1. Report No. TX-97/2929-2	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle RESULTS OF SIMULATION STUDIES RELATING TO THE OPERATION OF CLOSED-LOOP SYSTEMS IN A TRAFFIC RESPONSIVE MODE		5. Report Date September 1996 Revised: January 1997
		6. Performing Organization Code
7. Author(s) Kevin N. Balke, Keithireddipalli S. Rao, and Christopher L. Brehmer		8. Performing Organization Report No. Research Report 2929-2
9. Performing Organization Name and Address Texas Transportation Institute The Texas A&M University System College Station, Texas 77843-3135		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. Study No. 7-2929
12. Sponsoring Agency Name and Address Texas Department of Transportation Research and Technology Transfer Office P. O. Box 5080 Austin, Texas 78763-5080		13. Type of Report and Period Covered Interim : September 1994-August 1996
		14. Sponsoring Agency Code

15. Supplementary Notes
Research performed in cooperation with the Texas Department of Transportation.
Research Study Title: Guidelines for Implementing Traffic Responsive Mode in TxDOT's Computerized Traffic Signal Systems

16. Abstract
This report summarizes the results of simulation studies performed in the development of guidelines for implementing traffic responsive mode in closed-loop traffic signal systems. Simulation studies were performed to evaluate a revised interconnection desirability index which can be used to determine whether two traffic signals should be operated in a coordinated or an isolated mode. Additional simulation studies were performed to evaluate three different configurations of system detectors to support the use of traffic responsive mode in a closed-loop system. Another simulation study was performed to evaluate the benefits of operating a signal system in a traffic responsive mode as compared to a time-of-day mode.

17. Key Words Closed-Loop Traffic Signal Systems, Traffic Responsive, System Detectors, Coordination	18. Distribution Statement No restrictions. This document is available to the public through NTIS: National Technical Information Service 5285 Port Royal Road Springfield, Virginia 22161		
19. Security Classif.(of this report) Unclassified	20. Security Classif.(of this page) Unclassified	21. No. of Pages 92	22. Price

Form DOT F 1700.7 (8-72)   Reproduction of completed page authorized

RESULTS OF SIMULATION STUDIES RELATING TO THE OPERATION OF CLOSED-LOOP SYSTEMS IN A TRAFFIC RESPONSIVE MODE

by

Kevin N. Balke, P.E.
Assistant Research Engineer
Texas Transportation Institute

Keithireddipalli S. Rao
Assistant Research Scientist
Texas Transportation Institute

and

Christopher L. Brehmer
Graduate Research Assistant
Texas Transportation Institute

September 1996
Revised: January 1997

IMPLEMENTATION STATEMENT

This report summarizes the results of several simulation studies related to operating a closed-loop traffic signal system in a traffic responsive mode. The simulation studies focused on three primary areas:

- the evaluation of a proposed procedure for determining when to operate traffic signals in an isolated versus a coordinated mode,

- the placement of system detectors to support the operation of closed-loop signal systems in a traffic responsive mode, and

- the benefits of operating a closed-loop traffic signal system in a traffic responsive and in a time-of-day mode.

The results of these studies were used to develop the following:

- procedures for determining where to provide isolated or coordinated control between two traffic signals,

- guidelines for locating system detectors in a closed-loop signal system to support operating the system in a traffic responsive mode, and

- guidelines for determining when to operate a traffic signal system in a traffic responsive versus a time-of-day control.

The guidelines and procedures are summarized in Report 2929-3F.

DISCLAIMER

The content of this report reflects the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official view or policies of the Texas Department of Transportation (TxDOT). This report does not constitute a standard, specification, or regulation, nor is it intended for construction, bidding, or permit purposes. The engineer in charge of the project was Kevin N. Balke, P.E. # 66529.

ACKNOWLEDGMENT

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SUMMARY

Most closed-loop systems today are capable of operating in a traffic responsive mode; however, many systems still are left to operate in a time-of-day mode after installation. TxDOT believes that one reason why traffic responsive mode is not used more often is that there are no clear guidelines or procedures to help engineers set up a closed-loop system in a traffic responsive mode. Specifically, guidelines are needed to address the following issues:

- when to operate a group of traffic signals in an isolated versus a coordinated mode,

- where to locate system detectors to support traffic responsive control, and

- when to operate a closed-loop system in a time-of-day versus a traffic responsive mode.

This report summarizes the results of several simulation studies that were performed to support the development of guidelines and procedures for setting up a closed-loop traffic signal system to operate in a traffic responsive mode.

A simulation study was performed to address the issue of when to provide isolated versus coordinated control in a group of signal-controlled intersections. A review of the literature revealed that several procedures have been developed to determine when to provide coordination between two signals. Of these procedures, the Interconnection Desirability Index was selected for further evaluation because it was the only procedure that expressly quantified the two main conditions that impact the ability to provide progression: the amount of platoon dispersion that occurs between two intersections, and the amount of vehicles in the traffic stream traveling in platoons. Simulation studies showed, however, that the Interconnection Desirability Index was not a good indicator for choosing between isolated and coordinated control when mid-block volumes were high. As a result, a revised Interconnection Desirability Index was developed. Additional simulation studies were then performed to determine thresholds defining when to provide isolated and coordinated control. These studies found that when two intersections had a score of 0.4 or greater with the revised Interconnection Desirability Index, the system performed better under coordinated control. Likewise, the studies found that when two intersections had a score of 0.3 or lower, the signals produced less delay under isolated control. Between 0.3 and 0.4, system performance was approximately the same under isolated control as it was under coordinated control.

Simulation studies were also performed to examine where to locate system detectors to support traffic responsive mode in a closed-loop signal system. Three different configurations of system detectors were examined: one where system detectors were provided at the critical intersection only; another where system detectors were located on all approaches to the critical intersection and at the mid-point of the system; and a final configuration where, in addition to

system detectors at the critical intersection and the mid-point of the arterial, system detectors were placed at the downstream entry point of the system. Traffic conditions in the network were simulated with each detector configuration. Volume and occupancy measures from the detector configurations were used to provide input into the traffic responsive algorithm. The simulation studies showed that there was no significant difference in the amount of total system delay that was provided by the signal system operating in a traffic responsive mode under the different detector configurations. It was concluded, therefore, that as long as the system is small, system detectors placed only at the critical intersection should be sufficient to operate the system in a traffic responsive mode.

A final simulation study was performed to examine the issues of when to operate a closed-loop signal system in a traffic responsive versus a time-of-day mode. This simulation study showed that there was no statistically significant difference in the amount of total system delay when the traffic signal system was operated in a traffic responsive mode versus a time-of-day mode. The primary reasons provided for why no significant difference could be found between the two operating modes include the following:

- any reduction in the amount of total system delay that was achieved by operating the signals in a traffic responsive mode was offset by the delays associated with the transitions between timing plans, and

- since previous research has shown that delays are relatively insensitive to changes in cycle lengths, the cycle lengths used to operate the system in a traffic responsive mode were too close to permit significant delay savings.

Therefore, it was recommended that when setting-up a traffic responsive system, engineers consider using timing plans that have large differences in cycle lengths.

CHAPTER I.
INTRODUCTION

Unlike at an isolated intersection, the objective of timing traffic signals on an arterial street is to provide for the progressive flow of traffic through the system. This objective recognizes that an individual traffic signal releases a platoon of traffic that travels to the next intersection on the arterial. Accomplishing this objective requires that timings of the intersections on the arterial are coordinated to ensure that the traffic signal at the downstream intersection turns green as or before the platoon arrives. This requires that the traffic signals be operated as a system.

There are several different types of control systems that can be used to provide coordinated control on an arterial. The Texas Department of Transportation (TxDOT) is installing closed-loop traffic signal systems in many locations across the state. These systems can select different timing plans in response to measured traffic conditions. Unfortunately, the traffic responsive capabilities of many of these systems are not being fully utilized. As a result, most of these systems are left to operate in a time-of-day mode. TxDOT has hypothesized that one reason many of these systems are not operated in a traffic responsive mode is the lack of guidelines that indicate when and where traffic responsive systems are beneficial; however, before guidelines can be formulated, there are several critical questions that must be addressed. These include the following:

- How does a traffic engineer know when to operate a group of signals in an isolated or a coordinated mode?

- Where does a traffic engineer need to locate the system detectors that are used in a control system so that changes in traffic patterns might be detected quickly and accurately?

- When is it best to operate the signals in a system in a traffic responsive mode versus a time-of-day mode?

This report summarizes the results of simulation studies performed to aid in the development of guidelines for implementing traffic responsive mode in closed-loop traffic signal systems. Simulation studies were performed to evaluate a revised interconnection desirability index which can be used to determine whether two traffic signals should be operated in a coordinated or an isolated mode. Additional simulation studies were performed to evaluate three different configurations of system detectors to support the use of traffic responsive mode in closed-loop systems. Another simulation study was performed to evaluate the benefits of operating a signal system in a traffic responsive mode as compared to a time-of-day mode.

1

CHAPTER II.
COORDINATED VERSUS ISOLATED CONTROL

Some of the greatest benefits to be achieved in traffic signal control come from coordinating the operations of two or more traffic signals. The objective of providing coordination between traffic signals is to minimize the number of stops and delays experienced by traffic traveling in a particular direction on an arterial. However, there is a trade-off associated with providing coordination. Even with the best coordinated traffic signal systems, some vehicles (particularly those on the cross-streets) may experience slightly longer delays under coordinated operations than if the signals were operated in an isolated (or local) mode. Because of this, coordinated control should be provided only when there are system-wide benefits to be achieved. Therefore, it is critical to evaluate the benefits that can be achieved when considering whether or not to include an approach or intersection in a coordinated system.

BENEFITS OF PROVIDING COORDINATION

The purpose of providing coordination between two traffic signals is to facilitate the progressive flow of traffic. This is done by ensuring that a green indication at the downstream approach is provided in sufficient time to permit vehicles to travel through the intersection without stopping. Coordination can be either one-way (where traffic flow in one direction is favored over all other directions) or two-way (where traffic flow in two opposite directions is favored). While the primary benefits of providing progression are to minimize the number of stops and delays in a particular direction of travel in the corridor, there are other benefits associated with providing coordination between traffic signals:

- the conservation of energy by minimizing fuel consumption,

- the preservation of the environment by reducing air pollution,

- the maintenance of a preferred travel speed in a direction of flow,

- the promotion of smooth flow by a platoon of vehicles, and

- the prevention of queues from exceeding available storage capacity at specific turn bays and approaches.

3

Factors Affecting Progression

In order to achieve the full benefits of coordination, three conditions must be present on a roadway. First, a predominant movement must exist between the two intersections. In other words, one movement at the upstream intersection (either the through, left turn, or right turn movement) must have significantly more traffic than the other movements. This promotes the formation of a natural platoon at the downstream intersection. When there is not a predominant movement at the upstream intersection, arrival patterns at the downstream intersection tend to be uniform.

In addition to one movement being predominant at the upstream intersection, traffic patterns have to repeat every cycle. This means that the same arrival patterns must be exhibited consistently from cycle to cycle. Because phase patterns and offsets usually remain constant from one cycle to the next, it is difficult to provide good coordination if the predominant movement varies from cycle to cycle.

Finally, the physical conditions of the roadway and the traffic demands at the intersection must support progression on the roadway. For example, traffic patterns at the two intersections must be similar enough so that the two intersections can operate with the same cycle length. Furthermore, the intersections must be located so that effects of progression in one direction do not negate the effects of progression in the other direction. Other factors that affect the ability to provide good coordination between intersections include the following:

- inadequate roadway capacity,

- substantial side friction (such as parking and multiple driveways),

- complicated intersections that require multi phase control,

- wide variability in traffic speeds (like those caused by heavy truck traffic),

- very short signal spacings, and

- heavy turning volumes either into or out of the street.

All of these factors can cause platoons to disperse more rapidly than in situations where these factors are not present.

4

EXISTING EVALUATION TOOLS

A review of the literature reveals that three different evaluation tools have been developed for determining when to provide coordination between two intersections. Each of these methods are discussed below.

Cost Function

A "cost" or "penalty" function has been proposed by McShane and Roess (*1*) for evaluating when to provide coordination between two signals. As shown in the equation below, this cost function is a weighted combination of stops and delays.

$$Cost = A \times (total\ stops) + B \times (total\ delay)$$

The engineer sets the weighting factors A and B to reflect the estimated economic cost of each stop and delay. The amount that each timing plan reduces the cost of control in a corridor is used in a cost-benefit analysis to evaluate whether or not coordinated control should be provided. Additional terms can be added to the equation to account for other factors that may affect the decision of whether or not to provide coordination (such as fuel consumption, vehicle emissions, etc.).

Coupling Index

Yagoda, et al. (*2*) developed a coupling index to determine which links in a network should be grouped together in a coordinated system. The index is the ratio of the volume of traffic on a link to the distance between two intersections:

$$I = \frac{V}{L}$$

where
> I = coupling index,
> V = hourly approach link volume (vph), and
> L = link length to next signal (meters).

To determine which links should be coordinated in a system, the coupling index is computed for each link in the system. As shown in Figure II-1, links with low index values are selectively removed from the potential control area until the network degenerates into smaller, more manageable subareas. The threshold for retaining links using the coupling index is set to meet local conditions and requirements. For example, the City of Arlington, Texas, uses two
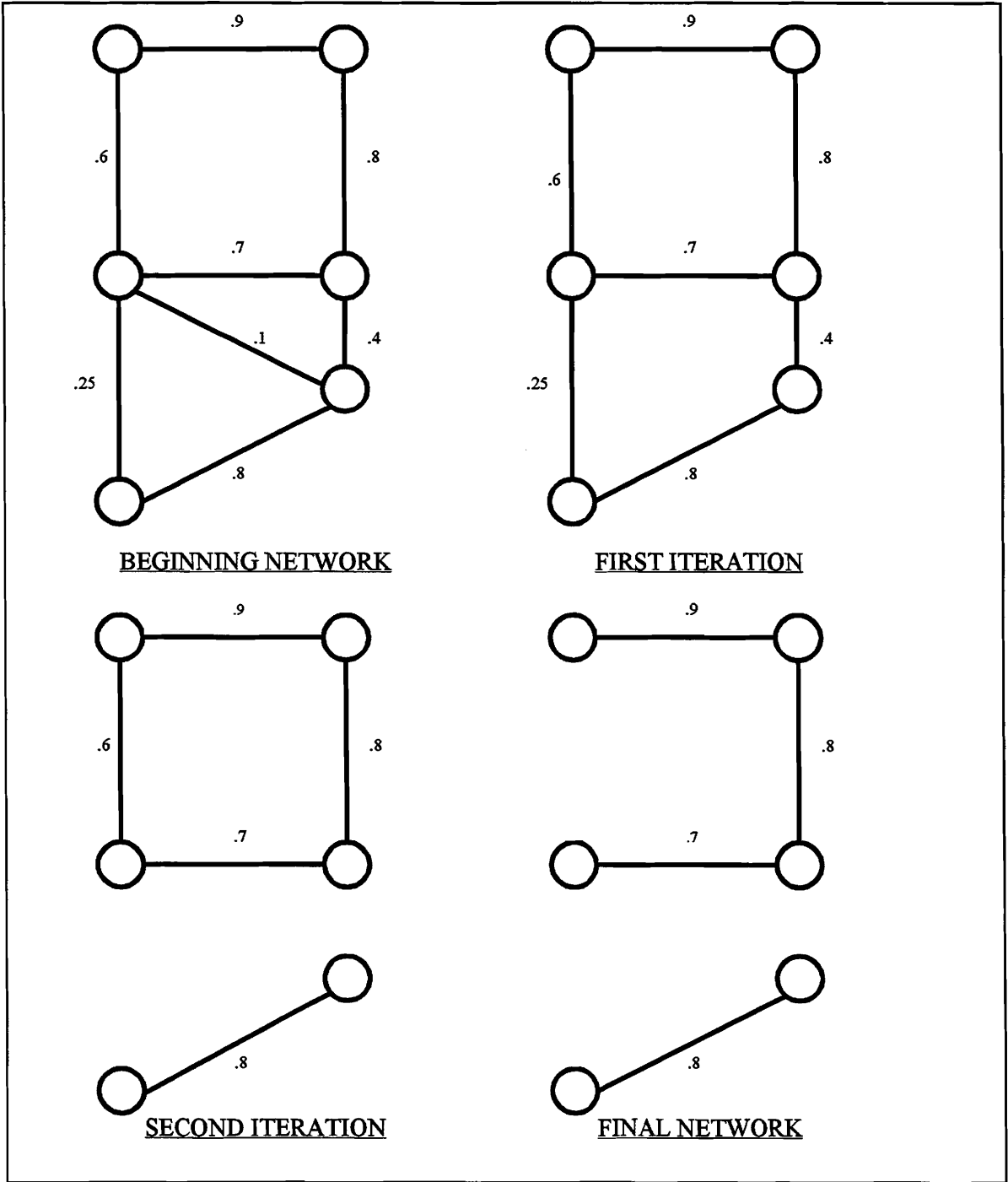
FIGURE II-1. Example of Subdivision Process Using Coupling Index (2)

different thresholds: a coupling index of 0.98 or more during any hour is used for planning purposes while an index value of 1.64 or more is used in operational analyses (3).

Interconnection Desirability Index

The interconnection desirability index is another approach that has been proposed for determining when to provide coordination between two signals (*4*). One-way link volumes are used to assess the need for progression in each direction on an arterial. The index also contains a factor to account for the effects of platoon dispersion. The formulation of the index is provided in the equation below:

$$I = \frac{0.5}{1 + t} \times \left[\frac{x \times q_{max}}{q_1 + q_2 + \ldots + q_x} - (N{-}2) \right]$$

where

$t =$ link travel time (link length divided by the average speed), expressed in minutes;

$x =$ number of departure lanes from upstream intersection;

$q_{max} =$ flow rate of the highest movement from the upstream intersection (usually the through movement);

$q_1{+}q_2{+} \ldots {+}q_x =$ sum of all the movements comprising the total flow arriving at the downstream approach; and

$N =$ number of arrival lanes feeding into the entering link of downstream intersection.

The value of the index ranges from zero to one. A value of one indicates a highly desirable condition for providing coordination. At the other end of the scale, an index value of zero represents a condition where coordination is least desirable. As shown in Figure II-2, if the index is below 0.25, isolated operation is recommended. When the index is 0.5 or more, interconnected operation is recommended. When the index falls between these thresholds, other factors should also be considered in assessing the need for interconnection (*4*). One study recommends that coordination should be provided when the index equals or exceeds 0.35 (*5*).

It should be noted that q_{max} represents the movement to be progressed from the upstream intersection through the downstream intersection. While this is usually the through movement, in some cases, a heavy turning movement may represent the majority of through traffic at the downstream intersection. The interconnection desirability index could also be used to identify those situations where it may be desirable to provide progression to a heavy turning movement.
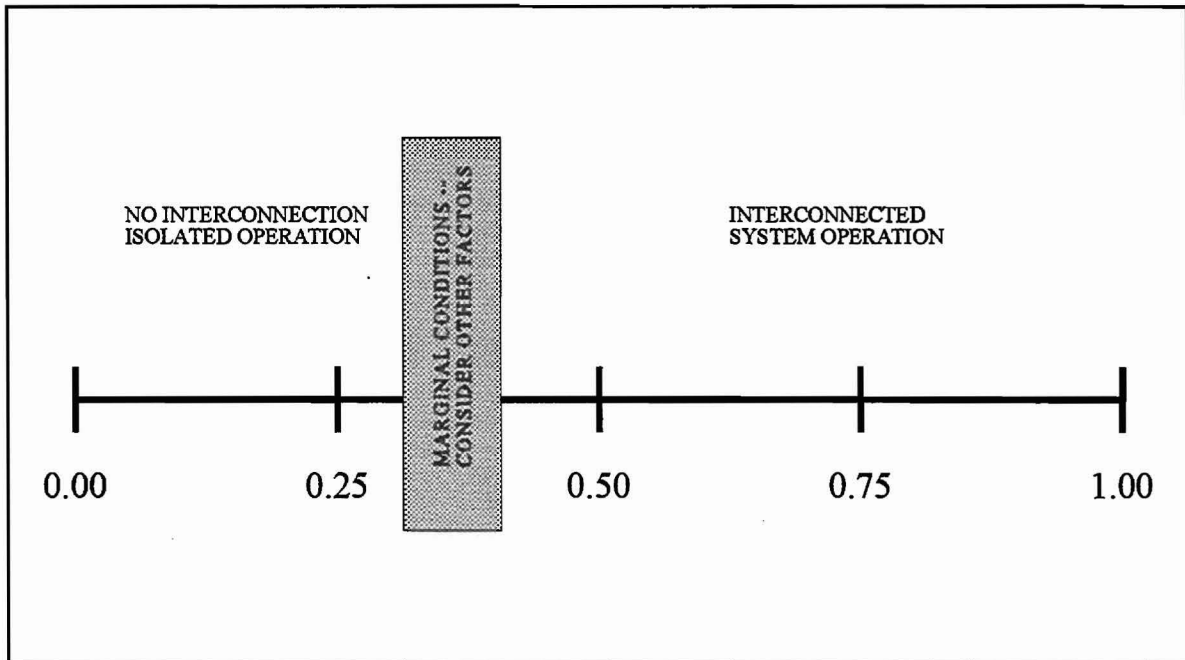
FIGURE II-2. Scale of Interconnection Desirability Index (4)

EVALUATION OF PROPOSED INTERCONNECTION DESIRABILITY INDEX

The objectives for evaluating the above index are as follows:

- determine whether or not the interconnection desirability index correctly identifies those periods where traffic performance is enhanced by providing coordination and those periods where traffic performance is enhanced by operating the traffic signals in an isolated mode;

- if the index is found unsatisfactory as a criterion for selecting between coordinated and isolated modes of operation, look at modifications to the index and other alternatives; and

- identify the appropriate threshold for determining when to operate the traffic signal system in a coordinated versus an isolated mode.

