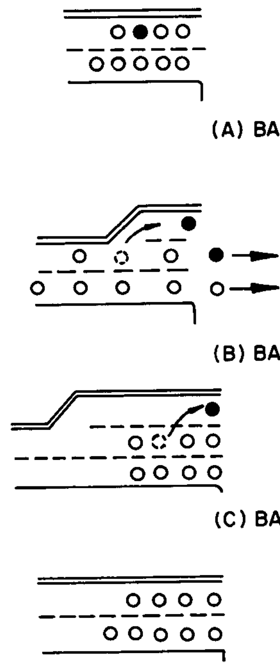
With respect to resolution of congestion and oversaturation, RTOR has the following advantages on the link on which it is allowed:

1. It reduces delay, with the most substantial savings being realized by those turning.
2. It increases throughput, with the percentage increase being of the same order as the percentage of right turners if an adequate turn bay exists. (The turn bay should be of the same order of length as the expected nonturning queue length (see Appendix VII of Ref. (J-6) for details, or earlier sections under "Turn Bays and Other Nonsignal Remedies"). Completely occupied turn bays, of course, would not be a component in the computations.)

A right-turn bay itself would provide the latter benefit, without RTOR. Likewise, it would provide much of the delay saving. Still, there is a substantial additional increment due to the use of RTOR with the turn bay.

A major disadvantage of RTOR in the control of congestion involves the fact that turning vehicles can enter heavily loaded links at will, "stealing" storage space from through vehicles on arterials and thus disrupting the intended control. In addition, the engineer must also be alert to the safety problems at locations of significant pedestrian activity.

Figures J-43 and J-44 illustrate the system impact of systemwide RTOR for the case study of Figure J-33. It is apparent from Figure J-43 that during the period represented, there appears to be some reduced delay per vehicle; the growth of delay per vehicle over time, as measured by the standard deviation of successive interval values, gives only a hint that there is a potential problem. The spillback history (Fig. J-44), however, clearly indicates a problem: RTOR creates oversaturation at the next upstream intersection.



$Q_A = \dfrac{N + (F-1)Np_T}{L}$

(A) BASE CONDITION: NO BAY

$Q_B = \dfrac{N}{L}$

(B) BAY ENTERED, GAP LEFT

$Q_C = \dfrac{N(1-p_T)}{L}$

(C) BAY ENTERED, VEHICLES MOVE UP

$Q_D = \dfrac{N}{L}$

(D) PROHIBITION

Figure J-40. Types of left-turn organization.

It should be recognized that this problem does not occur when the equity offset function provided by backward progression is in force; the queue extent is so long that the right turners who would "steal space" during the time provided by the RTOR just cannot fit during the added turning time given them. This is illustrated by Figure J-36, parts B and C (the added turning time is during Main Street green when the Main Street queue extent is maximum).

Two-Way Turn Lanes

Vehicles that make midblock left turns frequently create an impediment (a local point of congestion) for those who are continuing along the roadway. Two-way left-turn lanes (i.e., a lane that can be used by vehicles in either flow direction to execute their turns, as shown in Fig. J-45) can remedy such congestion impact, if space permits, and there is some evidence that they can substantially improve the accident situation (*J-13*). Should such turns occur during truly heavy flows, they can, indeed, virtually remove one moving lane for much of the peak period; this could happen in both directions simultaneously.

Two-way left-turn lanes also offer a remedy when confronted with suburban strip development patterns. However, one must conserve the directional capacity. Consider a situation in which there are two moving lanes in each direction, plus a center two-way left-turn lane. If the center is dedicated to one of the two directions, that direction has a net increase in capacity approximately equal to the percent time midblock left turners do not block the lane. Potentially, even more important, three discharge lanes are available at the intersection in that direction. The engineer must trade advantages off the

TABLE J-8

ESTIMATES OF PRODUCTIVITY INCREASE, LEFT-TURN BAY

P_T	FACTOR F = 2		FACTOR F = 3	
	Bay, Gap	Bay, No Gap	Bay, Gap	Bay, No Gap
0.05	4.7%	9.5%	9.1%	13.6%
0.10	9.1%	18.2%	16.7%	25.0%
0.15	13.0%	26.1%	23.1%	34.6%

Notes:

(1) P_T is left turn percentage

(2) "Bay, Gap" is shown in Figure J-49. B, "Bay, No Gap" in Figure J-49. C

(3) Effective Green Decrease: 100/(100 + Increased Percent)

potentially uneven quality of flow in the third lane, because of the left turners.

The option of a two-way turn lane not only may solve the midblock congestion problem, but it also may open the possibility, with the addition of another lane, that it can be used at the intersection by the through flow—if unbalanced flow is implemented. The increased discharge capacity per hour of green for that approach will allow a reduction in the total green needed for the approach (for a given demand) and, thereby, a decrease in the cycle length.

Of course, this use of unbalanced flow (also discussed later in the context of major land arrangement) requires proper planning—Can the opposing direction accommodate its own turners without unduly impeding its continuing vehicles? Figure J-46 shows a flow pattern that allows unbalanced assignment on lanes: 3 eastbound, and 2 westbound. The fact that a.m. midblock turners are light permits this. Had the morning westbound midblock turners also been moderate, the two-way left-turn lane may have been necessitated or a reversible lane implemented.

Figure J-47 summarizes the decision checklist when midblock congestion is involved, or when unbalanced flow and reversible lanes are alternate candidates. Note that any decision involving two-way turn lanes versus unbalanced flow in this context considers only the congestion/saturation issue. Data on accident advantages are not sufficient to say whether there is an accident benefit of two-way turn lanes (e.g., removal from the traffic stream as opposed to decreasing the density) which should override.

Dual Turn Lanes

When turning volumes are extremely heavy, both capacity and queue extent may dictate use of two lanes for a turning movement. Such a situation is illustrated in Figure J-48.

Data in the literature and discussions with engineers who implemented dual turning lanes indicate that these lanes have comparable discharge capacities and can be so treated in the signal timing. Of course, multiphase operation is logical if there is an opposing flow; standard markings and signal devices are needed and must be properly located.

MAJOR LANE ARRANGEMENT

If localized remedies fail, the engineer may have to consider systemwide remedies. The engineer should have already considered all of the candidate remedies previously discussed, including, especially, use of turn bays (which provide additional space).