The following sections discuss simulation experiments conducted to achieve the objectives mentioned above. The simulation experiments were performed using TRAF-NETSIM. PASSER II-90 was used to generate optimum signal timing plans for the networks used in the simulations.

8

First, the interconnection desirability index, as proposed by Chang and Messer (*5*) and discussed in the previous section, was analyzed. Anomalies were noticed in this index. Hence, based on the limitations observed in the interconnection desirability index, a new index similar to the interconnection desirability index was proposed.

The new modified interconnection desirability index was analyzed using TRAF-NETSIM simulations. It was found that the modified index was able to reasonably identify conditions suitable for coordinated and isolated modes of operation.

Although the modified interconnection desirability index provides a good empirical method to identify the appropriate mode of operation for a given traffic condition, it was felt that a more analytical method based on traffic flow theory would be desirable. Hence a review of platoon progression and dispersion model used in PASSER II was conducted to determine if it is possible to automatically identify conditions suitable for interconnection within the program, obviating the need to use extraneous methods to do the same. The following sections discuss the findings of this review.

Chang's Interconnection Desirability Index

As discussed earlier, in the interconnection desirability index, one-way volumes are used to assess the need for progression in each direction on an arterial. The index also contains a factor to account for platoon dispersion. The value of the index ranges from zero to one. A value of one indicates highly desirable conditions for interconnection, and a value of zero indicates a least desirable value for interconnection.

The formulation of the index is shown in the equation below.

$$I = \frac{0.5}{1 + t} \times \left[\frac{x \times q_{max}}{q_1 + q_2 + \ldots + q_x} - (N-2) \right]$$

where

t = link travel time (link length divided by the average speed), expressed in minutes;

x = number of departure lanes from upstream intersection;

q_{max} = flow rate of the highest movement from the upstream intersection (usually the through movement);

$q_1+q_2+ \ldots +q_x$ = sum of all the movements comprising the total flow arriving at the downstream approach; and

N = number of arrival lanes feeding into the entering link of downstream intersection.

The value of 'x' and 'N', as defined in the above relation, were both assumed to mean the number of movements comprising the total flow arriving at the downstream intersection. There is ambiguity in the definition of these terms in the original report (5).

A network consisting of two major north-south arterials in Nacogdoches, Texas, as shown in Figure II-3, was used in this simulation experiment. The objective was to validate the findings reported in the previous studies and to fine-tune the threshold.

For the purposes of analysis of the index, it would be desirable to have a network where all the links had the same index value. In such a scenario, all intersections could be timed using both isolated and coordinated approaches, and then the delays and other measures of effectiveness resulting from the plans could be compared. Since all the links/intersections have the same index values, it would be simpler to converge on a threshold value for the index and also to evaluate how well the index identifies conditions appropriate for each mode of operation.

In order to obtain the same index value at all the intersections, the actual volumes at each intersection were adjusted such that they would yield the desired index value. Keeping the turning movements at the observed levels, the through volumes in the north-south direction were adjusted. In some cases, the turning movements were also adjusted to keep the through movement as the dominant movement while still yielding the desired index value for the link.

Several sets of volumes were generated to yield index values ranging from 0.1 to 0.5 at increments of 0.05. As mentioned earlier, it was hypothesized that there is a threshold index value beyond which coordinated mode of operation would always perform better than isolated. Also, it was assumed that such a threshold would not be higher than 0.5 based on observations in previous studies (5).

For volumes representing each index value, PASSER II-90 was used to generate both isolated and coordinated optimum signal timing plans. These timing plans were coded into TRAF-NETSIM and simulated.

Link and intersection delays were obtained from TRAF-NETSIM for both isolated and coordinated modes of operation and compared. Only one TRAF-NETSIM run was made for each case. During preliminary stages, five runs were made with different random number seeds; however, it was observed that the difference between MOEs generated in these runs was very minimal. Hence, it was decided to make just one TRAF-NETSIM run. Figures II-4 to II-11 show the network-wide, intersection and directional delays for both isolated and coordinated modes of operation as a function of the index values. It is evident from Figure II-4 that the network-wide delay is marginally higher for coordinated mode of operation for index values less

than 0.45. For the interconnection desirability index values greater than 0.45, however, the network-wide delay is higher for isolated mode of operation.
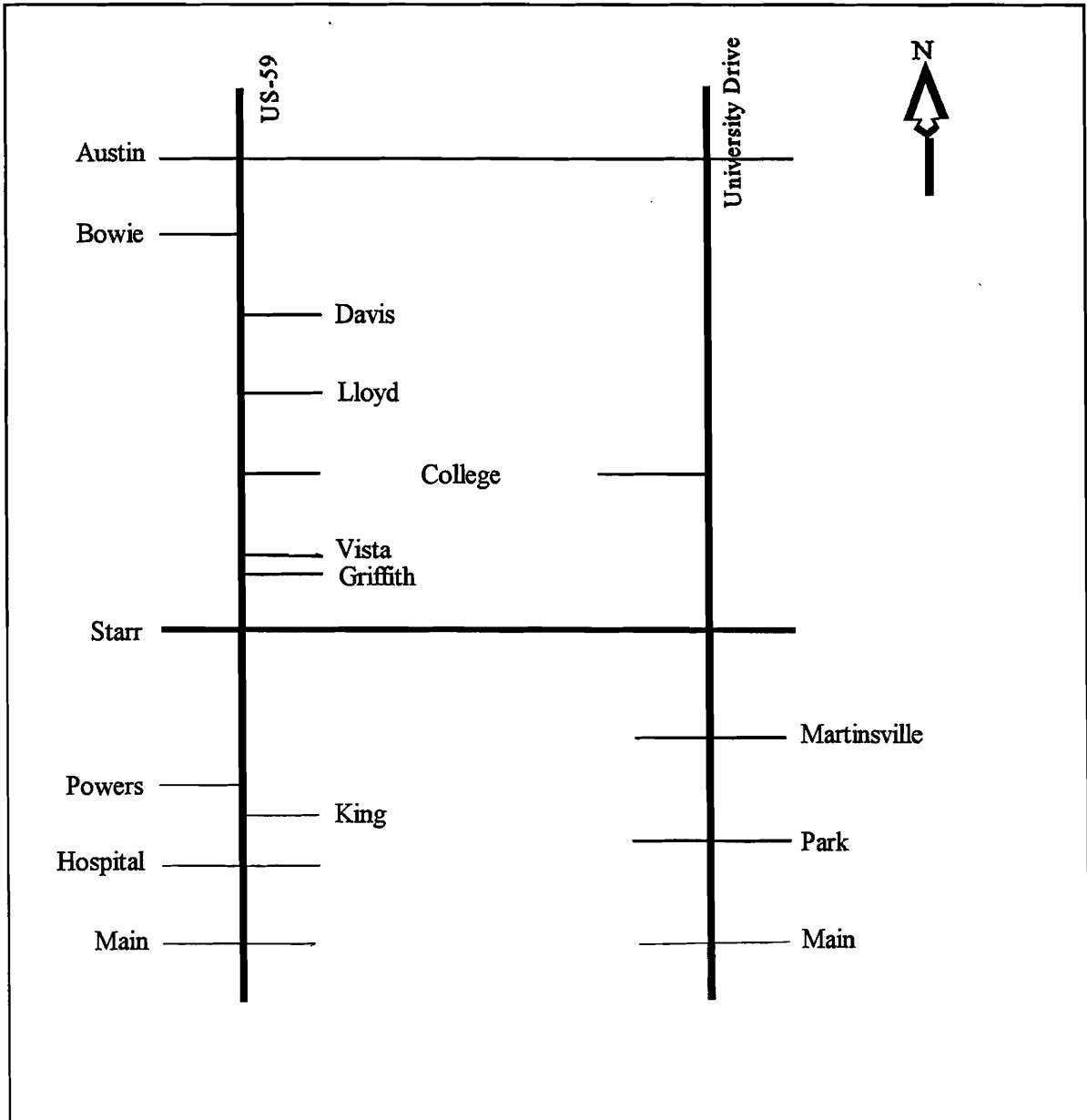
FIGURE II-3. Network Used in Evaluation of Interconnection Desirability Index
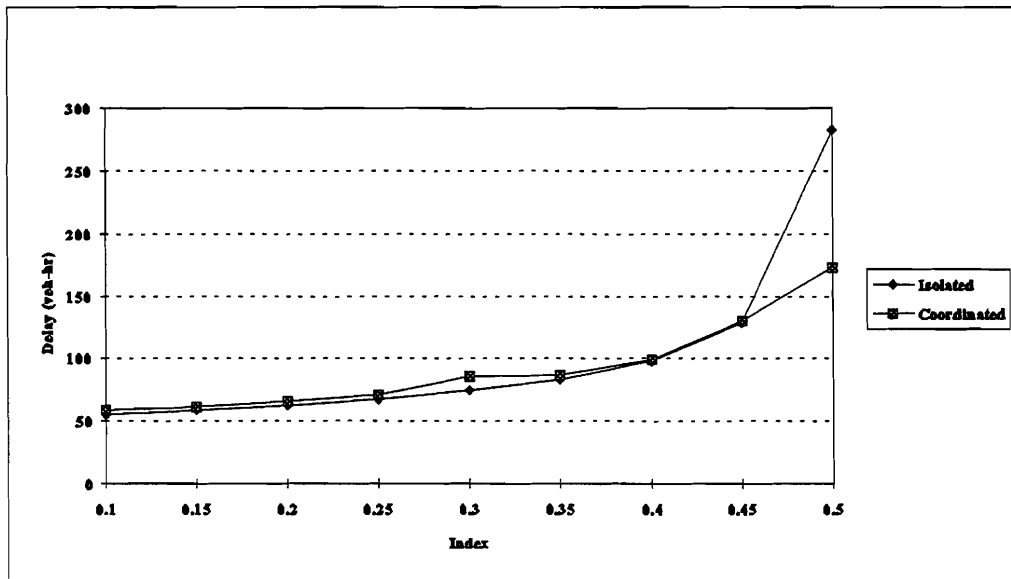
**FIGURE II-4. Comparison of Network-wide Delay for Isolated and
Coordinated Control**

Figures II-5 and II-6 depict the total intersection delays for each north-south arterial. This delay is the sum of all intersection delays along the arterials. Each intersection delay was estimated by aggregating the delays on each approach to the intersection. Intersection delays represent the delay experienced by the cross street as well as the arterial. It can be seen that the total intersection delays are very similar for both isolated and coordinated modes of operation on US-59; however, for University Drive, isolated mode of operation resulted in a higher total intersection delay for index values greater than 0.45.

Figures II-7 through II-10 depict the total directional delays on both arterials. Directional delays constitute the sum of all link delays in each direction. This represents the delays experienced by the through traffic alone and demonstrates if coordinated mode of operation would at least be beneficial for the progressed movement. It can be seen from Figures II-7 and II-8 that the directional delays on US-59 are very similar for both modes of operation. For index values between 0.4 and 0.45, the directional delays are marginally higher for isolated mode than coordinated mode. For University Drive, however, the isolated mode of operation is consistently better than the coordinated mode considering the directional delays. This may be due to the long separation between intersections on University and heavy mid-block volumes.

12
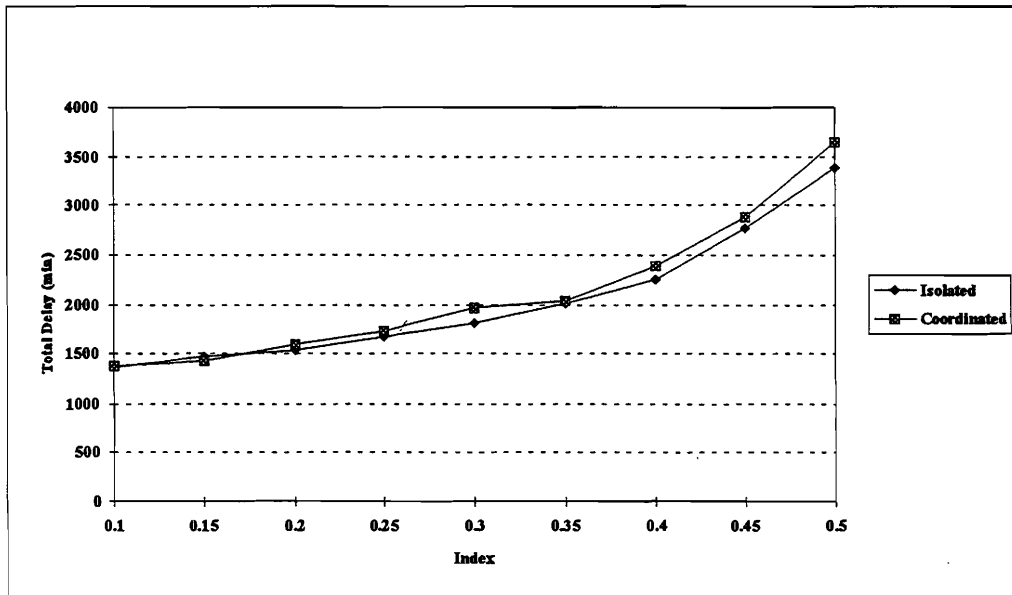
FIGURE II-5. Total Intersection Delay on US-59 Under Isolated and Coordinated Control
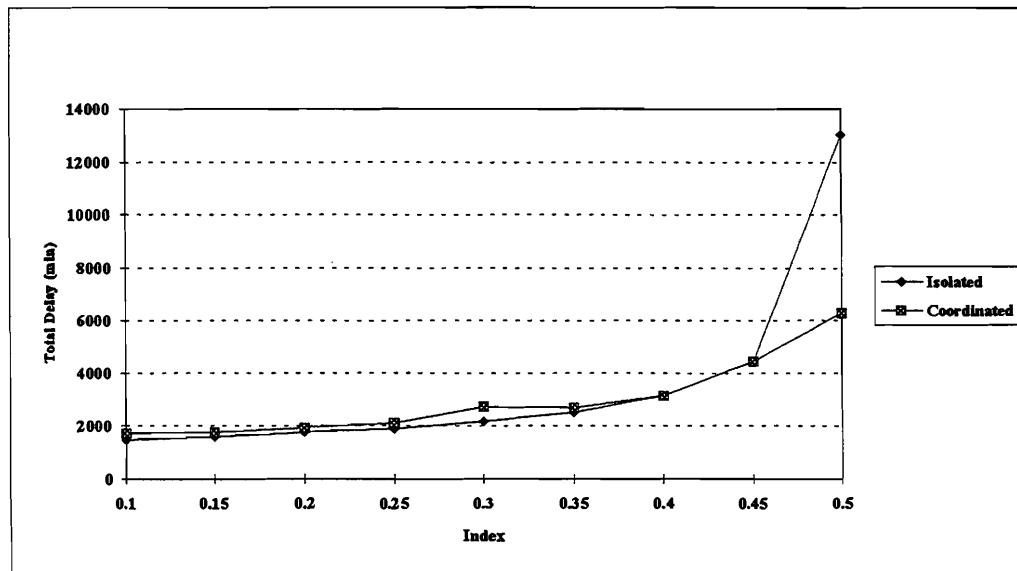
FIGURE II-6. Total Intersection Delay on University Drive Under Isolated and Coordinated Control

13
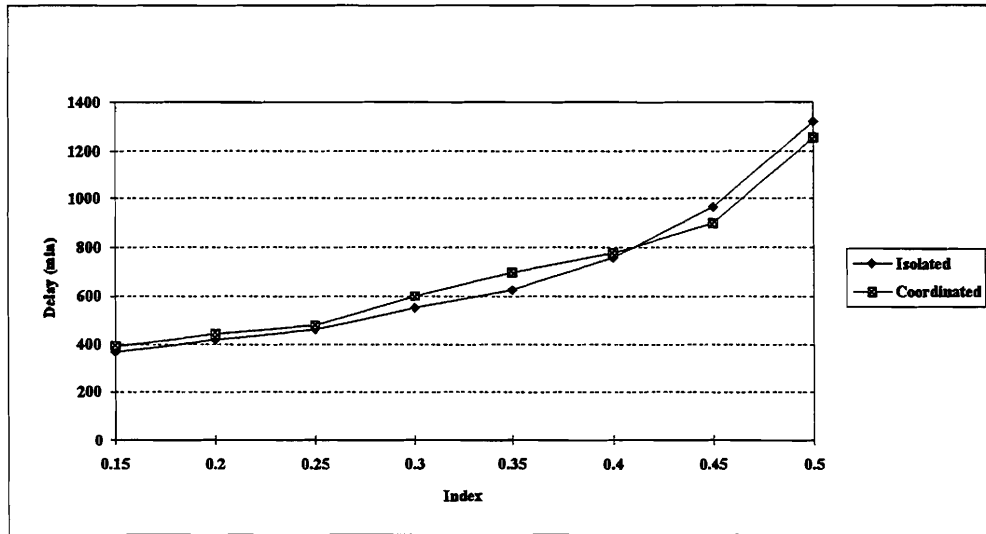
**FIGURE II-7. Total Directional Delay on US-59 Northbound
Under Isolated and Coordinated Control**
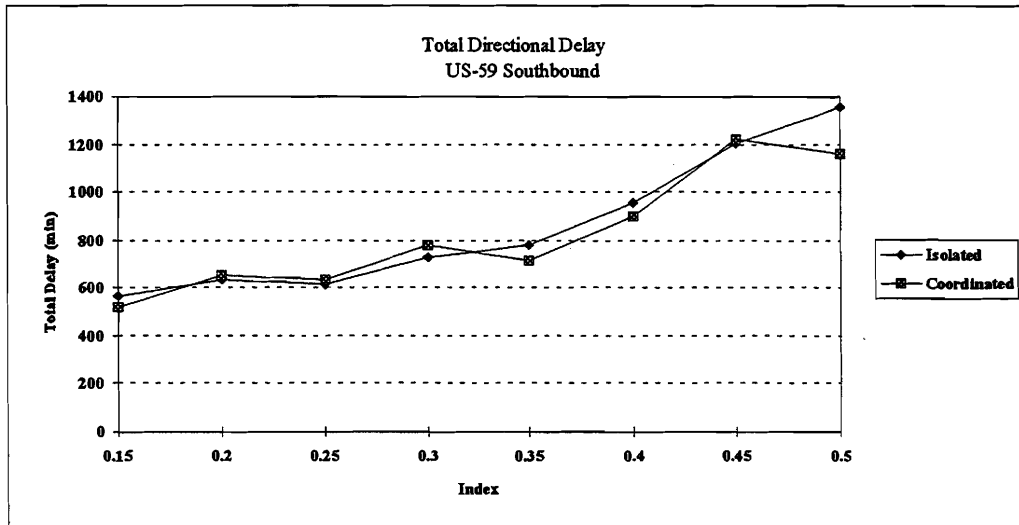
**FIGURE II-8. Total Directional Delay on US-59 Southbound Under
Isolated and Coordinated Control**