One-Way System

One-way operation generally increases the capacity of any street, except perhaps for two-lane streets with curb parking, where side frictions apparently preclude a gain.

One-way systems have distinct advantages for congestion/saturation: turns are unimpeded, lanes can be assigned by movement easier (since there are now more in the same direction), midblock crossing conflicts do not occur, and there is more space to pass standing queues. One-way signalization can be achieved on two-way systems, but only by disregarding the other direction totally (this is, in fact, essential if signalization is to be used aggressively to combat congestion/saturation). Thus, the spatial removal—and separate signalization—of the two directions is a distinct advantage.

Of course, changing to a one-way system is a major project in itself, in which many of the issues are not directly related to congestion and saturation. Impact on business and on bus routes (walking distances, etc.) are just two issues to consider. Increased travel times or circulation for significant flows is another. Most importantly, there must be return paths for the flows during "the other" peak.

At any rate, if other remedies with the existing street layout fail, and reverse paths appear to exist, the engineer must then undertake a feasibility study of a one-way street system (or a partial system). Use of these guidelines will also provide insight into signalization, conflicts, and remedies on the one-way system.

Unbalanced Flow

It may well be that striping of two-way streets for an odd number of lanes, with one direction having more lanes than the other, will solve some problems. Such a remedy generally requires that the demand in two directions not be symmetric; if the engineer has *X* vph westbound in the a.m. peak hour, and the same number in the p.m. peak hour, what logic would ever permit one direction to have fewer lanes?

One possible situation would arise from the turning pattern of these vehicles; namely, the equivalent flows may not be equal.

Another possible situation may be that the problem relates to block lengths more than to throughput. Consider Figure J-49, which depicts a situation in which the

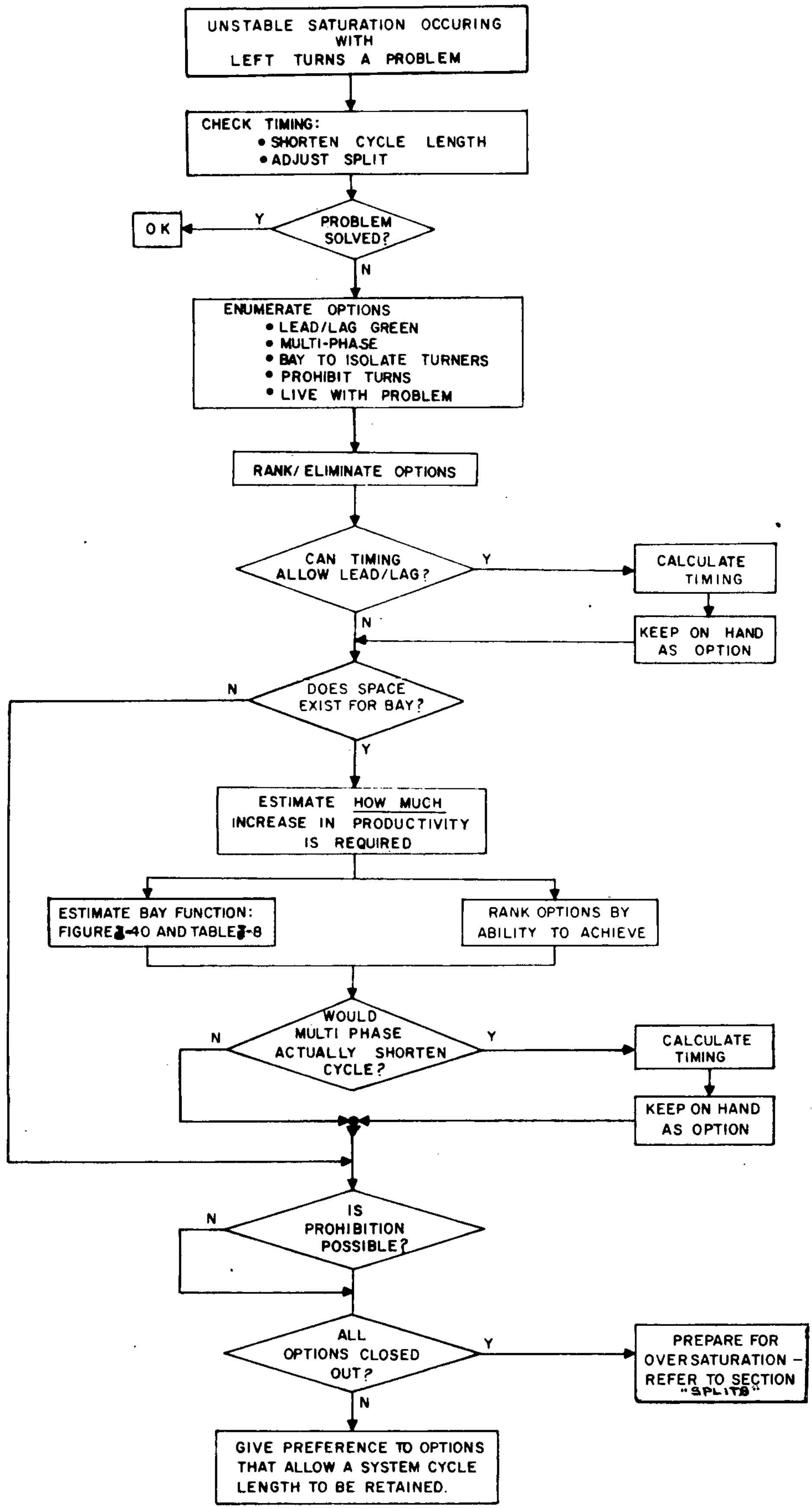


Figure J-41. Decision checklist for left-turn problems.

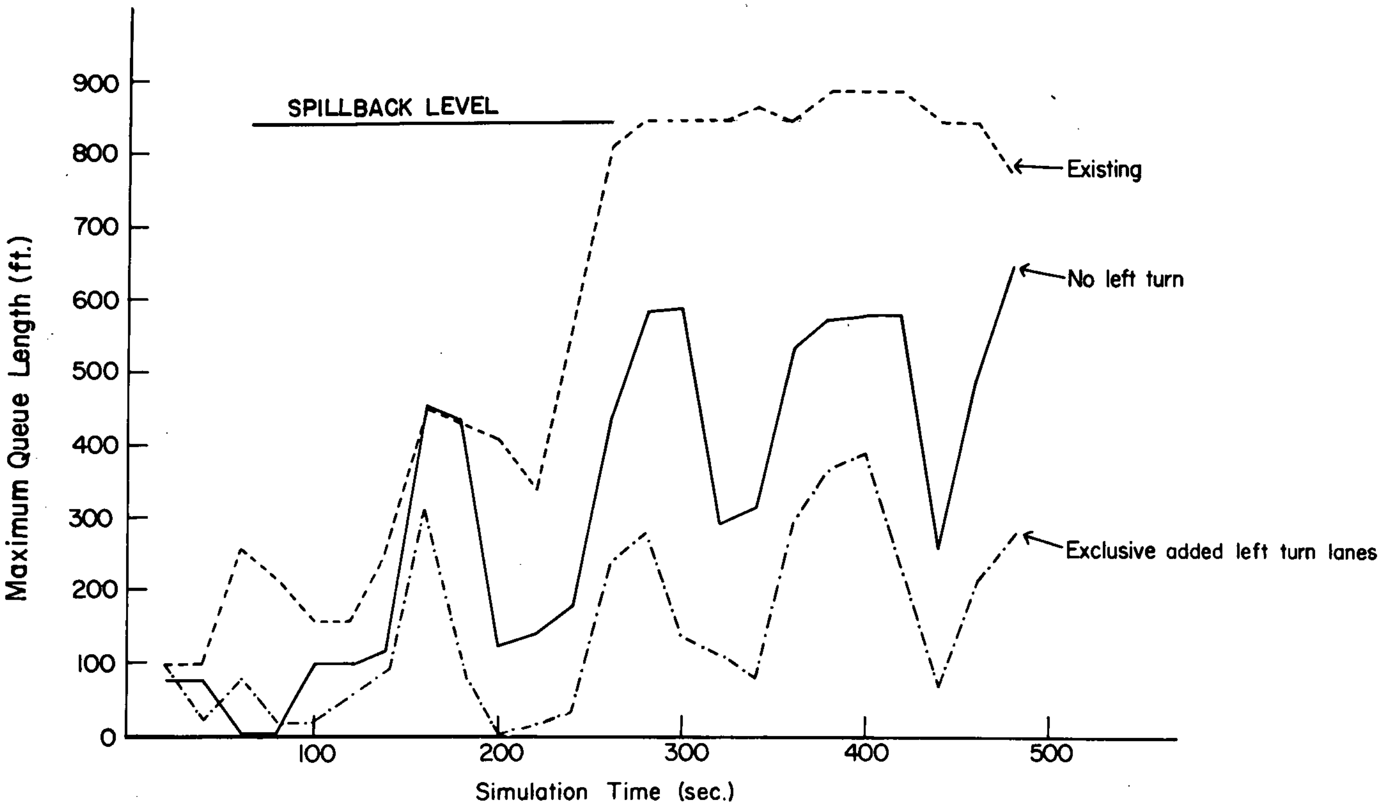


Figure J-42. Maximum queue length upstream of a typical CI with many left turns.

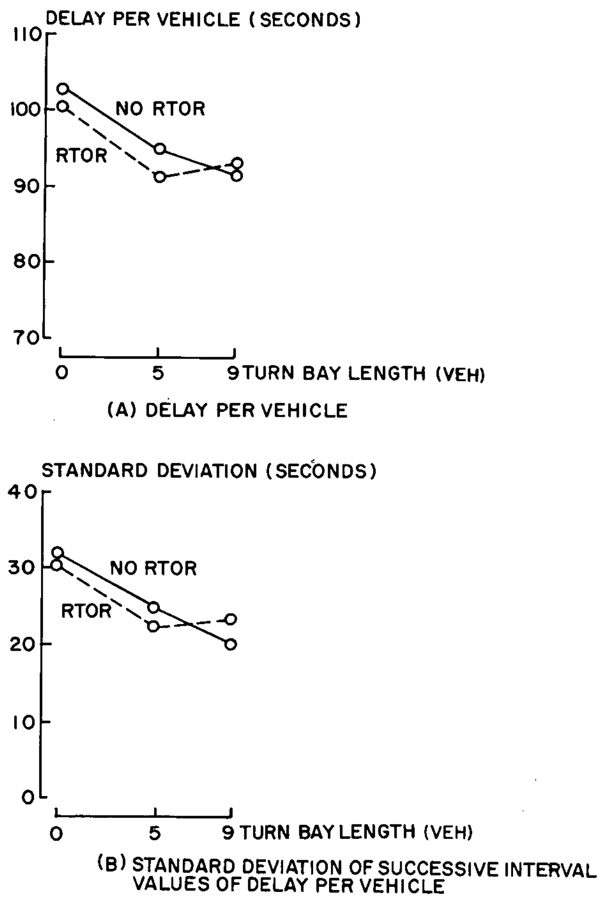


Figure J-43. System results with RTOR, simultaneous progression.

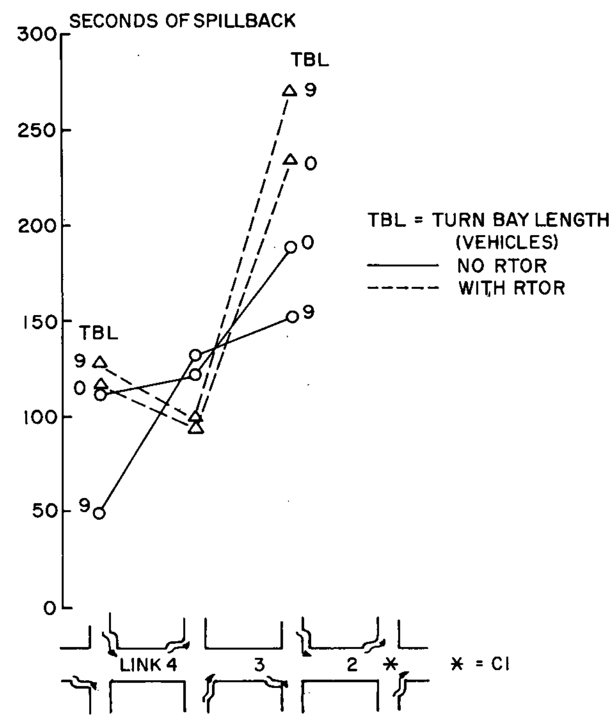


Figure J-44. Possible adverse impact with RTOR, simultaneous progression.