14
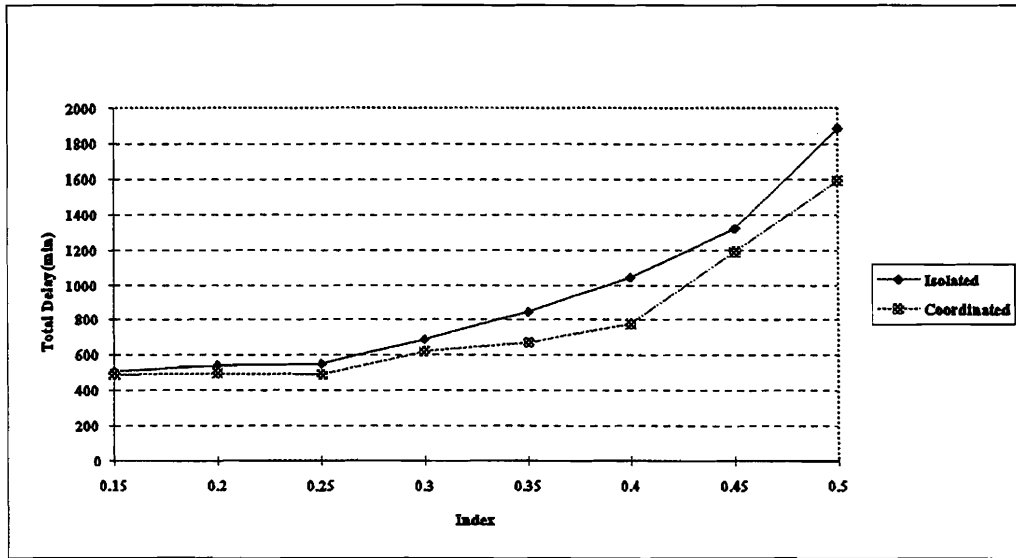
**FIGURE II-9. Total Directional Delay on University Northbound
Under Isolated and Coordinated Control**
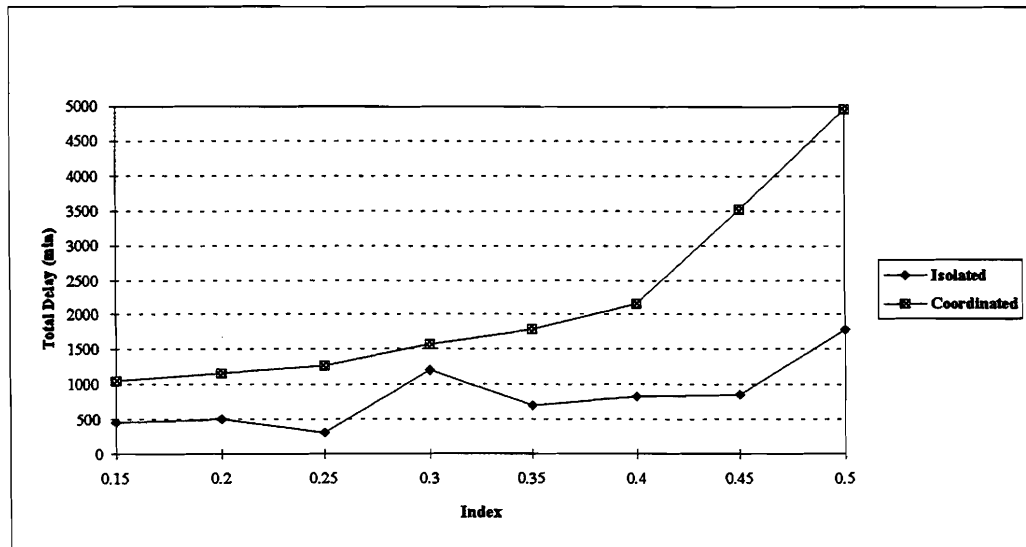
**FIGURE II-10. Total Directional Delay on University Southbound Under
Isolated and Coordinated Control**

15
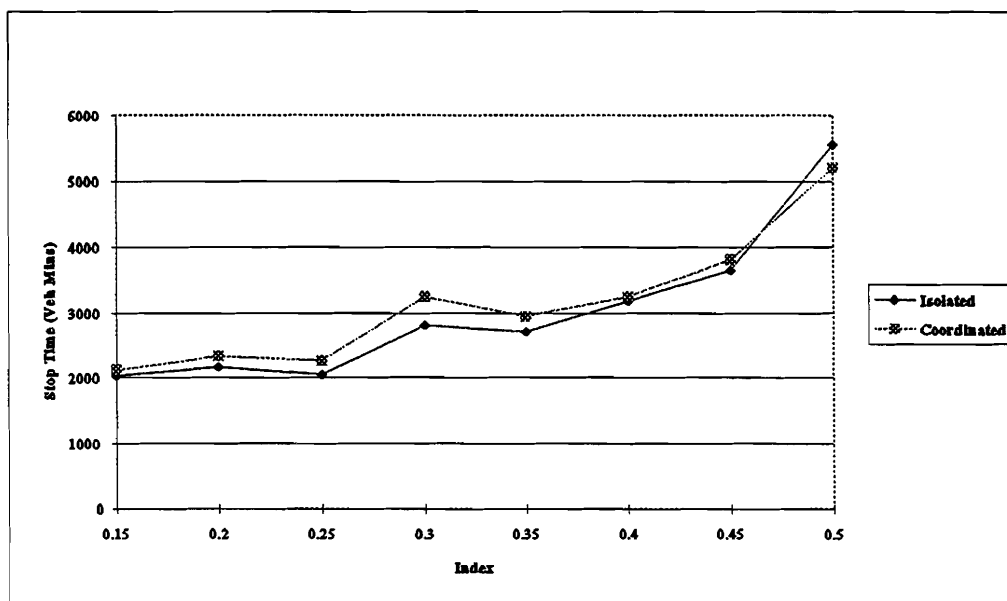
FIGURE II-11. Network-wide Stop Time Under Isolated and Coordinated Control

The network-wide stop time, which is indicative of the number of stops, is depicted in Figure II-11. It can be seen that the stop time for both modes of operation are very similar, although the stop time for isolated mode tends to increase faster than the coordinated mode of operation for index values greater than 0.45.

From these observations, it is apparent that an index value of about 0.45 may be the threshold for choosing between isolated and coordinated modes of operation; however, as observed on University Drive, when the mid-block volumes are very high, the index does not serve as a good indicator for choosing between isolated and coordinated modes of operation. Therefore, a slightly modified interconnection desirability index may be needed that will take into account the mid-block traffic also. The following section describes the proposed modified index that attempts to capture the effect of mid-block traffic while also accounting for the percentage of through traffic and platoon dispersion. The analysis of the modified index is also presented.

MODIFIED INTERCONNECTION DESIRABILITY INDEX

In order to overcome the limitations of the index discussed above, a new index was proposed. This new index also incorporates the platoon dispersion aspect through the travel time term as in the index proposed by Chang and Messer (5) discussed above. The modified index mainly differs from the original index in its treatment of mid-block traffic. Heavy mid-

16

block traffic can have a detrimental effect on platoons and thus render signal coordination ineffective. The following relation describes the index:

$$I = \left(\frac{1}{1+t}\right) * \left(\frac{q_{thru}}{Q}\right)$$

where
- t = link travel time in minutes;
- Q = total volume (left+through+right) at the downstream intersection; and
- q_{thru} = total through volume at the upstream intersection, vph (if Q is greater than upstream through volume); or
- = upstream through volume - net mid-block exiting volume for the link (if Q is less than upstream through volume).

The first term in the above formulation represents platoon dispersion. Higher travel time between intersections leads to higher dispersion, and hence a lower index value. The second term represents the fraction of the total traffic downstream that constitutes the progressed traffic component. Usually the through movement is progressed. As the volume of the progressed traffic component increases, the benefits of interconnection increase. The index has a range between 0 and 1, where 0 represents the least, and 1 represents the most favorable conditions for interconnect.

It should be noted that the total downstream intersection volume also includes mid-block traffic. When the total volume at the downstream intersection is less than the upstream through volumes, i.e., when there is a net exit of vehicles from the link mid-block, it is assumed that the exiting vehicles are from the upstream through (progressed) movement. Hence q_{thru} in the above equation is reduced by the number of vehicles exiting mid-block.

In order to test the applicability of this new index, several scenarios were developed. TRAF-NETSIM was used to compare both coordinated and isolated modes of operation for each scenario; however, it was decided that a simple two intersection network should be used to test the effectiveness of this index. It was felt that a larger network, such as the one used in the previous case, does not add to the accuracy of the findings. In any case, the index is only intended to serve as a general indicator of the appropriateness of a coordinated or isolated mode of operation for a given traffic situation. A simple network would permit evaluation of more varied conditions than a large network.

The two intersections in the test case were loaded with traffic such that the critical flow ratio (v/s) was equal to 0.5 or 0.65. A critical flow ratio of 0.5 would yield a v/c ratio of 0.7

17

when a minimum delay cycle, according to Greenshield's method, is provided (assuming a 4 phase operation). Similarly, a flow ratio of 0.65 would yield an intersection v/c of 0.85.

From these two traffic loadings, several scenarios were developed. First, both intersections were loaded such that their critical flow ratio was equal to 0.5 (Case A). Keeping the critical flow ratio at the predetermined level, the mid-block volume was varied from 0 to 40 percent of the through traffic. To maintain the predetermined flow ratios, the through traffic was reduced by the additional volumes introduced as mid-block traffic. Figure II-12 depicts this procedure for Case A. Similar procedures were adopted for other cases.

It should be noted that in Figure II-12 as the mid-block volume increases, the upstream through movement was correspondingly reduced. The total volume at the downstream intersection remained the same in all the cases.

Next, both intersections were loaded such that their critical flow ratio was 0.65 (Case B). Again, in each case, the mid-block volume was varied as described above.

In a third scenario (Case C), both intersections were loaded at a saturation flow rate of 0.5. But in this scenario, instead of varying the mid-block volume, the upstream turning movement volumes were varied. This scenario helped test the index for varying percentages of the upstream turning movement component of the link volumes.

The fourth scenario (Case D) was a situation where both intersections were loaded at different levels to test the validity of the index when the two interconnected intersections had different levels of traffic volumes, which was the most prevalent situation. In this scenario, there were only three variations in the mid-block volumes, as against five in the previous cases. The mid-block volumes were varied in steps of 10 percent of the corresponding through movement. The inherent imbalance in the levels of traffic at the two intersections rendered it difficult to provide more variations in mid-block traffic. In one of the cases (Case D-3), the mid-block volume was higher than the upstream through volume.

It should be noted that the actual field conditions could have any combination of traffic levels. The scenarios in this experiment were intended to provide different situations qualitatively.

The test network consisted of two intersections with an east-west arterial. The arterial was a two-lane facility with exclusive left-turn bays. The cross-streets were one-lane roadways with exclusive left-turn lanes in all the cases except case C. In case C, the cross-street geometry was modified to increase the saturation flow rate for turning movements so that the cross-street turning movement volumes could be raised without simultaneously increasing the flow ratio.
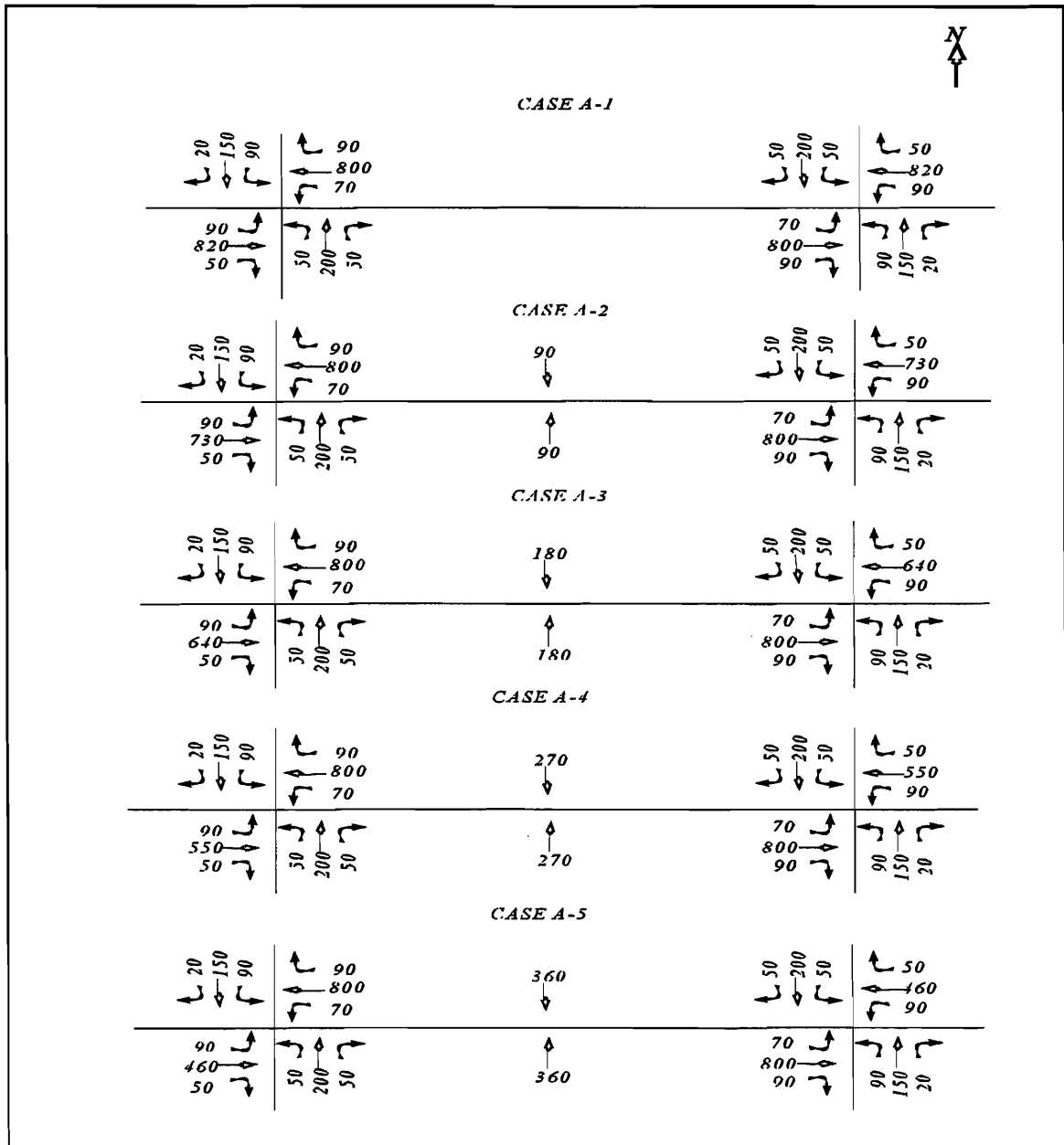
FIGURE II-12. Traffic Loading for Case A

19

For each of the scenarios above, the intersection spacing was varied from 18.3 to 1220 meters at increments of 18.3 meters. For each combination of traffic volumes and spacing, coordinated and isolated timing plans were developed using PASSER II-90. These optimum timing plans were coded into TRAF-NETSIM and simulated; thus, 126 different scenarios were tested, which would not have been possible on a larger network.

Figures II-13 through II-34 depict the results of the simulation discussed in the following paragraphs. The figures depict the difference in delay and stops between coordinated and isolated modes of operation. This difference is the value obtained after subtracting the coordinated mode delay/percent stops from the corresponding quantity for the isolated mode. A negative value indicates that coordinated mode is advantageous over isolated mode.

The establishment of a threshold value for the index so that the threshold is good under most traffic and geometric conditions was the main objective in analyzing the information presented in Figures II-13 through II-34. If such a threshold can be established successfully, it implies that the index is suitable for identifying conditions for interconnection with some degree of accuracy. The interconnection desirability index as formulated here is only an empirical value and cannot be expected to be an accurate indicator of whether operating the signals in a coordinated or isolated mode is more suitable for every given traffic condition.

In Case A, it can be seen from both link and network-wide measures of effectiveness (Figures II-13 to II-18) that the coordinated mode of operation is better, in most cases, where the index is greater than 0.4; however, for index values between 0.3 and 0.4, it appears that both modes of operation perform equally (i.e., an equal number of observations on either side of the difference axis).

As in Case A, in Case B, similar observations can be made from Figures II-19 to II-24. As stated earlier, Case B is similar to Case A except that the overall traffic level at the intersections is higher. Both the network-wide and link delays and stops are lower in the coordinated mode of operations when the index value is greater than 0.4.

In Case C where the cross-street turning movement volumes were varied instead of mid-block traffic, similar trends as those discussed above were noticed (see Figures II-25 to II-30). Although the isolated mode was better in more observations with an index of less than 0.4, it should be noticed that when the index is higher, the coordinated mode of operation results in less delay and fewer stops.