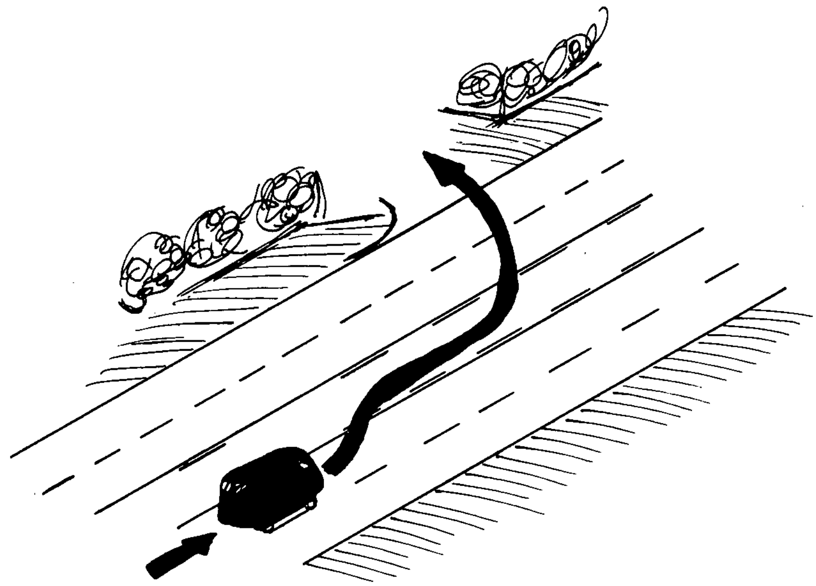


Figure J-45. Use and layout of a two-way left turn lane.

desire to avoid spillback may motivate unbalanced flow (width permitting, of course). Let there be equal flows in the eastbound (EB) and westbound (WB) directions, each capable of causing oversaturation. The engineer may wish to use both split and unbalanced assignment of lanes to attempt to equalize the time at which intersections F and G are affected. This may be particularly important when $D_2 \gg D_1$. Failing that, he may attempt to equalize the times at which intersections E and H are affected. The formulas given in the subsection on "Splits" apply.

In any case, the use of an odd number of lanes will be over several blocks. It is not a local treatment. Figure J-50 depicts a situation in which unbalanced lane arrangement is the only feasible solution, the 1375 sec of available green having been dictated by considerations at heavily loaded intersections. Although not shown explicitly, it must be ascertained that flows in other periods can be accommodated. The heavy right turn into Avenue C from Main Street occupies virtually one lane. A lane is therefore marked for right turners.

Unbalanced lane arrangement was also discussed earlier in the context of two-way turn lanes. Figure J-47 summarizes some of this discussion.

Reversible Lanes

If all other potential remedies cannot be applied, the engineer may wish to consider reversible lanes; that is, one or more lanes which are assigned to one direction or the other, depending on the time of day. This remedy requires substantial associated signing, marking, and lane signalization. It does present an added (and, for many, an unfamiliar) task for the motorist: cognizance of such a situation and avoidance of the lane when appropriate. It has the advantage of flexibility, and certainly provides additional capacity where needed.

DISRUPTIONS TO THE TRAFFIC

Disruptions to the traffic stream can occur from numerous sources; pedestrian interferences, bus boarding, parking/unparking maneuvers, and midblock activities are perhaps the most common sources.

This section illustrates the impact of some of these disruptions; some were addressed earlier.

Pedestrian Interference

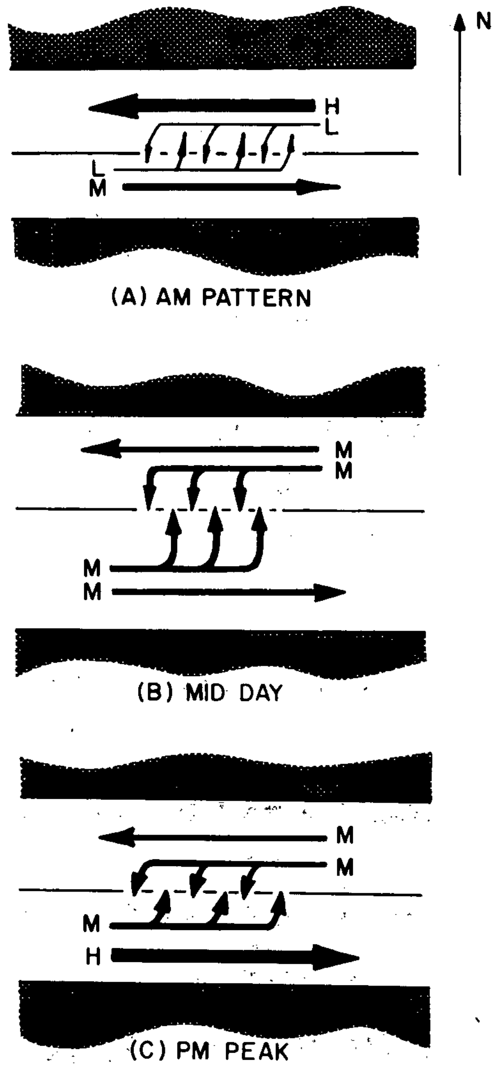
There is no question that pedestrian interference not only can degrade traffic performance but also can create safety problems. Taken as a traffic entity in its own right, the pedestrian flow creates a challenging design and operations problem. Explicit design for pedestrians is addressed in the literature (*J-14, J-15*).

In designing for the pedestrian movement, three characteristics must be considered: (1) minimum pedestrian green time, (2) depth of the storage space (reservoir), and (3) minimum space allocation in the crosswalk. These and other aspects, including sidewalk design, are addressed in the literature cited.

For the benefit of both pedestrian and traffic movement, it is best to separate the movements. Unfortunately, grade separation is rarely feasible within the context in which these guidelines are used. Prohibition of pedestrian crossing on certain legs is also frequently impractical.

One-way streets aid the solution of pedestrian-vehicular problems by reducing the number of conflict locations per intersection, by widening the approach so as to enable through vehicles to avoid standing turning queues, and by giving the turners a broader turning front. The reduction of the number of legs with conflicts makes changes of path more attractive for some pedestrians. Figure J-51 illustrates these aspects.

Another solution to the pedestrian-vehicular problem involves turn bays, which provide an opportunity to re-



LEGEND:

H HEAVY

M MODERATE

L LIGHT

NEEDS: CONTINUING +	TURN →	LANES
H	L	2
H	M	3
M	M	2
M	L	<2

Figure J-46. Possible midblock demand situation permiting unbalanced flow.

move vehicles from the traffic stream when significant turning delays are noted or anticipated. A separate pedestrian phase generally lengthens the cycle considerably, and is only truly effective in localities that have strict pedestrian compliance.

In terms of actual delay and productivity decrease, pedestrian activity has to be rather heavy—and combined with heavy volumes (virtually capacity)—before any harm is done at normal turning rates. This does not mean that the quality of service provided to both motorist and pedestrian remains the same as at lower volumes; it simply means that the productivity does not change.

Figure J-52 depicts the impact of pedestrian interference. Note that numbers 100 percent, 90 percent, and 80 percent refer to percent of capacity on a two-lane link. Turns were the same percentage in each case (about 6 percent) giving rise to turn volumes of 100 vph, 90 vph, and 80 vph, respectively The pedestrian rates inherent in the UTCS-1 simulator (which was used to generate these curves) are representative of those in Washington, D. C. The heaviest level (4) is said to be typical and representative of intense CBD pedestrian traffic. Note also that substantial increases in travel time or delay/vehicle only occurs for the highest pedestrian volumes and the highest V/C ratios. Part C of the same Figure indicates that, in the worst case, approximately a 100 percent increase in the average number of stops has occurred. This reflects on quality rather than on productivity. Indeed, over the range illustrated, productivity did not change as a function of pedestrian volume.

Bus Stops

Buses frequently stop in the traffic stream. A basic question arises as to whether this adversely affects productivity or delay in a significant way. Also, if there are "adverse uses" (e.g., goods delivery, passenger drop-off, etc.), are they of sufficient benefit to the community to justify any such disbenefits?

This latter question touches on site-specific benefit-cost or cost-effectiveness analysis beyond the scope of these guidelines. Suffice it to say that simulation studies run in the course of this work ascertained productivity increases only in the order of 2 to 3 percent under heavy flow conditions. (In other words, discharge increased by only 2 to 3 percent when buses used bus stops.)

Figure J-53 illustrates the individual disruption caused by buses stopping in the traffic stream. Certainly such events, particularly at farside bus stops, cause local disruptions and even spillback.

In the following, the impact of bus stops—their existence or not (which can be interpreted as their use or not because of the simulation runs made)—is investigated in a number of cases. Because the cases studied comprise a range of private auto volumes, some interpretation as to the impact of moving people from autos to buses is possible. While this transit planning is generally outside the focus of these guidelines, some pitfalls of such an interpretation are worth noting.

Consider the two-way street of Figure J-54 with a base condition as follows:

- 600 passenger cars per hour per lane (pcphpl).
- 20 buses/hr.
- Two moving lanes in each direction.
- Simple progression, set at 25 mph.
- Bus occupancy, 50 persons; auto occupancy, 1.5 persons.