As discussed earlier, in Case D the traffic levels at the two intersections are different, thus leading to different index values in each direction. Therefore, it was not possible to plot the network-wide measures of effectiveness as a function of the interconnection desirability index.
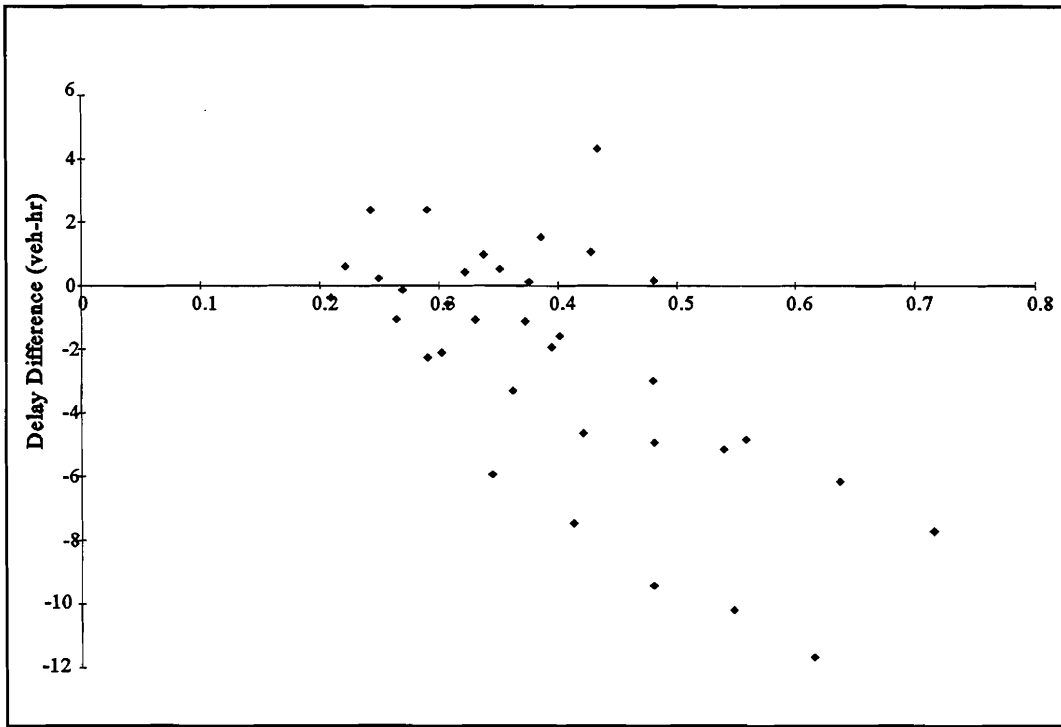
FIGURE II- 13. Difference in Network-wide Delay — Case A
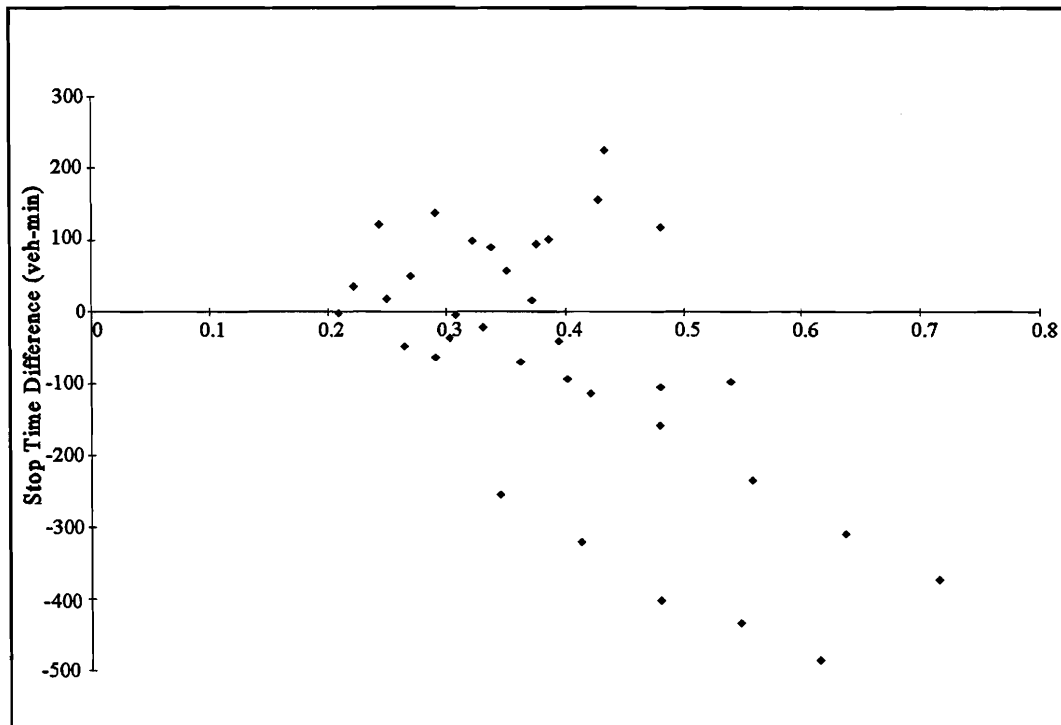
FIGURE II-14. Difference in Network-wide Stop Time — Case A
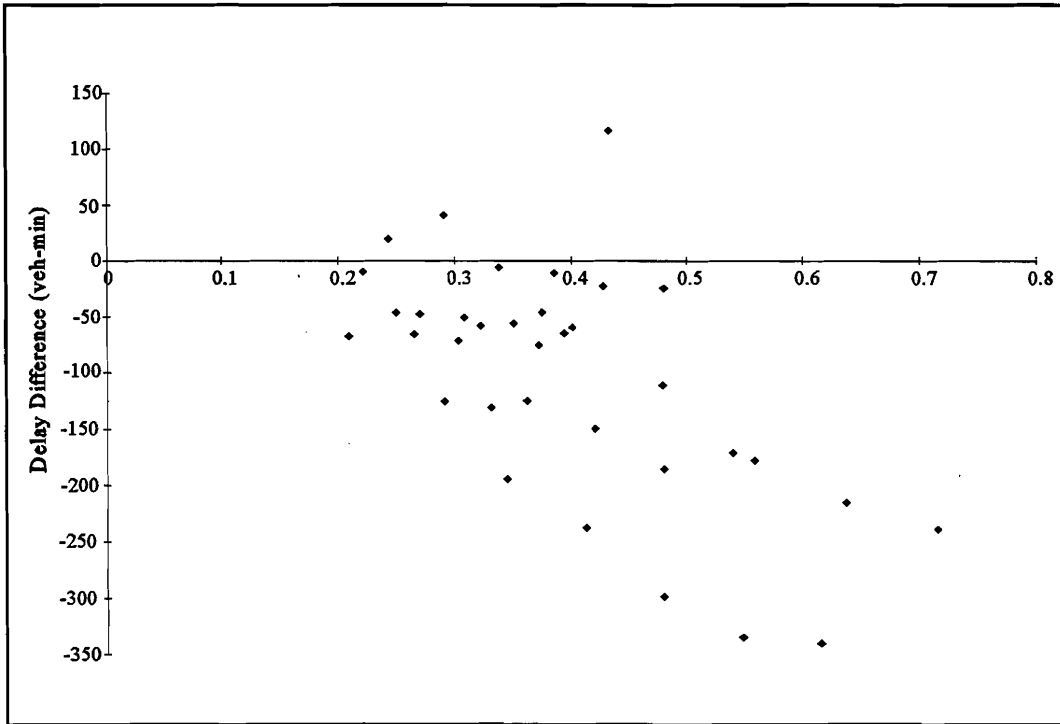
FIGURE II-15. Difference in Eastbound Link Delay — Case A
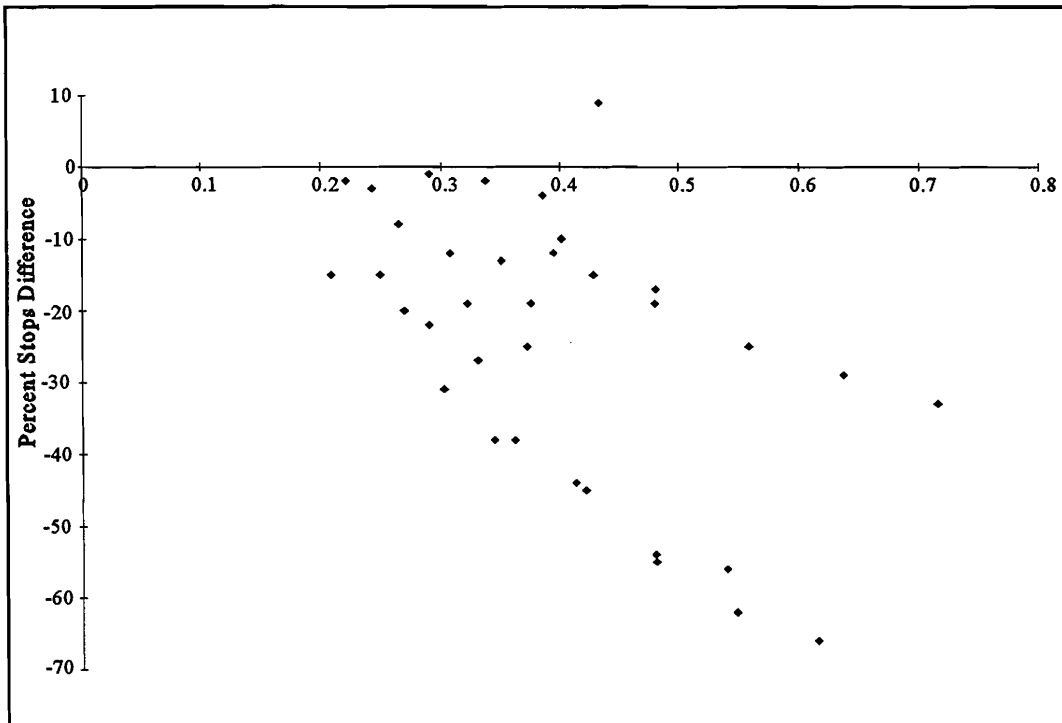
FIGURE II-16. Difference in Eastbound Link Percent Stops — Case A
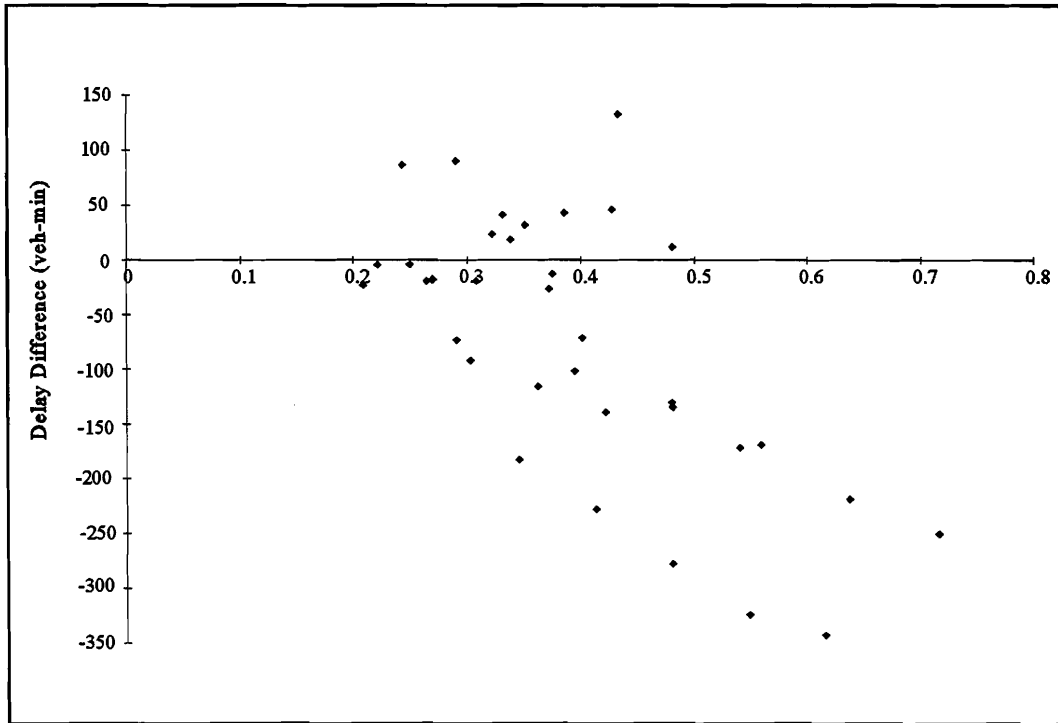
FIGURE II-17. Difference in Westbound Link Delay — Case A
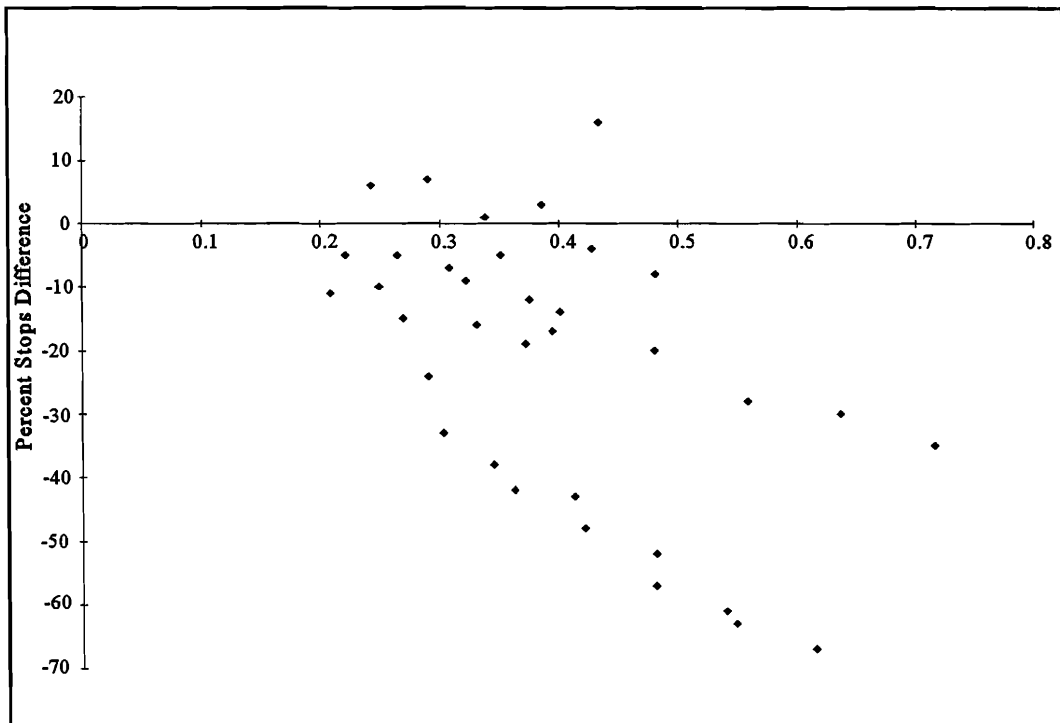
FIGURE II-18. Difference in Westbound Link Percent Stops — Case A
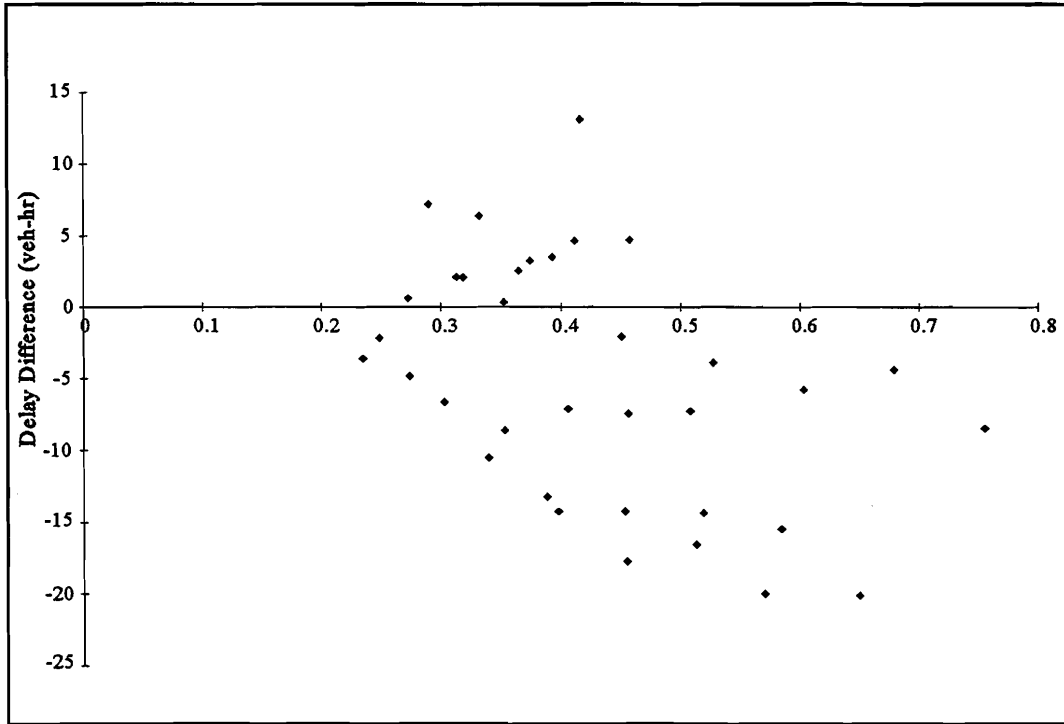
FIGURE II-19. Difference in Network-wide Delay — Case B
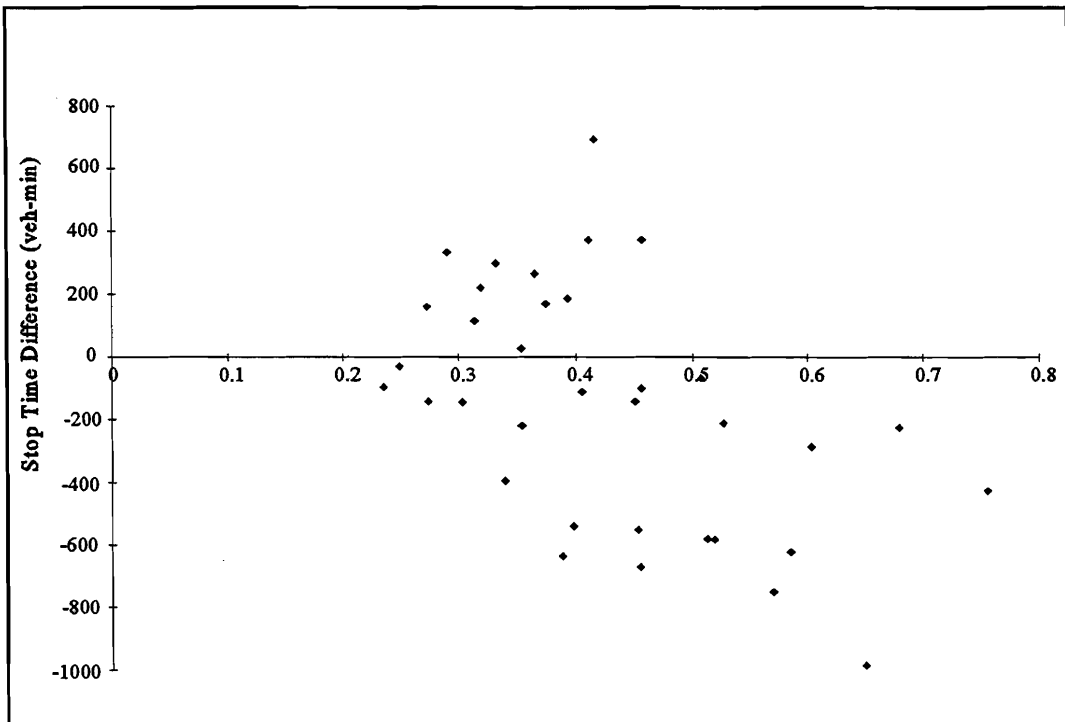
FIGURE II-20. Difference in Network-wide Stop Time — Case B
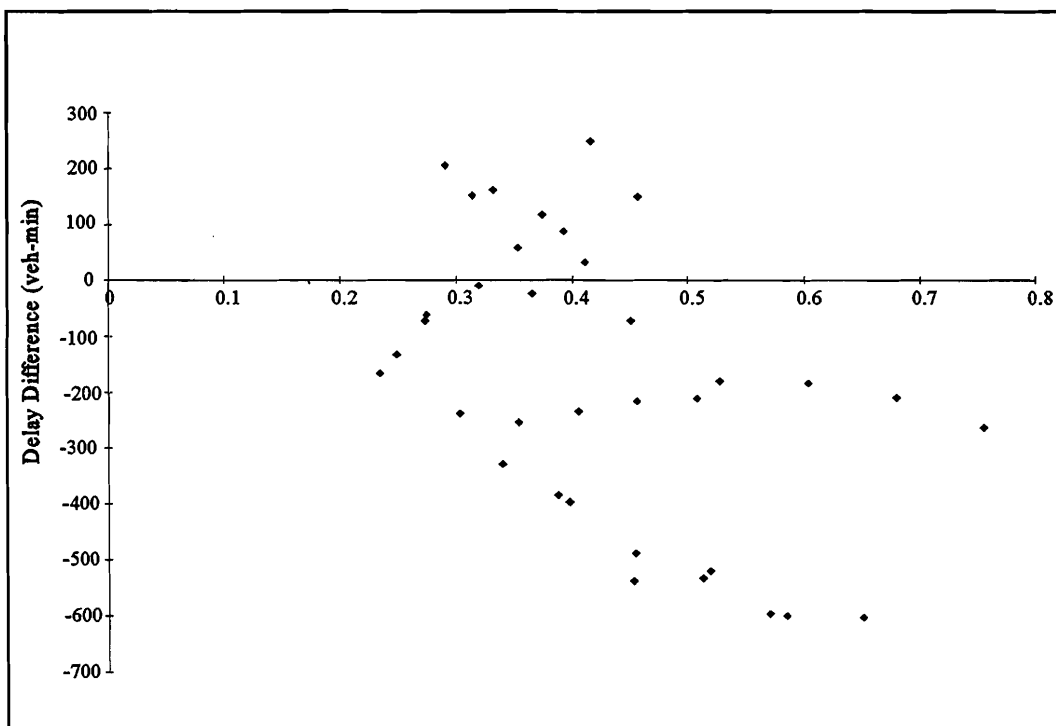
FIGURE II-21. Difference in Eastbound Link Delay — Case B
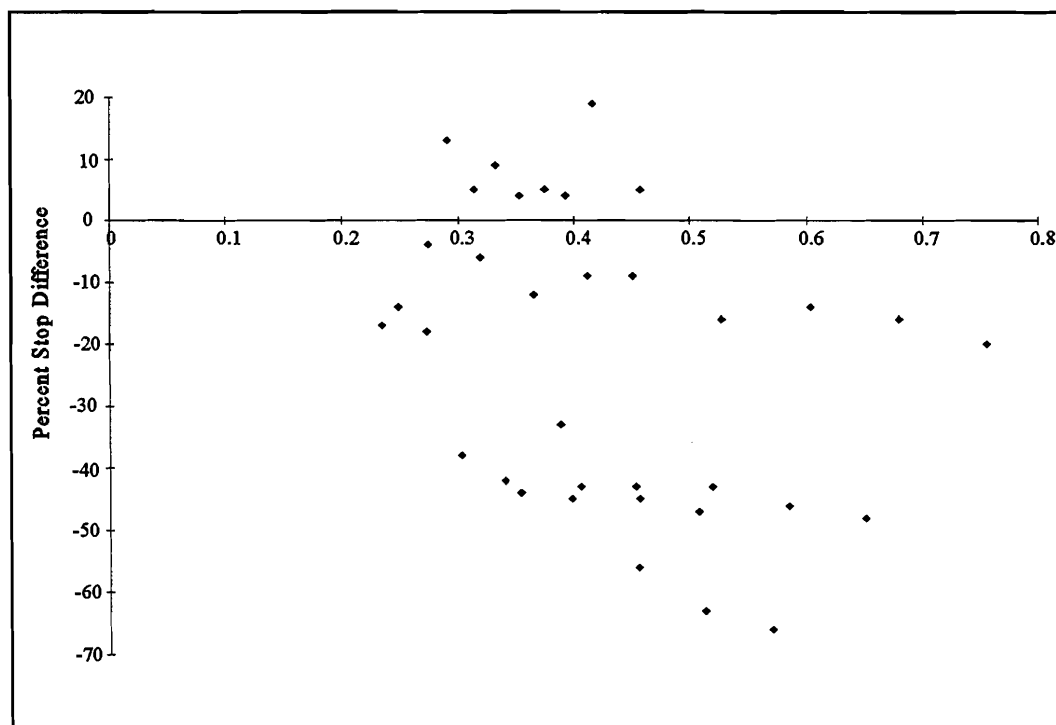
FIGURE II-22. Difference in Eastbound Link Percent Stops — Case B
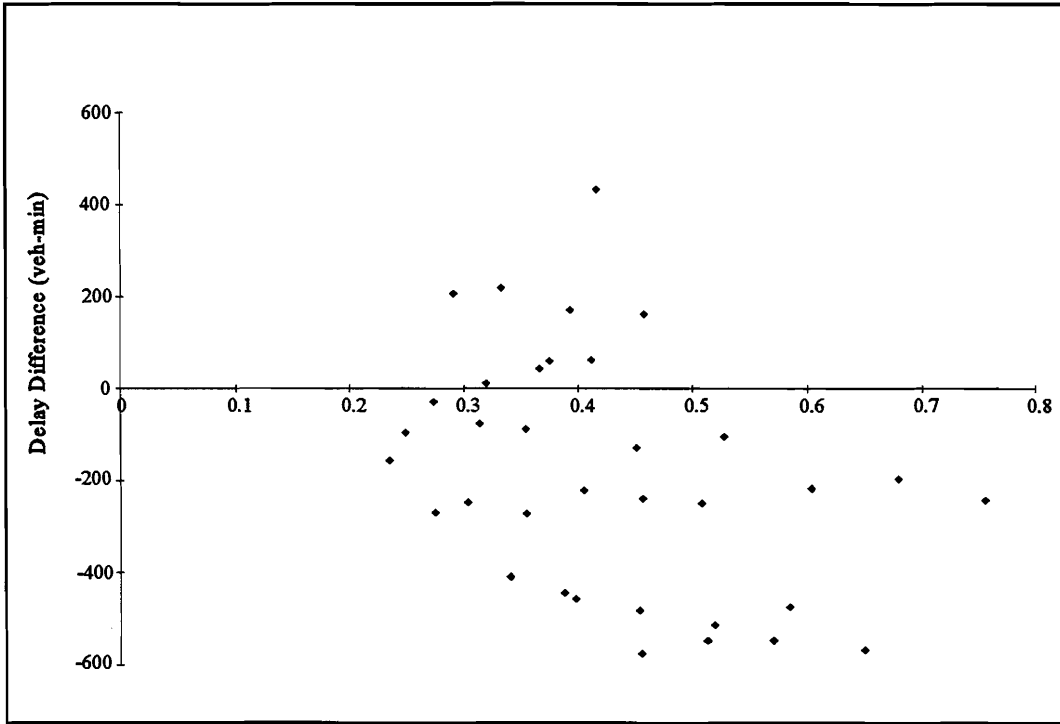
FIGURE II-23. Difference in Westbound Link Delay — Case B
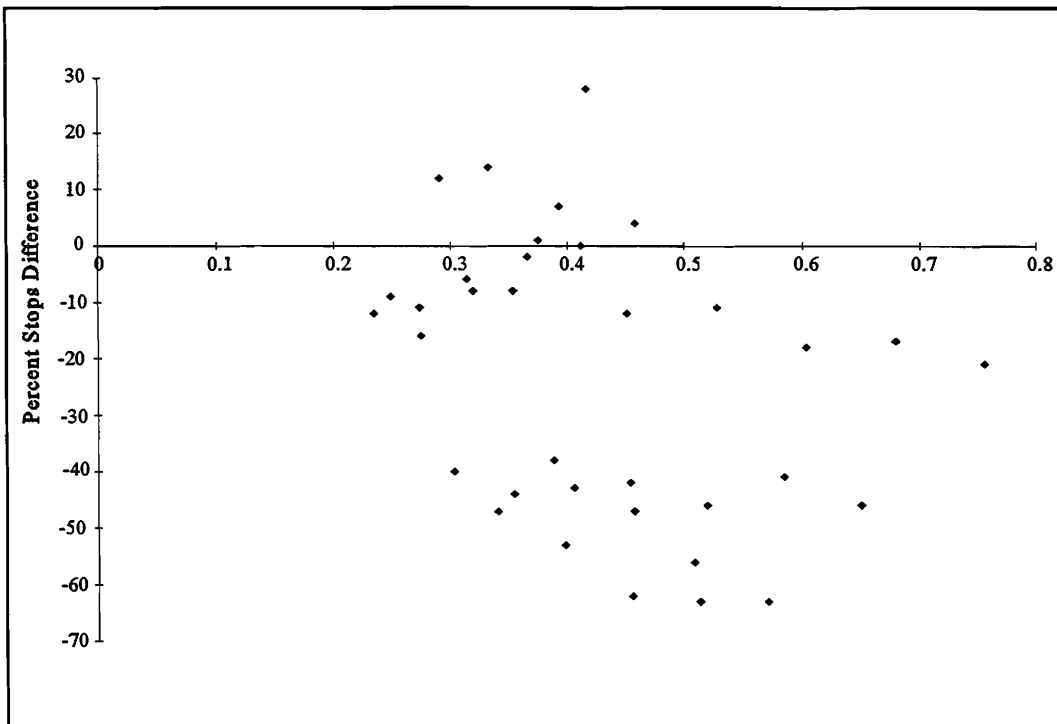
FIGURE II-24. Difference in Westbound Link Percent Stops — Case B