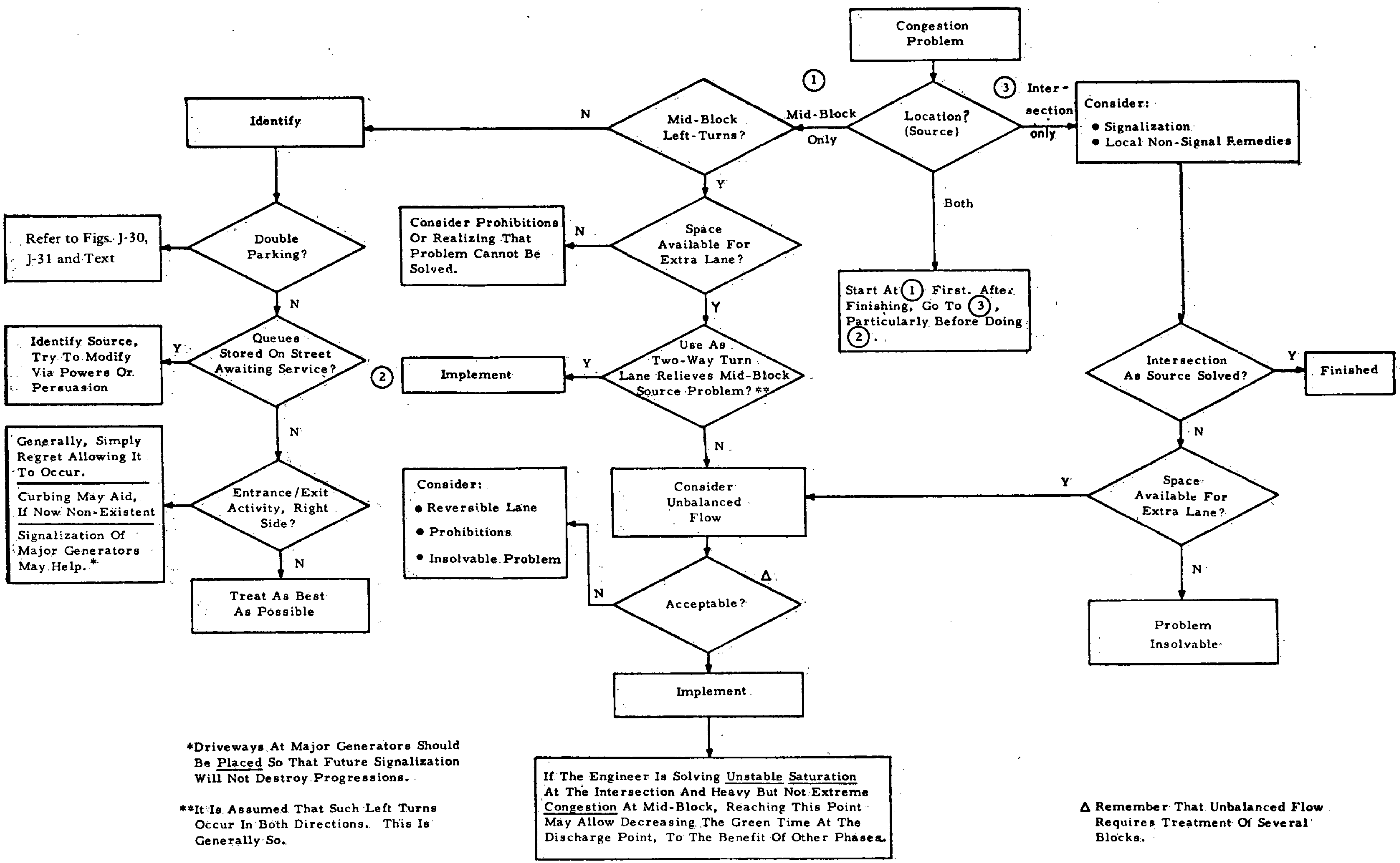


Figure J-47. Decision checklist involving two-way turn lanes and unbalanced flow.

Figure J-48. Two lanes for left turns.

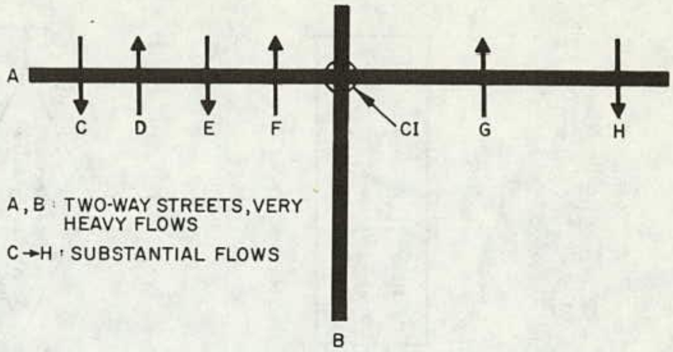


Figure J-49. Block length may influence unbalanced flow.

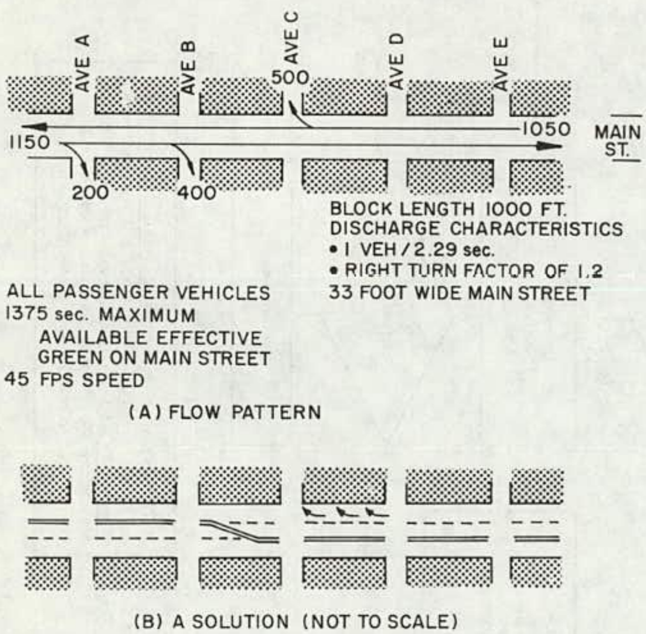
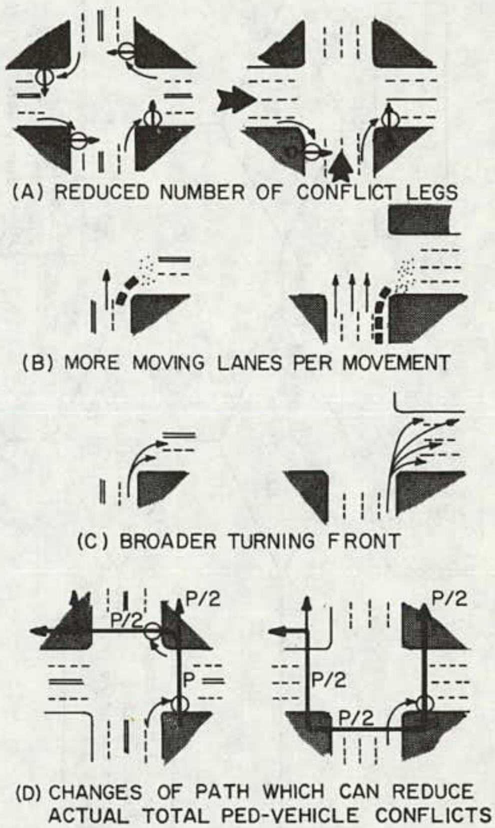


Figure J-50. Case in which flow pattern dictates unbalanced lanes.



Notes:
1. ⊖ denotes a ped-vehicle conflict point
2. P is a pedestrian volume; in some cases, it has alternate paths; a 50 - 50 division is then assumed for illustration

Figure J-51. Advantages of one-way over two-way operation—pedestrian-vehicular activity.

From this base, a number of volumes are considered. For convenience, the number of people is kept constant by increasing the number of buses. The buses needed to do this is shown in part A of Figures J-55 and J-56.

Figure J-55 shows that the use of bays has the most impact at moderate volumes on the street studied. This makes sense: at low volumes, the buses stopped in a moving lane can be easily by-passed; at high volumes, the passenger cars themselves are so slow moving that the buses do not contribute a substantial increment of delay.

Figure J-56 shows that the use of a bus bay has much less of an impact when the street has three moving lanes in each direction, even at moderate volumes. This would appear to be due to the greater opportunity to by-pass stopped buses.

An attempt to evaluate the benefits of moving people onto transit (thus decreasing person-delay) would reveal the following limitations of these case studies. Everyone on the buses is still required to stop at all bus stops; the signal offsets were not tailored to the buses, despite (in some cases) their greater numbers of people; the use

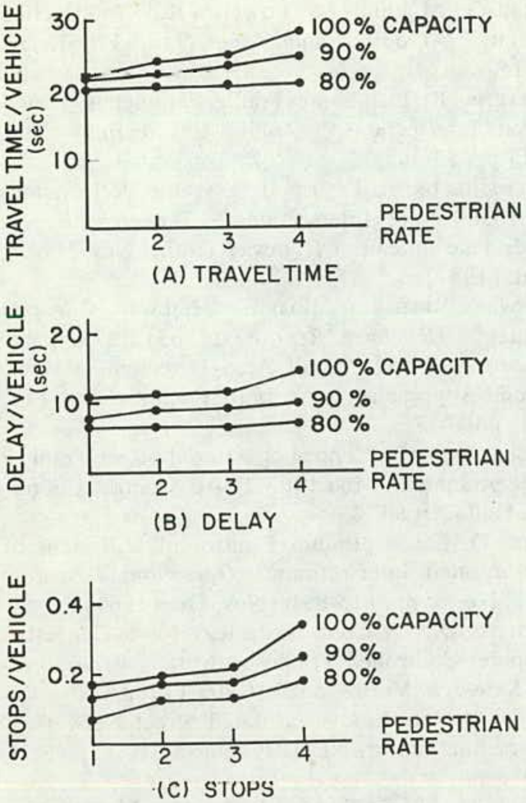
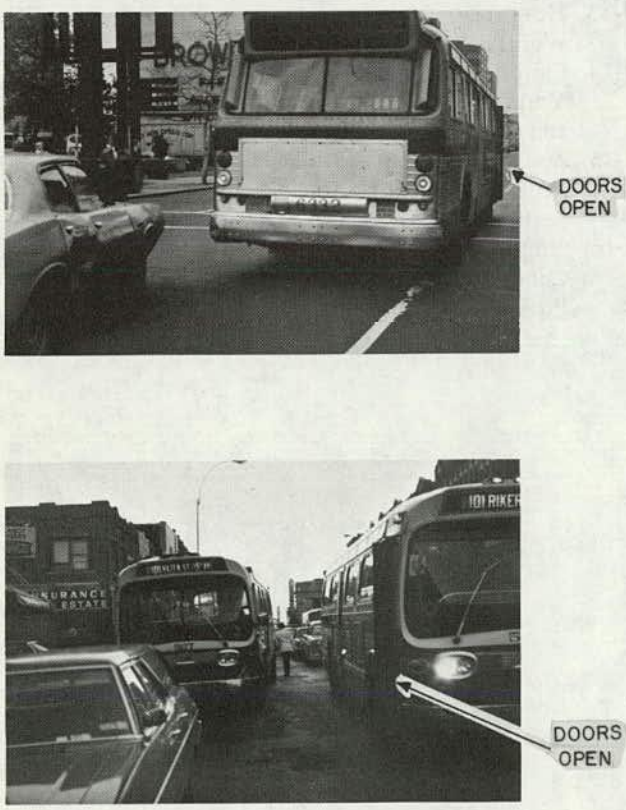


Figure J-52. A case of pedestrian interference on a two-lane link with right turns.



BUS IN TRAFFIC STREAM, HEAVY FLOW

Figure J-53. Disruptions due to buses.

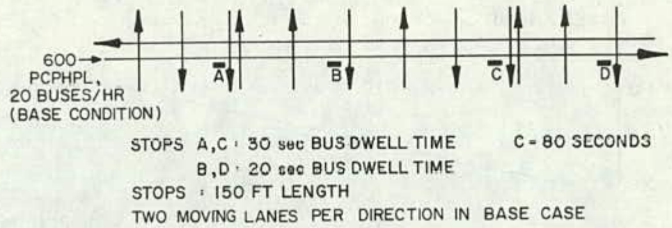


Figure J-54. A two-way arterial with bus stops.

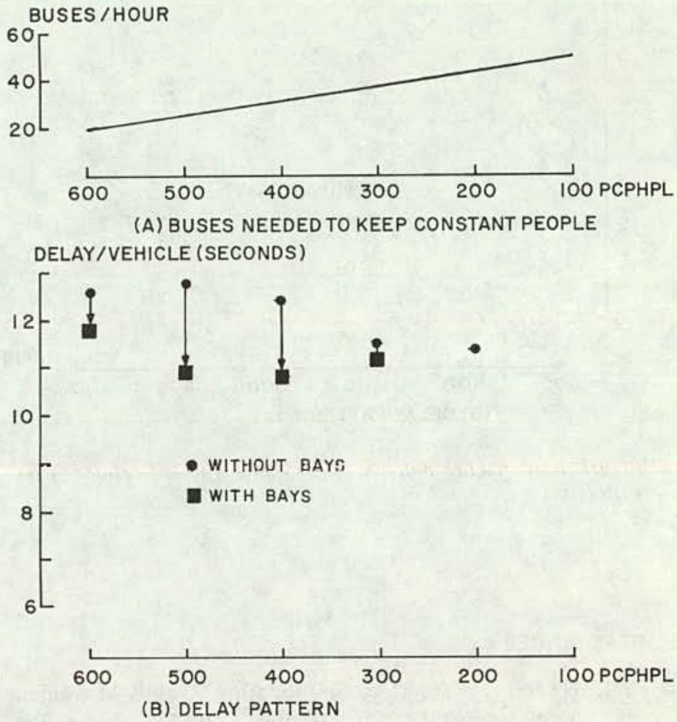


Figure J-55. Delay pattern on a 2-lane (per direction) street with buses.

of some stops (particularly in the 3-lane case) becomes intense. Should the engineer become involved in such planning or analysis, he is cautioned to use suitable alternatives (nonstopping buses, perhaps, to substitute for through vehicles, unless a full route is evaluated) and proper signalization. In the present context, the engineer should restrict his attention to the impact of the use (or not) of bus bays in Figures J-55 and J-56 as a traffic stream disruptor.

Parking/Unparking/Double Parking/Midblock Activity

Intense parking/unparking can create short-term disruptions of a local nature. While perceived as a problem, it is often the accompanying double-parking or other midblock activity that is the true problem (refer to appropriate sections discussed earlier—subsection "Issues and Considerations" under section on "Nonsignal Controls—Enforcement and Prohibition," and subsection "Two-way Turn Lanes" under section on "Turn Bays and Other Nonsignal Remedies").