26
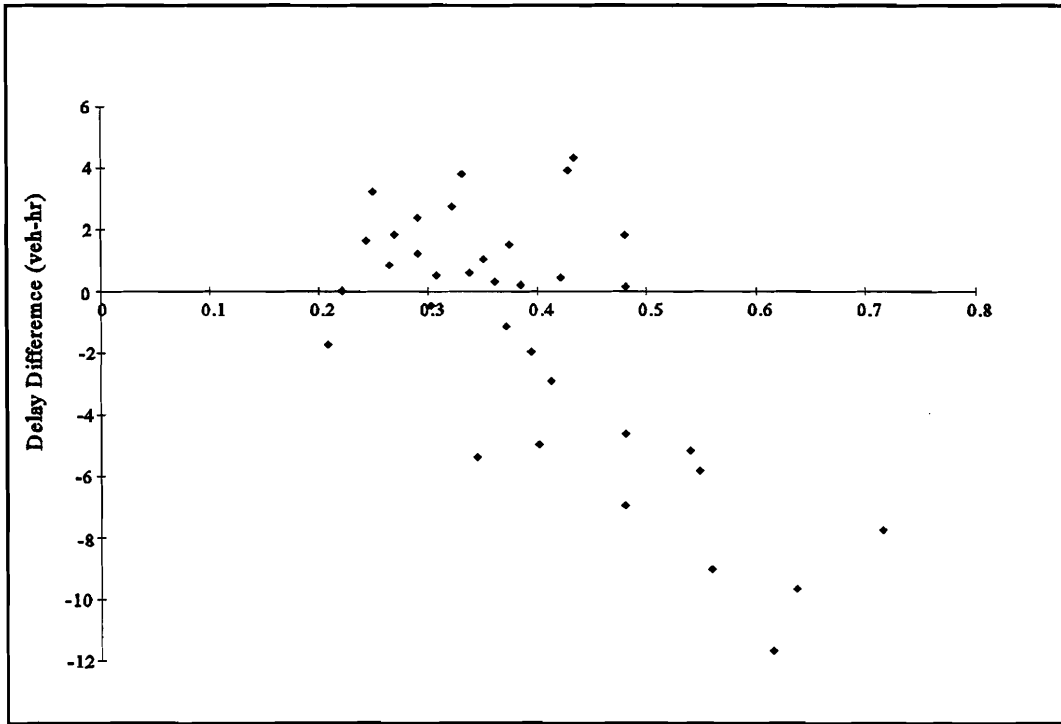
FIGURE II-25. Difference in Network-wide Delay — Case C
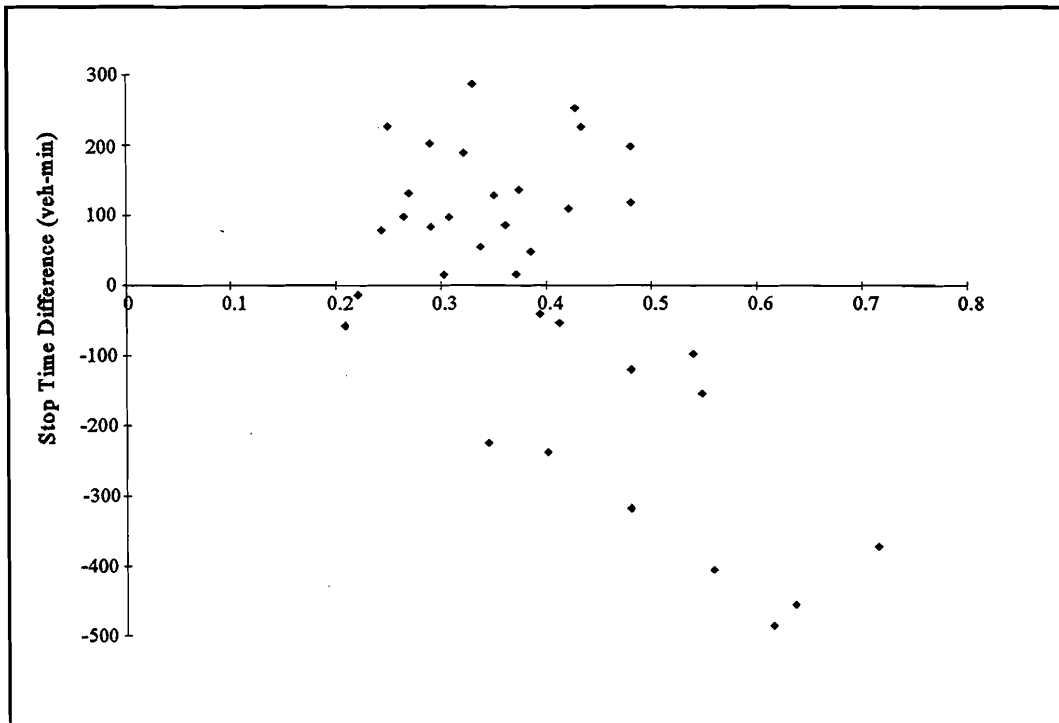
FIGURE II-26. Difference in Network-wide Stop Time — Case C
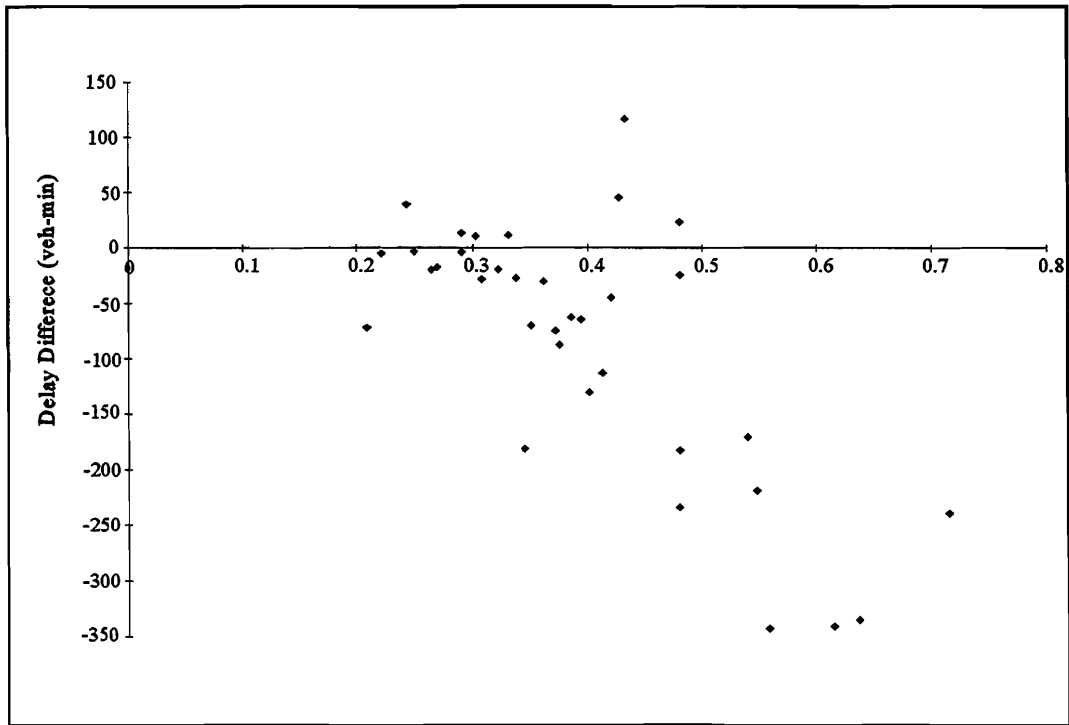
FIGURE II-27. Difference in Eastbound Link Delay — Case C
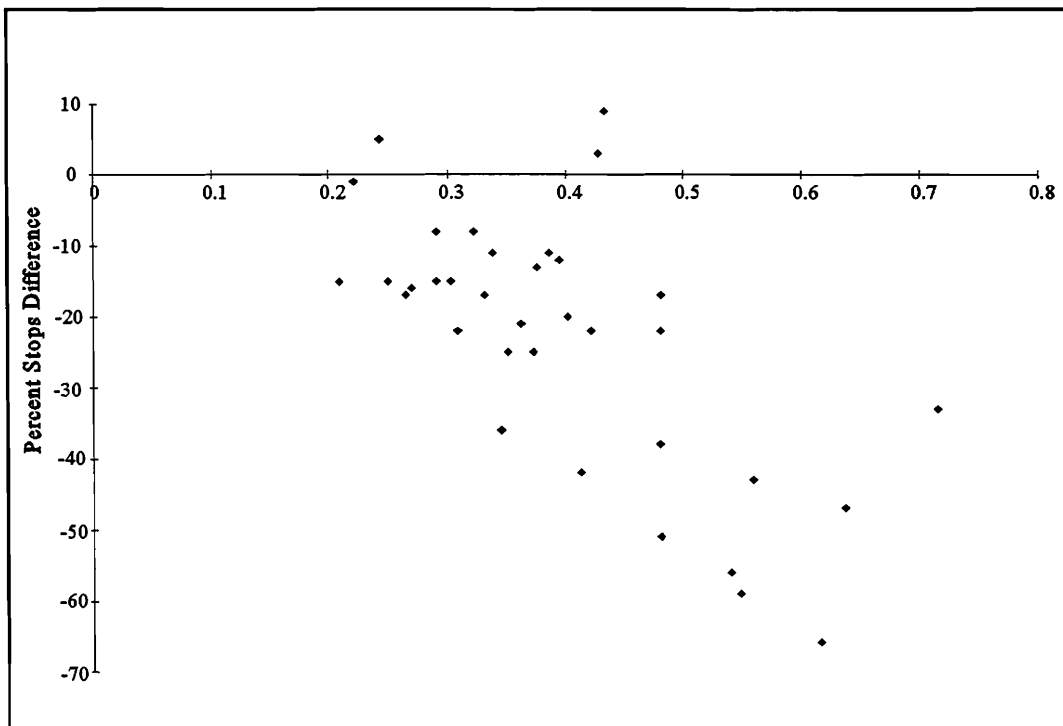
FIGURE II-28. Difference in Eastbound Link Percent Stops — Case C
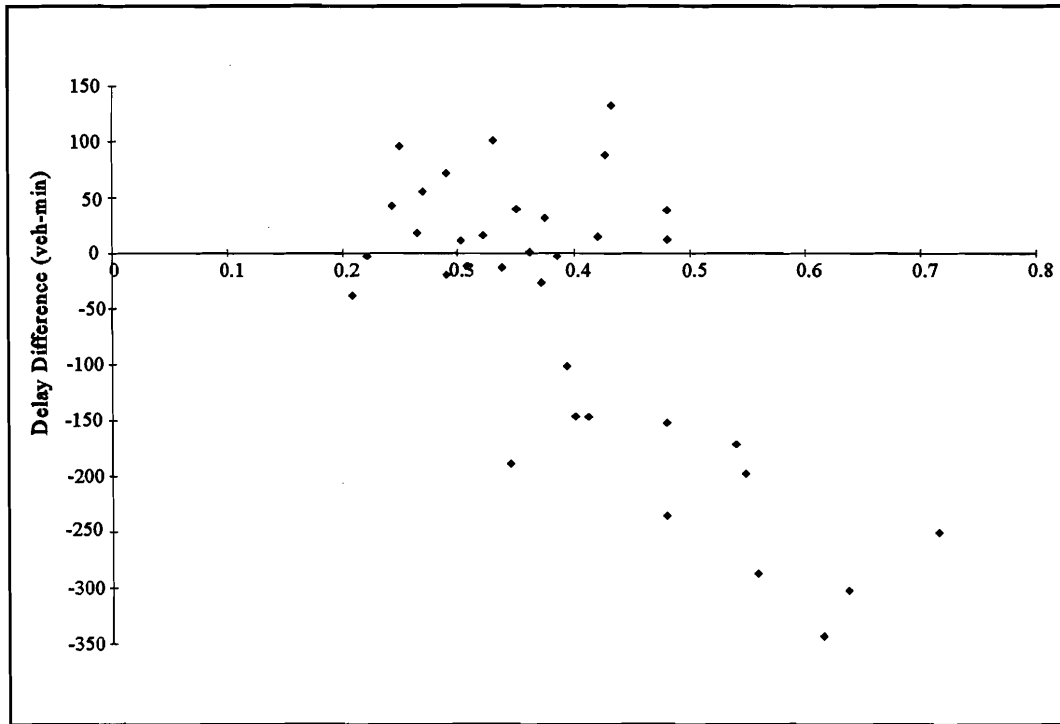
FIGURE II-29. Difference in Westbound Link Delay — Case C
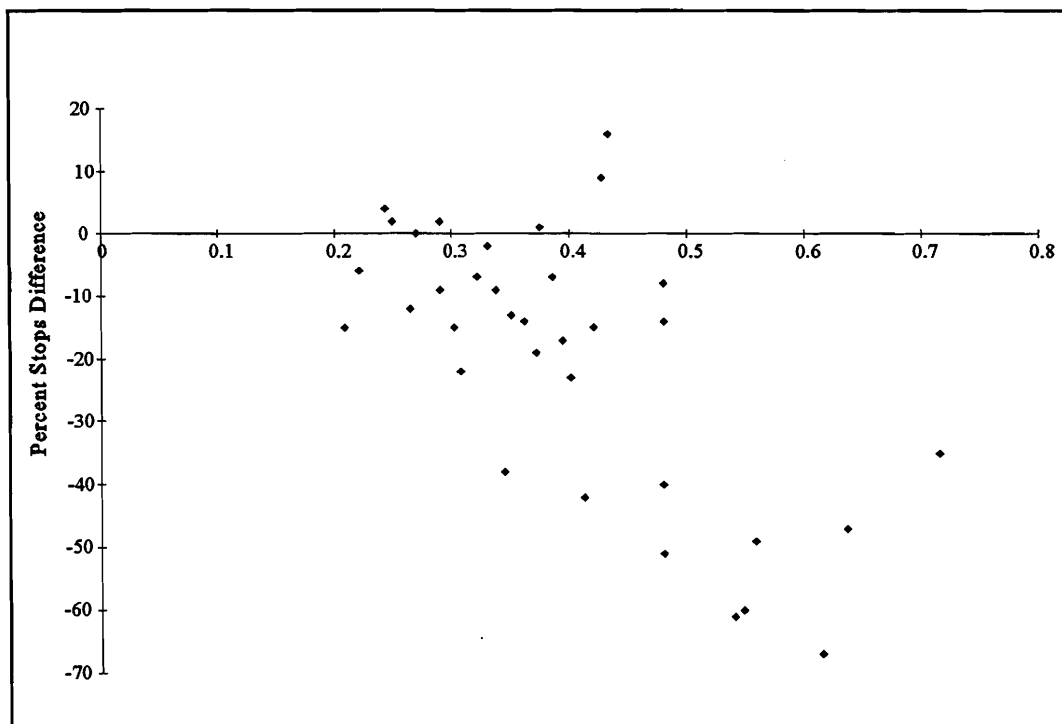
FIGURE II-30. Difference in Westbound Link Percent Stops — Case C

29
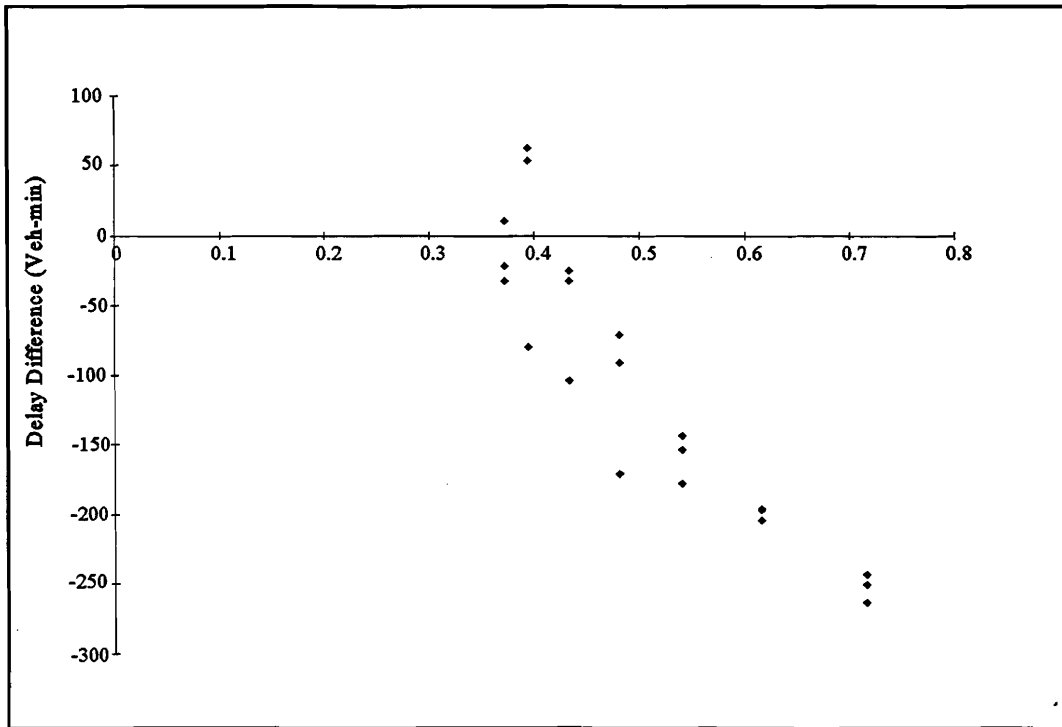
FIGURE II-31. Difference in Eastbound Link Delay — Case D
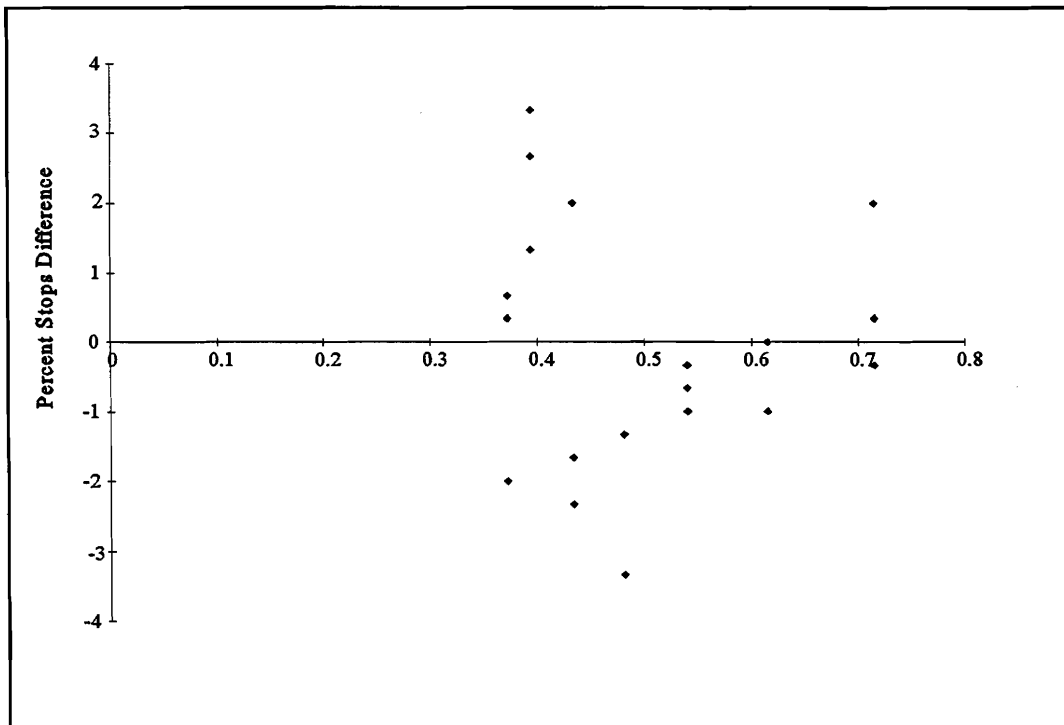
FIGURE II-32. Difference in Eastbound Link Percent Stops — Case D
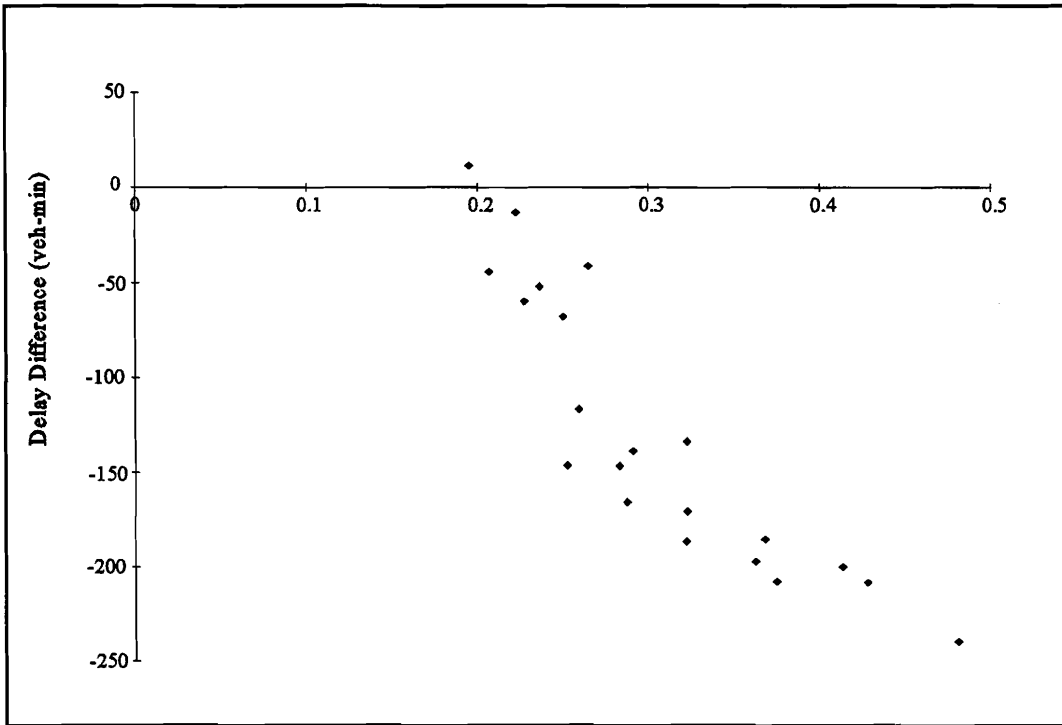
FIGURE II-33. Difference in Westbound Link Delay — Case D
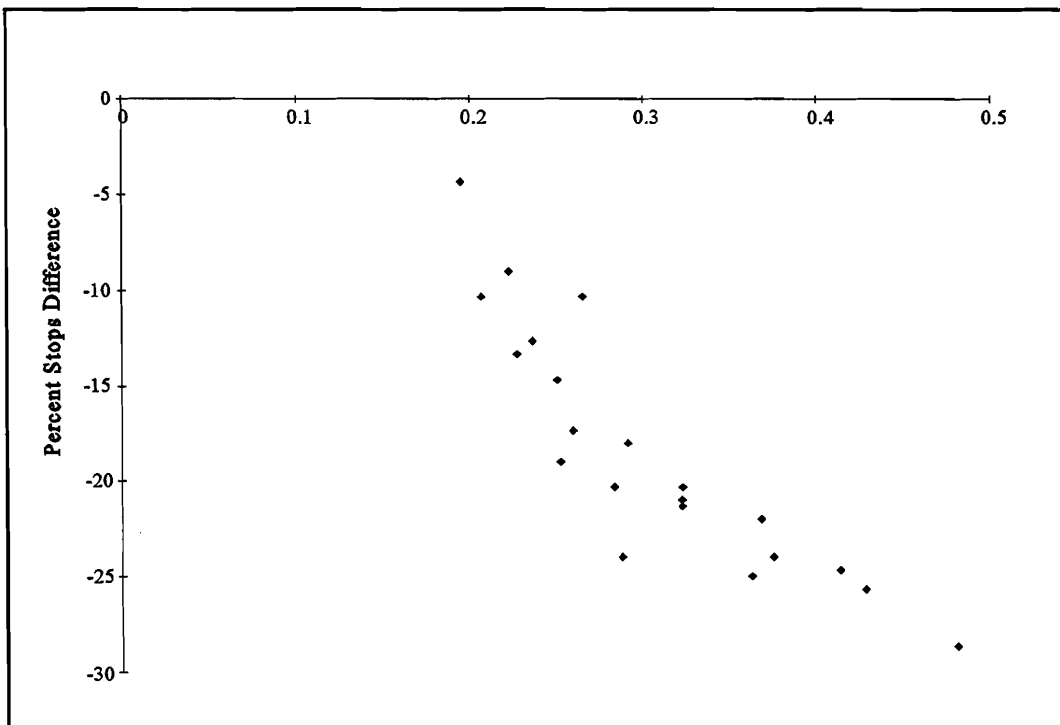
FIGURE II-34. Difference in Westbound Link Percent Stops — Case D

31

As in earlier cases, the coordinated mode of operation is more beneficial than the isolated mode of operation in the eastbound direction for index values greater than 0.4. It can be observed from Figures II-33 and II-34 that for the westbound arterial segment, the coordinated mode of operation appears to be more beneficial over the isolated mode of operation even at very low index values. This may be due to the fact that mid-block traffic enters the street before or after the platoon traffic and thus does not interfere with platoon progression (exiting traffic, on the other hand, slows down and tends to slow down the other vehicles in the platoon). Because of the way the index is formulated, mid-block traffic reduces the index value by increasing the total downstream approach volume with respect to the through volume from the upstream intersection.

In summary, from all the different cases, it can be observed that when the index value is over 0.4, the coordinated mode of operation is always beneficial. For index values between 0.3 and 0.4, both modes of operation perform similarly. At lower index values, an isolated mode of operation appears to be generally beneficial. It should also be noted that the above analysis was only qualitative and not quantitative. In many cases, the difference between the two modes of operation is very small.

The index, as formulated in this study, aims to capture the effect of both platoon dispersion as well as the predominance of through movement (or any other movement that is coordinated); however, there are several other factors that determine the desirability of coordination. For example, a link with several mid-block access points and a high number of mid-block entries and exits may not be a good candidate for coordination. The index only captures the net movement in or out of the link but does not account for both mid-block entries and exits. Engineering judgement should be used in such cases.

The minimum green requirements for pedestrian and other extraneous considerations apart from vehicular volume alone may also influence the selection of the mode of operation. The least delay occurs when Greenshield's minimum delay cycle is provided; however, due to extraneous considerations, it may be necessary to provide a much larger cycle length. A larger cycle length would lead to higher delays. Whenever the cycle lengths at adjacent intersections are similar due to traffic requirements or otherwise, it may be prudent to coordinate the predominant movement as much as possible, even though the index may indicate that isolated mode is more beneficial.

Although the coordinated mode of operation is generally found to be beneficial for index values above 0.4, in many cases, it is also beneficial at lower values. A threshold of 0.4 is conservative. At lower values, between 0.4 and 0.30 (sometimes lower), the decision should be based on engineering judgement and field conditions.

The interconnection desirability index discussed above is an empirical relationship that attempts to capture the main factors that influence interconnection of intersections. As the

results show, it can only provide a conservative threshold for selecting between coordinated or isolated modes of operation. A more analytical approach based on traffic theory may produce a better means of selecting between the two modes of operation. It may be more appropriate to have a methodology built into coordinated mode signal timing design tools to decide whether coordination is appropriate for the given traffic conditions. The following section explores the theory behind traffic signal coordination and the possibility of utilizing the theory to aid in deciding between coordinated and isolated modes of operation.

INTERCONNECTION DESIRABILITY: TRAFFIC THEORY REVISITED

The vehicular arrivals at a totally isolated intersection are random. The flow rates during the green period are the same as those during red. On the other hand, if two adjacent intersections are within close proximity of each other, the arrival rate at the downstream intersection is dependent on the signal control at the upstream intersection.

If the progression between the two signals is good, most of the traffic will arrive at the downstream intersection during the green phase of the signal. This results in the average arrival rate during green at the downstream intersection being greater than during red. Poor interconnection could result in greater arrival rates during red than green.

The main objective of signal coordination is to ensure that the flow rate during green is greater than the flow rate during red (6). At an isolated intersection:

$$\frac{PVG}{PTG} = 1$$

where
PVG = Percent of volume that arrives during green, and
PTG = Percent of the cycle that is green.

Through proper signal timing coordination at the two intersections, it can be ensured that:

$$\frac{PVG}{PTG} > 1.$$

When this situation is achieved, more vehicles arrive during green which causes fewer stops and less delay than when vehicles arrive randomly. It should be noted that when intersections are closely spaced, non-coordination may result in the platoons from the upstream intersection arriving at the downstream intersection during red (PVG/PTG < 1), thus causing higher delays than at an isolated intersection with similar levels of traffic volumes.

33

The percentage of vehicles arriving in green at the downstream intersection is a function of several factors, including platoon length at the upstream intersection, platoon dispersion, percent of traffic that travels through (and hence can benefit from progression), and the green overlap for the platoon traffic at the downstream intersection. This is shown in the following relationship:

$$PVG_j = \left[PTT_j * \frac{GO_j}{LP_j} \right] + \left[(1-PTT_j) * \frac{RO_j}{(C-LP_j)} \right]$$

where

PVG_j	=	percent of vehicles arriving in green at j (downstream intersection);
PTT_j	=	percent of total through traffic arriving from i at j (thru traffic at i / total traffic at j);
GO_j	=	green overlap for the through traffic from i at j;
RO_j	=	green overlap for non-platoon traffic at downstream intersection $= G-GO_j$; and
LP_j	=	platoon length at downstream intersection, seconds.

The length of the platoon at the downstream intersection, LP_j, is a function of the platoon length at the upstream intersection and platoon dispersion factor. This is shown in the following relationship:

$$LP_j = LP_i * PD_{ij} + 0.8*(0.9 + 0.056t_{ij})$$

where

LP_i	=	platoon length at upstream intersection, seconds;
t_{ij}	=	travel time between intersection i and j, seconds;
PD_{ij}	=	platoon dispersion factor;
	=	$1.0 + (0.026 - 0.0014*NP) * t_{ij}$; and
NP	=	number of vehicles in the platoon at upstream intersection.

The length of the platoon at the upstream intersection, LP_i, is given by the following equation (7):

$$LP_i = g_0 + PVG * \frac{(g-g_0)^2}{g}$$

34

where

g_0 =   time required for queued vehicles to clear the intersection at i, sec; and

g =   effective green for the through movement at i.

As can be seen from the above equations, the computations related to progression and platoon dispersion are very complex and are not amenable to simple spreadsheet type calculations that can be easily performed by the practicing engineer. However, these same relationships are used in PASSER II to model the platoon flow and estimate the delay experienced by the progressed movements. Therefore, it is possible to evaluate the PVG/PTG value discussed above and automatically recommend to the user if it is found that signal coordination does not result in PVG/PTG values much greater than 1.

It should be noted that the PVG/PTG ratio may be small due to two reasons: 1) poor coordination, resulting in low green overlap for the platoon (GO_j) and hence a low PVG value, and 2) high platoon dispersion due to long distances between intersections leading to large lengths for the downstream platoon. It is possible to improve coordination by appropriately modifying the timing plan if green overlap is a problem; however, if the lack of progression is due to long distances between intersections and low traffic volumes, isolated mode of operation may be more appropriate.

CONCLUSIONS

An analysis of the interconnection desirability index as proposed by Chang et al. (5) was performed. Based on the deficiencies found, a new interconnection desirability index is proposed. The new interconnection desirability index is relatively simpler than the original interconnection desirability index.

An analysis of the modified interconnection desirability index using TRAF-NETSIM revealed that the coordinated mode of operation is more beneficial than the isolated mode for index values greater than 0.4. For index values between 0.3 and 0.4, both modes are equally desirable. Engineering judgement should be used when the index is within this range.

A review of the platoon progression modeling in PASSER II was performed to evaluate the feasibility of automatic determination of the desirability of interconnection. It was found that it is possible to determine the desirability of interconnection using the information generated within PASSER II. Updating PASSER II to perform this is not within the scope of this study. Further analysis should be performed, and automatic determination of interconnection desirability should be implemented within PASSER II.

CHAPTER III.
LOCATION OF SYSTEM DETECTORS

The objective of operating a closed-loop traffic signal system in a traffic responsive mode is to make sure that the system is operating with a timing plan that closely matches existing traffic conditions on the network. Most closed-loop traffic signal systems use system detectors to measure traffic conditions in the network. The master controller of the system uses volume and occupancy measurements from these detectors to compute traffic parameters for selecting appropriate timing plans. The master controller compares these parameters to thresholds established for each timing plan. The thresholds define the acceptable operating conditions for each timing plan. By continuously monitoring data from the system detectors, the master controller operating in a traffic responsive mode can respond to changing traffic conditions by implementing new timing plans when needed. Ideally, system detectors need to be located in a position to measure when and how traffic conditions are changing in the network.

This chapter examines where system detectors should be placed to measure changing traffic conditions in the network quickly and accurately. Traffic operations were simulated to test three different configurations of system detectors in a network. Each detector configuration was designed to detect specific types of changes in traffic conditions in the network. Recommendations for locating system detectors in a network were then derived from these simulation results.

BACKGROUND

In the early 1970s, the Federal Highway Administration (FHWA) provided the following general guidelines for locating system detectors in an Urban Traffic Control System (UTCS) operating in a traffic responsive mode (8):

- System detectors should be on all approaches leading to the critical intersection(s) in a network. Critical intersections are those intersections that involve the crossing of two or more links while carrying volumes that result in operation at or near saturation for substantial periods.

- Additional system detectors should be provided on every fourth link of an arterial and every third link for a grid network.

- System detectors should be in the critical lane only. If the critical lane varies by time-of-day, then multiple lanes should be detectorized, and time-of-day factors should be used to select the detector that best represents travel conditions during that period. The

37

critical lane is the lane carrying the largest traffic volumes and can be identified by observing queue lengths at intersections.

● System detectors should be outside the area of influence of adjacent signalized intersections. In other words, system detectors should be located far enough upstream of an intersection to be outside the area where standing queues form from the downstream intersection. They should also be far enough downstream to be outside the acceleration zone of vehicles leaving the upstream intersection. Figure III-1 provides general recommendations for placing system detectors on a link.

● System detectors should also be at least 15 meters downstream from a major traffic generator. A major generator is one that adds at least 40 vehicles per hour to the critical lane.

● System detectors should not be within areas that involve extensive weaving or other forms of unstable traffic flow.
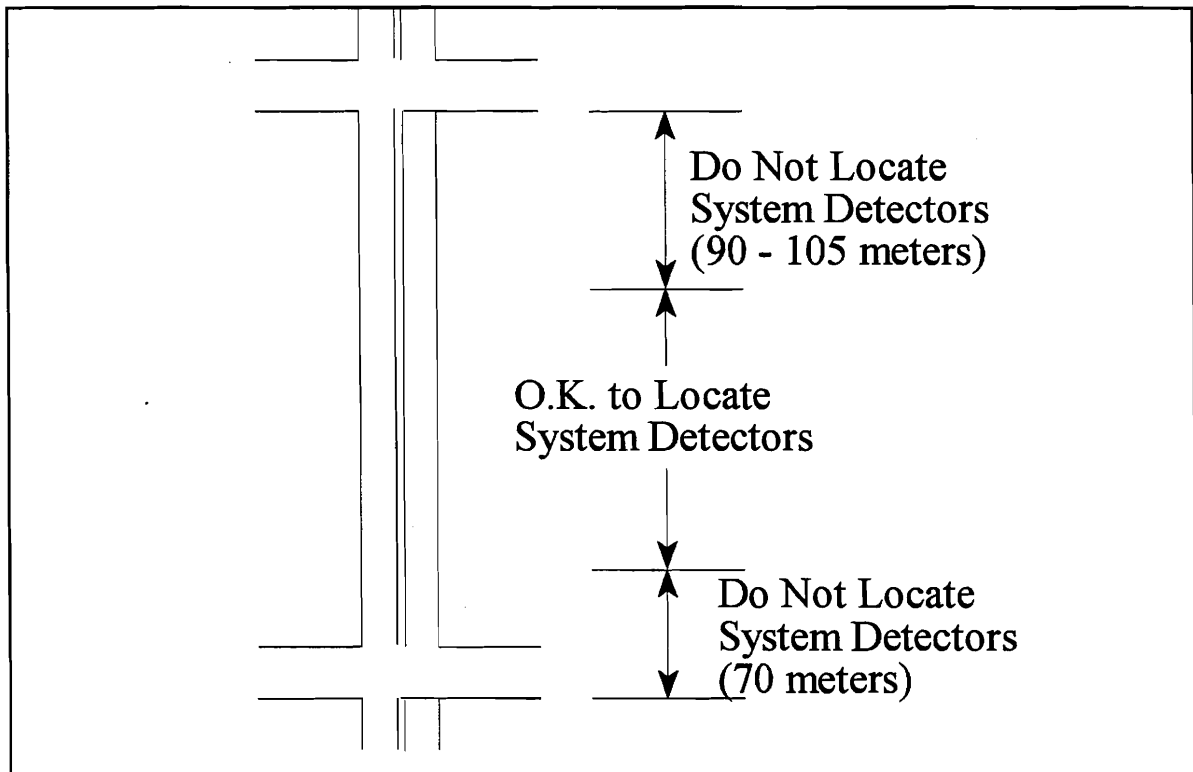
FIGURE III-1. Recommendations for Placing System Detectors on a Link

The Federal Highway Administration (FHWA) developed an additional procedure for determining where system detectors should be located in a network operated by an Urban Traffic Control System (UTCS). The procedure uses an Offset Benefit technique to determine whether or not a system detector should be provided at heavily traveled links in a network. Essentially, the Offset Benefit is the ratio of the time occupied by the progression band to the total green time available at a downstream intersection on a link. The numerator of the ratio estimates the time duration (in seconds) of platoons arriving at the downstream intersection while the denominator estimates the "green window" available for the platoon on a link. The ratio provides a measure of how much traffic arriving at an intersection is platooned. Links with the highest Offset Benefit are given priority of detectorization. The authors of this research suggest that all links with an Offset Benefit value of 0.75 or greater be detectorized.