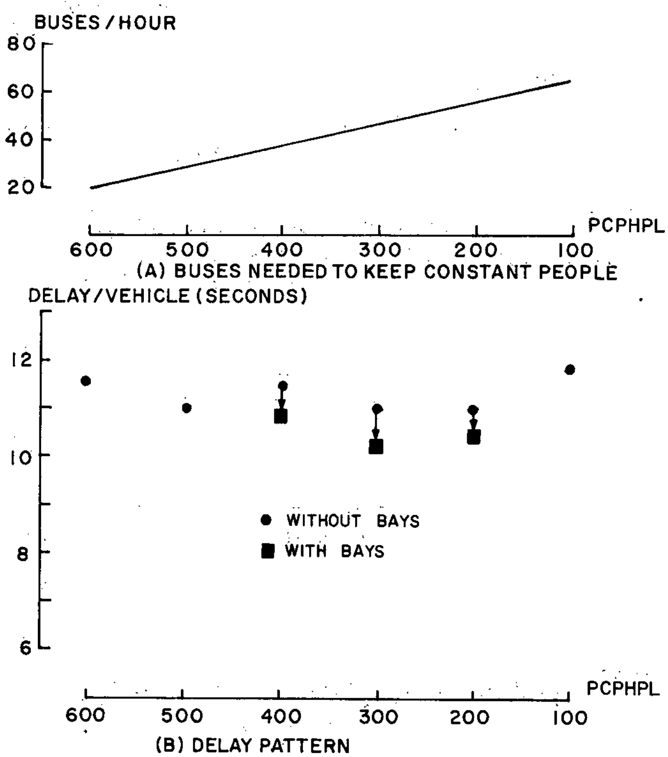


Figure J-56. Delay pattern on a 3-lane (per direction) street with buses.

REFERENCES

J-1. HABIB, P. A., "Accommodating Goods-Movement Vehicles in the City Center." Ph.D. dissertation, Polytechnic Institute of New York (June 1975).

J-2. MOSKOWITZ, K., "Long Cycles, Short Cycles, and "Lost Time" at Signalized Intersections." Paper presented at the Workshop on Intersection Capacity, Highway Research Board (Jan. 1974).

J-3. WEBB, G. M., and MOSKOWITZ, K., "Intersection Capacity." *Traffic Engineering* (Jan. 1956), pp. 147-152.

J-4. CARSTENS, R. L., "Some Traffic Parameters at Signalized Intersections." *Traffic Engineering* (Aug. 1971) pp. 33-36.

J-5. GREENSHIELDS, B. D., ET AL., "Traffic Performance at Urban Street Intersections." *Technical Report No. 1*, Yale Bureau of Highway Traffic, New Haven, Conn. (1947).

J-6. HIGHWAY RESEARCH BOARD, "Highway Capacity Manual." *HRB Spec. Rept. 87* (1965) 397 pp.

J-7. PARSONSON, P. S., "Small Area Detection at Intersection Approaches." *Traffic Engineering* (Feb. 1974) pp. 8-17.

J-8. YAGODA, H., "The Control of Arterial Street Traffic." Paper presented at the 1967 TFAC Control Conference, Haifa, Israel.

J-9. GAZIS, D. C., "Optimum Control of a System of Oversaturated Intersections." *Operations Research*, V. 12, No. 6, pp. 815-831 (Nov.-Dec. 1964).

J-10. LONGLEY, D., "A Control Strategy for a Congested Computer-Controlled Traffic Network." *Transportation Research*, Vol. 2, pp. 391-408 (1968).

J-11. "Traffic Control in Oversaturated Street Networks." Agency final report, NCHRP Project 3-18(2) (Sept. 1973).

J-12. MCGEE, H., "Right Turn on Red," Alan M. Voorhees and Associates, Inc., Highway Administration, Contract CN-DOT-FH-11-8251.

J-13. HOFFMAN, M. R., "Two-way Left-Turn Lanes Work!" *Traffic Engineering* (Aug. 1974) pp. 24-27.

J-14. FRUIN, J. J., *Pedestrian Planning and Design*. New York, Metropolitan Association of Urban Designers and Environmental Planners (1971).

J-15. ZUPAN, J. M., and PUSHKAREV, B., "Capacity of Pedestrian Facilities." Paper presented at the Transportation Research Board, Wash., D.C. (Jan. 1975).

J-16. "Improved Control Logic for Use with Computer-Controlled Traffic." Agency final report, NCHRP Project 3-18(1).

THE TRANSPORTATION RESEARCH BOARD is an agency of the National Research Council, which serves the National Academy of Sciences and the National Academy of Engineering. The Board's purpose is to stimulate research concerning the nature and performance of transportation systems, to disseminate information that the research produces, and to encourage the application of appropriate research findings. The Board's program is carried out by more than 150 committees and task forces composed of more than 1,800 administrators, engineers, social scientists, and educators who serve without compensation. The program is supported by state transportation and highway departments, the U.S. Department of Transportation, and other organizations interested in the development of transportation.

The Transportation Research Board operates within the Commission on Sociotechnical Systems of the National Research Council. The Council was organized in 1916 at the request of President Woodrow Wilson as an agency of the National Academy of Sciences to enable the broad community of scientists and engineers to associate their efforts with those of the Academy membership. Members of the Council are appointed by the president of the Academy and are drawn from academic, industrial, and governmental organizations throughout the United States.

The National Academy of Sciences was established by a congressional act of incorporation signed by President Abraham Lincoln on March 3, 1863, to further science and its use for the general welfare by bringing together the most qualified individuals to deal with scientific and technological problems of broad significance. It is a private, honorary organization of more than 1,000 scientists elected on the basis of outstanding contributions to knowledge and is supported by private and public funds. Under the terms of its congressional charter, the Academy is called upon to act as an official—yet independent—advisor to the federal government in any matter of science and technology, although it is not a government agency and its activities are not limited to those on behalf of the government.

To share in the tasks of furthering science and engineering and of advising the federal government, the National Academy of Engineering was established on December 5, 1964, under the authority of the act of incorporation of the National Academy of Sciences. Its advisory activities are closely coordinated with those of the National Academy of Sciences, but it is independent and autonomous in its organization and election of members.

TRANSPORTATION RESEARCH BOARD
National Research Council
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

ADDRESS CORRECTION REQUESTED

NON-PROFIT ORG.
U.S. POSTAGE
PAID
WASHINGTON, D.C.
PERMIT NO. 42970

000015M001
JAMES W HILL

IDAHO TRANS DEPT DIV OF HWYS
P O BOX 7129
BOISE ID 83707