A recent study conducted by TTI revealed that many operating agencies have found it more cost effective to install loop detectors at all feasible sites at the time of initial installation and then conduct a correlation study after the system is brought on line to determine which of the installed detectors best measure the changing traffic conditions in a network. In addition to providing flexibility at the initial time of installation, placing system detectors on all approaches provides redundancy in the event of detector failures. Furthermore, the extra detectors can be used to meet future traffic patterns (9).

This same TTI study provided a methodology for locating system detectors on an arterial street. Figure III-2 shows this methodology. Essentially, the methodology recommends that system detectors be located approximately every 800 meters on an arterial street to measure demand at points that are indicative of changing traffic conditions. In addition, system detectors should be placed at points of major cross-section changes, points upstream and downstream of major traffic generators, and points near intersecting major streets.

PROPOSED SYSTEM DETECTOR PLACEMENT

With most closed-loop signal systems, the user is required to assign the system detectors to each of the timing plan selection parameters. Often, there is a limit to the number of system detectors that can be assigned to each selection parameter. For example, the Naztec closed-loop system allows up to 10 system detectors to be assigned to each flow parameter (10), while the Econolite system allows only 4 detectors to be assigned to each of the selection parameters (11). Because of these restrictions, the engineer needs to be careful about placing the system detectors where they can measure the prevailing traffic conditions in the system. As a general guideline, the system detectors need to be located throughout the system where they can best detect the following changes in traffic conditions:

- increases or decreases in overall demand levels that might require modifying the cycle length for the system,

39
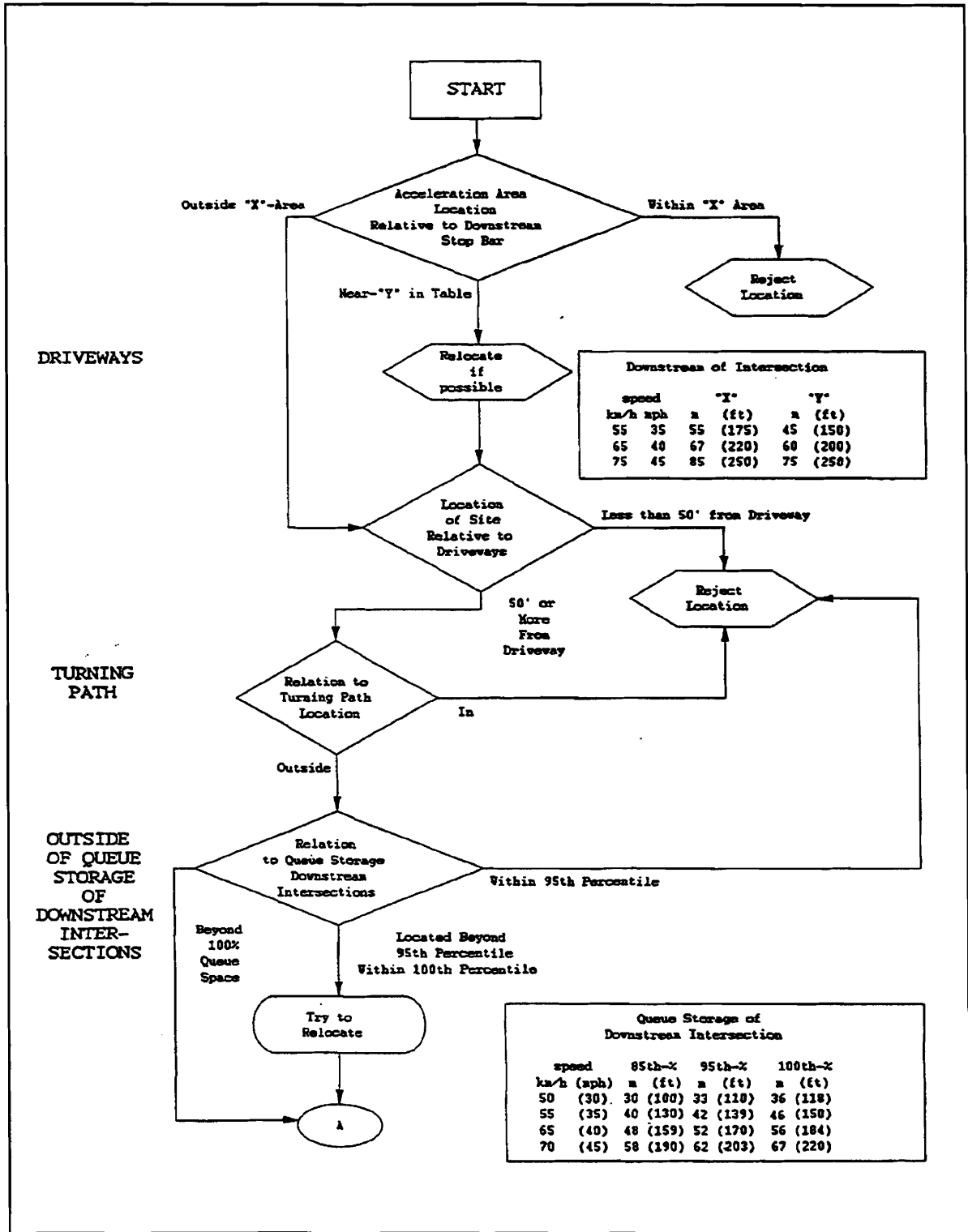
FIGURE III-2. Wood's Procedure for Identifying Strategic Locations for System Detectors

FIGURE III-2. Wood's Procedure for Identifying Strategic Locations for System Detectors (Cont.)
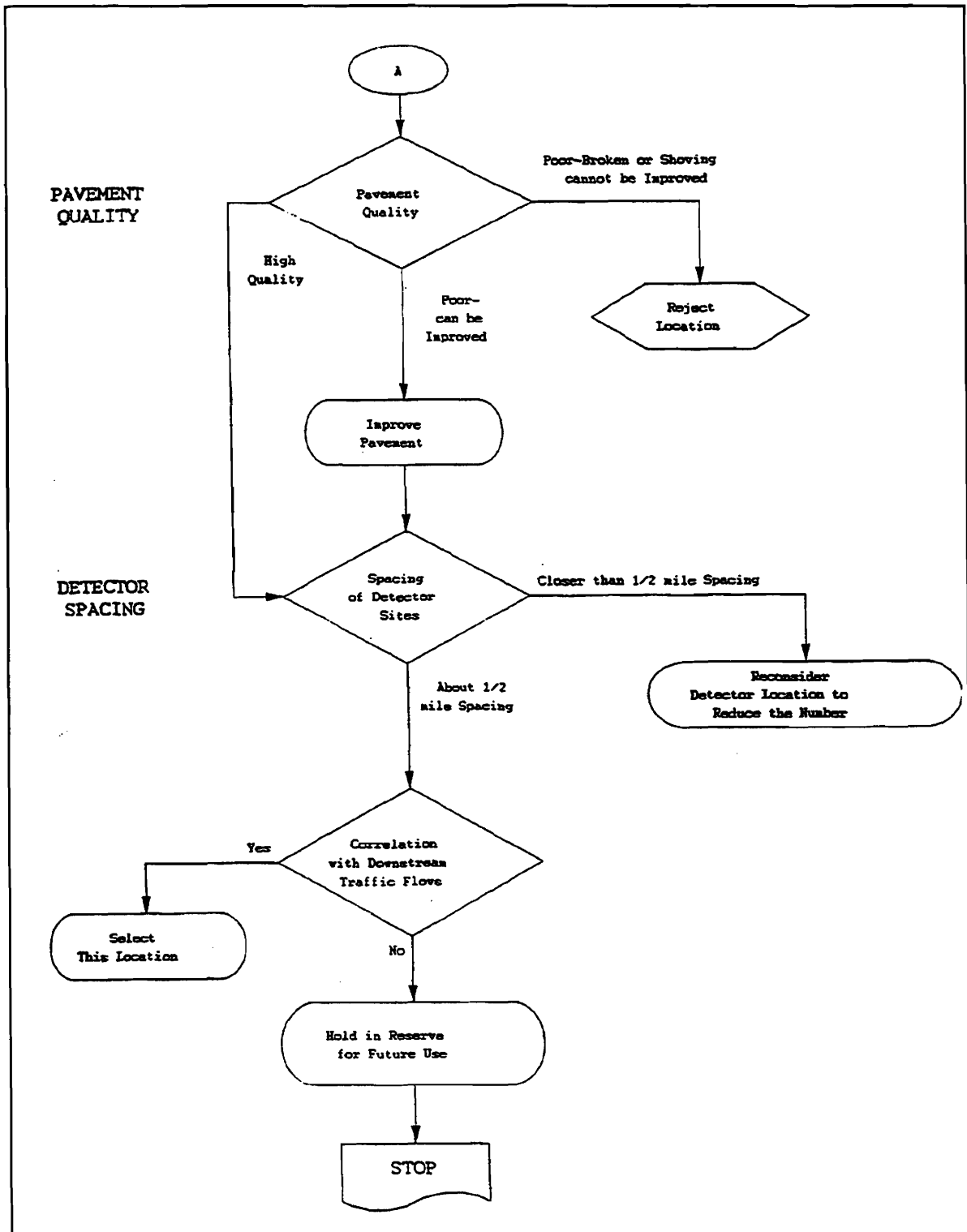
41

- shifts in directional demand that might require different offset plans, and

- changes in cross-street directional demand that might require different split plans.

Because all signals in a coordinated system are required to operate on the same cycle length, there is usually one intersection in any system that dictates the cycle length for the remaining intersections in the system. This intersection is typically called the critical intersection. It generally experiences the greatest demands and is most likely to become congested first in the system. Because different intersections can become critical (i.e., control the timings of the other intersections) at different times during the day, it is recommended that system detectors be placed on all approaches to the critical intersections in the system. These detectors need to be assigned to those traffic parameters that are responsible for determining the cycle length in a traffic responsive mode.

System detectors are also needed to measure changes in the directional distribution of traffic in the system. These system detectors need to be assigned to the traffic parameters that are responsible for selecting the offset in a traffic responsive mode. As a general guideline, these system detectors need to be located as close to the source of the directional change as possible. For example, if the source of a change in the directional distribution occurs outside the limits of the system (i.e., from traffic entering at the ends of the system), then the system detectors should be placed at the ends of the system; however, if the source of change in directional distribution occurs at a location within the system boundaries (i.e., from turning traffic entering the system from an internal signalized intersection or driveway), system detectors need to be located near this location to measure this change in demand. In these situations, the system detectors should be located on the main-street downstream of where the traffic is entering the system.

Some closed-loop systems permit different split plans to be implemented based on a comparison of cross-street traffic to main-street demand; therefore, system detectors need to be assigned that measure the cross-street demands in the system. As in the cycle length, there is generally one intersection that is critical in terms of the amount of time that must be provided to the cross-street. Generally, this is the same intersection that dictates the cycle length requirements for the system. The same system detectors that are used to measure changes in cycle length can also be used to measure changes in split requirements, except that the detectors on the side-street need to be assigned to the cross-street selection parameter while the detectors on the primary street need to be assigned to the main-street selection parameter.

SIMULATION STUDY

The objective of the simulation study was to evaluate different potential configurations of system detectors that could be used when operating a signal system in a traffic responsive mode. Three different detector configurations were tested using TRAF-NETSIM. The following sections describe the methodology and results of the simulation studies.

Overview of Methodology

To model the effects of different detector configurations on traffic responsive operations, an arterial study site was first chosen. Traffic turning movement data and corresponding timing plans were obtained from the study site and coded into the TRAF-NETSIM software traffic simulation package for study. The software package was calibrated to represent traffic conditions observed at the study site. Three configurations of system detectors were chosen and coded into the simulation software at various locations for study. Traffic loadings and their corresponding time-of-day timing plans were then simulated with the software package, and timing plan selection thresholds were calibrated for each detector configuration based on volume and occupancy data obtained from the simulated detectors.

Having established operating thresholds for various timing plans, different traffic conditions were simulated with the software. As each of the simulations progressed, timing plan changes were made based on detector data obtained from the software and the established thresholds. Based on several simulated operational periods, cumulative system delay information generated by the software was used to evaluate the effectiveness of the detector configurations in operating traffic responsive mode.

Measures of Effectiveness

The performance of the different detector configurations was evaluated using two measures of effectiveness. The primary measure of effectiveness was total system delay. A detector configuration was judged to be superior if it resulted in the implementation of timing plans that minimized the amount of delay that occurred in the network. The second measure of effectiveness used in the simulation study was the number of system detectors required. Because of maintenance requirements, a detector configuration that had the fewest number of system detectors required to operate the system while allowing the signal system to maintain an adequate level of operation in the responsive mode was judged to be superior.

43

Study Site

NASA Rd. 1 in Houston, Texas served as a test network for evaluating the different detector configurations. The portion of NASA Rd. 1 under investigation consisted of eight signalized intersections along an arterial street that was previously controlled by TxDOT's Arterial Traffic Management System, a first generation traffic responsive control technology. This site was chosen based on the availability of both traffic volume/turning movement data and the corresponding time-of-day operation signal timing plans. Figure III-3 depicts NASA Rd. 1.
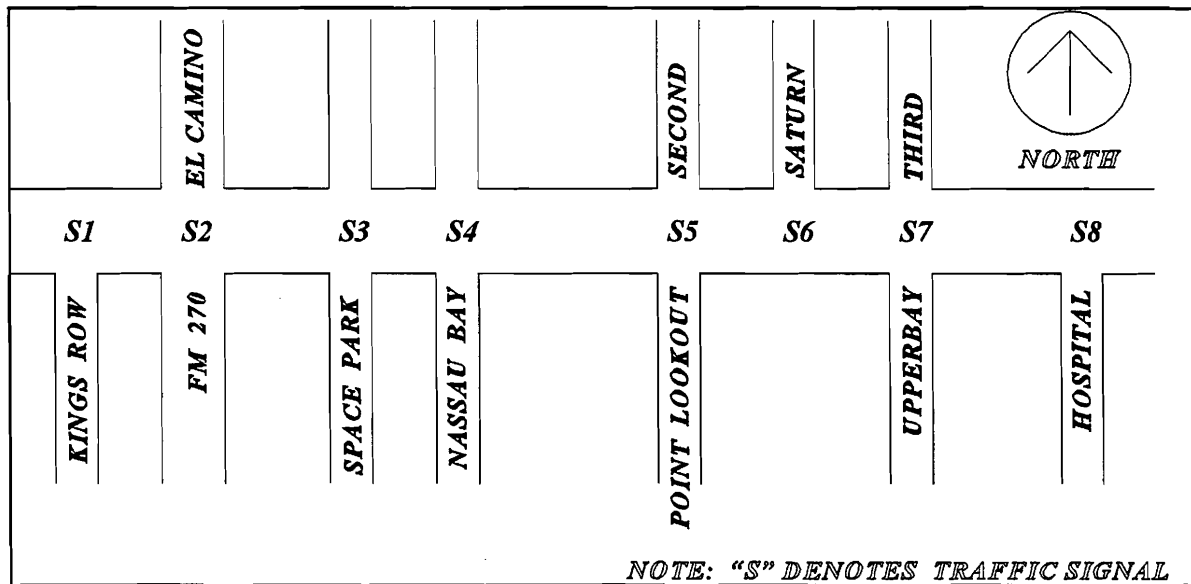
FIGURE III-3. Schematic of NASA Rd. 1 in Houston, TX

Detector Configurations

Three detector configurations were selected to be tested during the simulation study. Figures III-4, III-5, and III-6 show the three system detector configurations examined. With the first detector configuration (Detector Configuration #1), system detectors were placed only on each approach to the critical intersection (the intersection of NASA Rd. 1 and El Camino/FM 270). This configuration represented the minimum amount of system detectors that could be used to operate a system in a traffic responsive mode. With this configuration, all the system detectors were used to compute the cycle length parameter, while only those system detectors on the eastbound and westbound approaches to the critical intersection were used to determine the offset parameter.
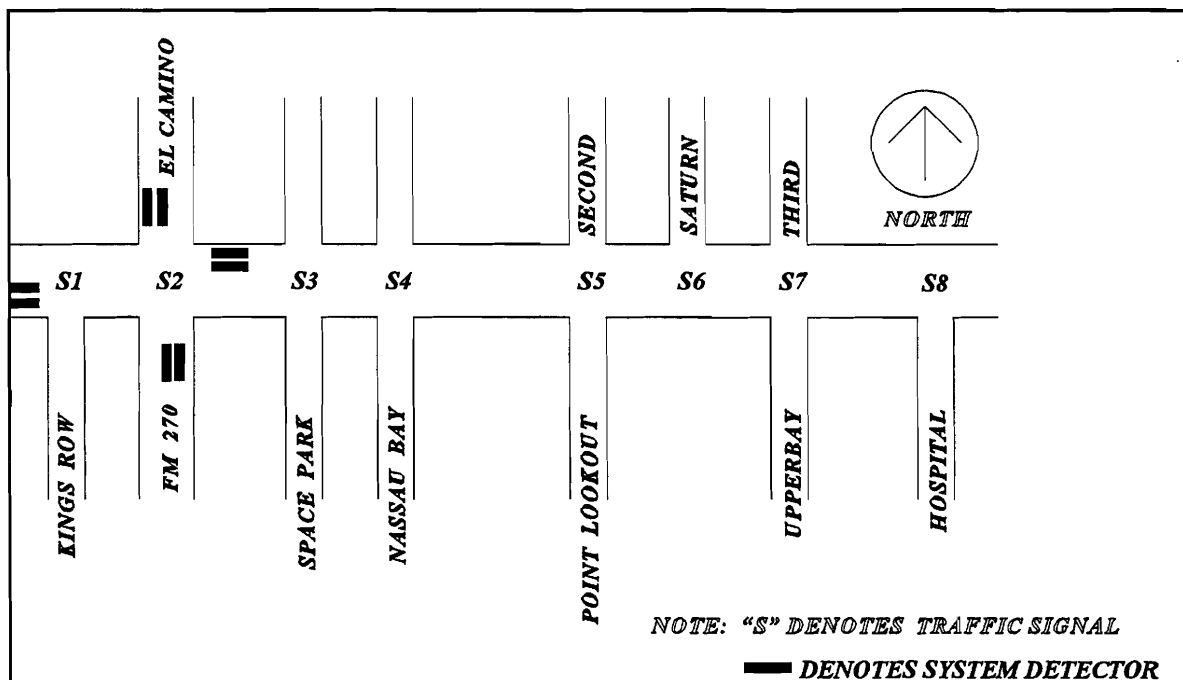
FIGURE III-4. Location of System Detectors in Detector Configuration #1
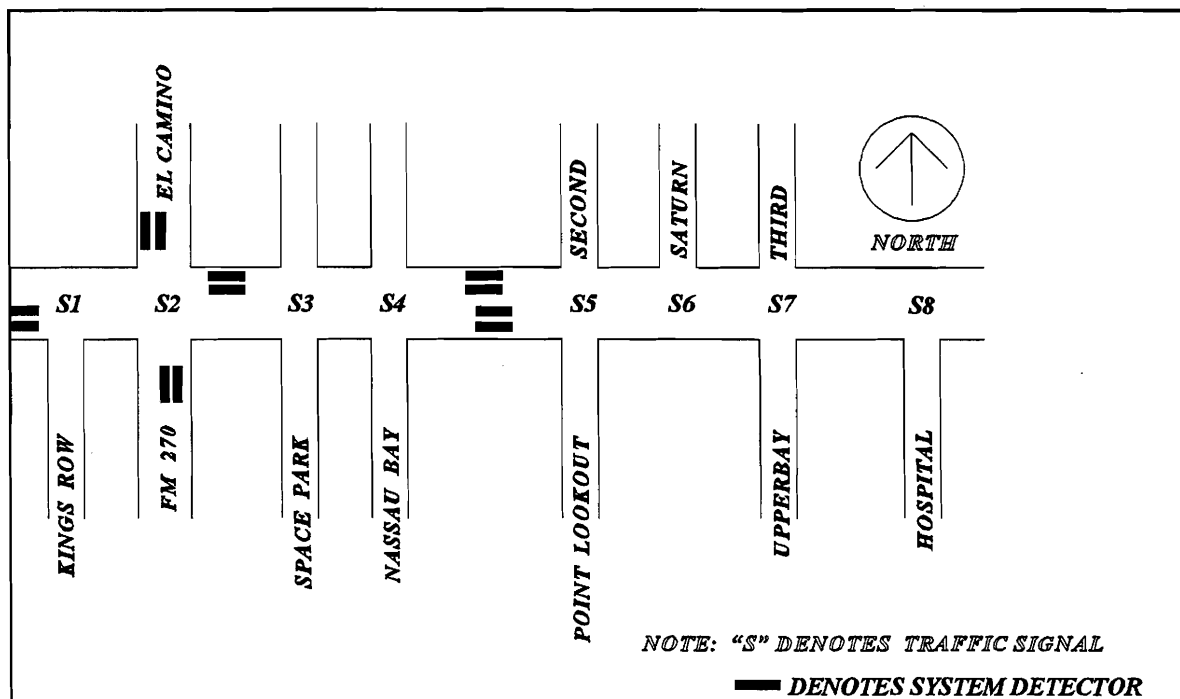
FIGURE III-5. Location of System Detectors in Detector Configuration #2
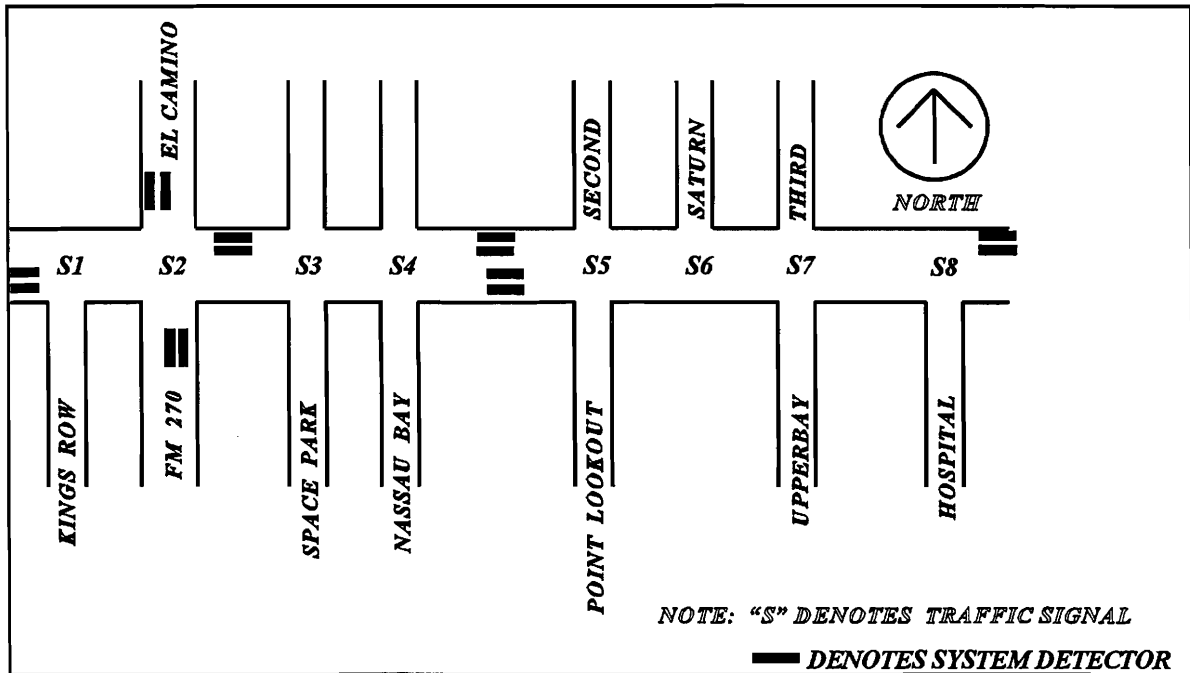
FIGURE III-6. Location of System Detectors in Detector Configuration #3

The second detector configuration (Detector Configuration #2), which is shown in Figure III-5, represented the current detector configuration used by TxDOT to operate the NASA Rd. 1 system in a traffic responsive mode. In this configuration, a total of six system detectors were used to select timing plans: four located around the critical intersection in the system (i.e., the intersection of NASA Rd. 1 and El Camino/FM 270) and two additional system detectors east of the critical intersection.

With the third detector configuration (Detector Configuration #3), additional system detectors were placed at the east end of the system. These detectors were used to measure directional demand entering the system from the east. Figure III-6 illustrates the placement of the system detectors in this configuration. This detector configuration roughly conforms to the detector configuration that would be developed using Wood's procedure.

Model Calibration

Before the actual simulations could be performed, the traffic model had to be calibrated. Using the animation package built into the TRAF-NETSIM package and knowledge of typical operations gained during site visits, it was recognized that the model produced abnormally long queues at some intersections. The network was observed to be saturating early into the operational periods, and none of the detector configurations were performing satisfactorily.

operational periods, and none of the detector configurations were performing satisfactorily. Upon investigation of the model outputs, it was determined that the default saturation flow rate was 1636 vehicles per hour, based on a mean queue discharge headway of 2.2 seconds. In order to rectify this situation, the default mean queue discharge headway was changed from 2.2 seconds to 1.9 seconds (equivalent to a saturation flow rate of approximately 1895 vehicles per hour) to more accurately reflect the 1994 *Highway Capacity Manual*'s ideal saturation flow rate of 1900 passenger vehicles per hour. A default vehicle mix consisting of all passenger cars was used, as were the default discharge characteristics. After completing this modification, further runs of the software and the animation package revealed that the model was producing traffic conditions that were more representative of those experienced during actual system operations.

Establishing Timing Plan Selection Thresholds

Before the detector configurations could be tested, cycle length and offset level selection parameters had to be established for each timing plan to be implemented by the system. With first generation traffic responsive systems, cycle length and offset level selection parameters are used to select appropriate timing plans from a library of plans. In order for the system to be able to identify an appropriate timing plan, detector data in the form of volume and occupancy measurements are manipulated to generate cycle length and offset level parameters. These parameters are then compared with threshold values to select a timing plan that corresponds to the measured traffic conditions. Two thresholds were established for each cycle length to denote the range of parameter values that resulted in selection of a given cycle length. Similarly, threshold values were established to correspond to three offset conditions: inbound, outbound, and average traffic demand. Once the cycle length has been chosen, these offset level thresholds allow for selection of a timing plan that corresponds to the traffic demand.

ATM Timing Plan Selection Parameters

This study was performed using TxDOT's Arterial Traffic Management (ATM) control logic. Because the ATM system was initially in operation on NASA Rd. 1 when the project began, it was decided to use the ATM algorithms in this evaluation for computing the timing plan selection parameters.

In operating the ATM in a traffic responsive mode, volume and occupancy data from each system detector are used to first compute a detector parameter (DP). The detector parameter consists of the sum of the volume and weighted occupancy measure from each detector. Two sets of detector parameters are computed with the ATM system: one for cycle length, split and offset selection (I=1), and the other for offset level selection (I=2). The formula used to compute the detector parameter is as follows:

47

$$DP_{ij} = \frac{(Vol_j + W_i * Occ_j)}{L}$$

where

$DP_{i,j}$ = the value of the detector parameter for the jth system detector,

Vol_j = the one minute average of the volume measurement from the jth system detector,

Occ_j = the one minute average of the occupancy measurement from the jth system detector,

W_i = the weighting factor for the two sets of parameters, and

L = the number of lanes covered by the system detector.

For the purposes of this study, the occupancy weighting factor for both the cycle length and the offset selection parameter was set to 0.5.

In the ATM system, the cycle length selection parameter (CLSP) is computed by averaging the detector parameter values from the corresponding system detectors assigned to the parameter in each configuration. Table III-1 shows the system detectors assigned to the cycle length selection parameter for each detector configuration.

For the offset level selection parameter (OLSP), two parameters are used in the ATM system: an inbound offset level selection parameter (Inbound) and an outbound offset level parameter (Outbound). As shown in Table III-2, different system detectors in each detector configuration were assigned to these two parameters. The actual offset level selection parameters were computed using the following equation:

$$OLSP = \frac{Inbound}{Inbound + Outbound} * 100.$$

Establishing Thresholds

To establish the thresholds, TRAF-NETSIM was used to generate volume and occupancy measurements at each system detector. Based on the traffic count data provided by TxDOT and the corresponding time-of-day operational plan, ten timing plans were identified for which data were available for simulating traffic operations. These ten plans represented those used by TxDOT over a continuous 11 hour (7:00 AM to 6:00 PM) operational period in time-of-day mode and were assumed to be sufficient to satisfy the traffic demand conditions.

48

As shown in Table III-3, the timing plans encompassed 100, 110, 120, and 140 second cycle lengths.

TABLE III-1. System Detectors from Each Detector Configuration Assigned to the Cycle Length Selection Parameter

Detector Configuration	System Detectors																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
#1	x	x	x	x	x	x	x	x												
#2	x	x					x	x					x	x						
#3	x	x	x	x	x	x	x	x	x	x			x	x					x	x

Note: An "x" denotes those system detectors assigned to the cycle level selection parameter. A shaded box indicates that the system detector was <u>not</u> used in computing the selection parameter.

TABLE III-2. System Detectors from Each Detector Configuration Assigned to the Offset Level Selection Parameter

Detector Configuration	System Detectors																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
#1	I	I					O	O												
#2	I	I	O	O	I	I	O	O												
#3	I	I					O	O	O	O			I	I					O	O

Note: "I" denotes those system detectors assigned to the inbound offset level selection parameter, while "O" denotes those system detectors assigned to the outbound offset level selection parameter. A shaded box indicates that the system detector was <u>not</u> used in the selection parameter.

49

TABLE III-3. TxDOT Time-of-Day Timing Plans Used to Calibrate Thresholds

Cycle Length	Offset Plan		
	Inbound	Average	Outbound
100	X	X	Not Used
110	X	X	Not Used
120	X	X	X
140	X	X	X

Each period was coded into a TRAF-NETSIM file that contained traffic volumes that corresponded to the given hour of the day. For example, the 11:00 AM - 12:00 PM traffic volumes were coded into TRAF-NETSIM and run for 15 minutes under the time-of-day timing plan that TxDOT developed for this time period (120 second cycle length with an outbound offset plan). Five randomly seeded replications of 15 minute durations were made of each of the cycle length and offset plan combinations bringing the total number of TRAF-NETSIM runs to 50. The volume and occupancy data for these 50 runs served as the basis for calibrating the selection parameters for each detector configuration.

During the course of running the TRAF-NETSIM files, it was found that excessive spillback occurred during both the 100 second cycle simulations and the 140 second inbound simulation. Spillback occurs when traffic fills a given link within a network. While this was not considered a problem for the 140 second cycle as 140 seconds was the maximum cycle length operated, it became a serious concern for the 100 second cycle. The 100 second cycle was felt to represent more of a concern because it would be operated in low volume conditions and would not be the subject of such spillback situations. In order to evaluate the appropriateness of the 100 second cycle timing plan for the given traffic conditions, the volume and phasing data were entered into PASSER II-90 Version 2. Based on the analysis of PASSER II-90, it was determined that the traffic conditions did not reflect those that would be best served by a 100 second cycle. PASSER II-90 indicated that the optimal cycle length should be at least 110 seconds for the given volume and phasing conditions. As a result, it was decided to proceed through the calibration process without using the 100 second cycle timing plans.

Having computed the selection parameters for each cycle and offset combination, threshold values were calculated. These values represent the ranges within which a cycle of given length and offsets is operated. Currently, no particular technique by which to select these thresholds has been established and proven. It was originally conceived that the thresholds could be established for the purposes of this research by constructing a 95% confidence interval

50

about the means of each of the cycle length and offset conditions. For example, the cycle length selection parameters corresponding to a 120 second cycle would be aggregated, and a 95% confidence interval would be developed about them. In this manner, the bounds of the confidence intervals would be used to delineate the range within which a 120 second cycle would be chosen. A similar process was envisioned to calibrate the offset thresholds. Unfortunately, it was found that the confidence interval created about the means of each cycle length and offset level was too narrow and that many of the data points were not encompassed by them. As a result, the process of calibration through use of a 95% confidence interval about the mean was abandoned. A second attempt to establish the thresholds by using the boundaries that encompassed all data points within two standard deviations (95% of the data points) of the mean was identified. This process resulted in threshold values of the cycle lengths that nearly encompassed each other. Because of the excessive overlap that resulted, this process was also deemed to be ineffective.

Finally, it was decided to establish thresholds through the use of the respective average upper and lower data points that resulted from each of the randomly seeded runs. The upper and lower data points for 15 randomly seeded runs (five simulations in the inbound direction, five in the average condition, and five in the outbound condition) were averaged to yield an upper and lower threshold boundary. In order to preserve consistency and avoid biasing the results of the research, the process of using the average value of the highest and lowest points from each run was used for each detector configuration in establishing both cycle length and offset thresholds.

Some modifications to the thresholds were made after the initial thresholds were reached. First, the offset selection parameters for Detector Configuration #2 did not allow the cutoff thresholds between outbound and average conditions to be clearly defined since the points for these two conditions directly overlapped each other. It was decided, therefore, to proceed using only the inbound and outbound conditions rather than including the average condition. This decision was made on the basis of visual inspection of the animation associated with these files and the resulting conclusion that the outbound condition would allow for better overall movement of traffic.

The second major area of modification involved the overlap of timing plans. Based on the threshold calibration process used, the lower limit of the 120 second cycle for both Detector Configuration #1 and Detector Configuration #3 did not overlap with the upper threshold of the 110 second cycle. This lack of overlap resulted in an undefined area where a computer operating the system would not be able to choose a cycle length if a parameter within this range was discovered. In order to rectify this situation, the threshold calibration data was reviewed, and a decision was made to reduce the lower thresholds of the 120 second cycle to 0.1 below the upper threshold of the 110 second cycle. This was felt to better represent the data than raising the 110 second cycle's upper limit. A similar change was made with the upper offset selection threshold for average conditions under Detector Configuration #1. That threshold was

51

raised so that outbound traffic conditions would not necessarily skip directly to the inbound conditions.

The third and final modification involved provisions for the system to drop into a 100 second cycle, even though calibration to that cycle length was not possible because of the aforementioned data problems. As the 100 second cycle was the shortest cycle length to be considered, the upper level threshold for it was set at 0.1 above the lower threshold for the 110 second cycle of each configuration.

Tables III-4 and III-5, respectively, show the cycle length and the offset selection thresholds developed for each detector configuration. These thresholds were used to determine when to make cycle length and offset changes for operating the NASA Rd. 1 system in a traffic responsive mode.

TABLE III-4. Cycle Length Selection Thresholds Calibrated for Each Detector Configuration

Cycle Length (seconds)	Detector Configuration #1		Detector Configuration #2		Detector Configuration #3	
	Lower	Upper	Lower	Upper	Lower	Upper
100	0.0	13.4	0.0	8.2	0.0	10.8
110	13.3	26.5	8.1	16.7	10.7	18.3
120	23.5	35.0	16.6	24.2	18.2	31.1
140	26.3	-	19.6	-	20.8	-

Evaluation of Detector Configuration Performance

In order to evaluate the performance of each of the three detector configurations, it was first necessary to determine when each configuration would detect changes in the traffic conditions within the arterial system. In practice, once the system detectors have measured a change in traffic volumes, an appropriate timing plan for the new condition is implemented. Thus, by determining the point in time when each detector configuration would institute a new timing plan, the manner in which signal timings would be operated over the course of a given period could be established. Having established when the signal timing plans would change, simulations of typical daily operations were made and corresponding measures of effectiveness examined.

TABLE III-5. Offset Level Selection Parameter Thresholds for Each Detector Configuration

Offset Level	Detector Configuration #1		Detector Configuration #2		Detector Configuration #3	
	Lower	Upper	Lower	Upper	Lower	Upper
Outbound	0.0	51	0.0	33	0.0	35
Average	-	-	24	34	34	48
Inbound	50	100	31	100	46	100

Using the existing time-of-day timing plans and the corresponding traffic volumes as a starting condition, TRAF-NETSIM files were developed to represent alterations in traffic conditions that were known to occur throughout the course of a typical operational period. Three time periods were selected to be simulated. Those time periods were from 7:00 AM to 8:30 AM, 9:00 AM to 10:30 AM, and 12:00 PM to 1:30 PM, encompassing the morning peak, a relatively low volume condition, and the noon hour peak traffic conditions. The 12:00 PM to 1:30 PM time frame was chosen specifically because of the decrease in operating cycle lengths indicated in the time-of-day operational plan. For each new file, the simulation began with the timing plan that would be in effect according to the time-of-day operational schedule.

Each new file was executed in TRAF-NETSIM, and the detector data generated was isolated for evaluation through the use of Statisical Analysis Software (SAS). The isolated detector data was analyzed to determine when the first change in timing plans was required by comparing the detection parameters generated to the established threshold values for the given detector configuration.

Through an iterative process, the previously coded file was modified to include the appropriate timing plans at the point in time at which the detector parameters indicated their use. Essentially, the TRAF-NETSIM file was run, and the detector data was analyzed to determine when timing plan changes were first required and what timing plans were chosen by the given detector configuration. The appropriate timing plans were then implemented at that point in time, and the files were rerun to determine when the next changes in timing plans would be required. This process was repeated until each operational period was simulated. For each of the three detector configurations, the same five randomly seeded TRAF-NETSIM files were subjected to this iterative process so that each detector configuration was tested under the same five traffic demand conditions.

To keep the system from changing timing plans while the system was transitioning to the new timing plan, the cycle length and offset selection parameters were ignored from the two control periods immediately after the point at which the timing plan was first changed. This two minute period was deemed to be a signal transition interval. Furthermore, either the cycle length or offset selection parameters (or both) had to exceed their respective thresholds for two consecutive iterations. In essence, because of this transition policy, once a new timing plan was initiated, a minimum of four minutes had to transpire before another change could be made.

As an illustration of the cycle and offset level selection process, consider what would occur if the system was operating a 100 second cycle under inbound conditions. If the new control period's cycle length selection parameter was greater than the upper threshold of the 100 second cycle (for illustrative purposes, in the range of a 110 second cycle) for two consecutive iterations, the cycle length would shift to a 110 second cycle length. If during the second control period, the cycle length selection parameter did not exceed the threshold or if it indicated an even higher cycle length than the first cycle length selection parameter (perhaps 120 seconds), no change would be made. In the latter case, the next cycle length selection parameter would be compared to see if, in fact, a transition should be made to the 120 second cycle.

During the period that cycle length was being chosen, a review of offset level was also underway. If the selection parameters indicated a need for both a new offset plan that favored a given directional movement and a cycle length change, both changes would be made simultaneously. However, if the need to change cycle length had been confirmed by two consecutive cycle length parameters and only one of the offset selection parameters indicated the need for a change in offset, only the change in cycle length would be made. If the offset selection parameters continued to indicate the need to change offset level, the minimum two control period transition phase would have to transpire before change would be implemented. Two consecutive offset selection parameters indicating the need for a different directional split would then be required before an offset level change could be made. This essentially means that a minimum of four minutes from the time when the new cycle length was implemented would have to pass before any changes were made in the offset level. The same conditional process was used if offset selection parameters indicated a change in offset before a change in cycle length had been confirmed by two consecutive cycle length selection parameters outside the current operating threshold range.

Early into the simulation process, it was determined that software limitations would restrict the length of the TRAF-NETSIM runs to 1.5 hours. TRAF-NETSIM allows a maximum of 19 timing intervals. Some of the detector configurations resulted in frequent timing plan changes, thereby surpassing the 19 interval limit. In addition, the global array within TRAF-NETSIM that stores vehicle data often became overloaded, causing the computer to stop the program after simulating a period slightly longer than 1.5 hours. A decision was thus made to stop the data collection process after intervals of 1.5 hours so that a uniform analysis period

could be used. In total, this process resulted in the development of 60 complete 1.5 hour TRAF-NETSIM files: five for each of the detector configurations and the time-of-day operational plan.

Data Analysis

The results of the TRAF-NETSIM runs were compared on an individual detector configuration basis. The total system delay that resulted under each of the three detector configurations over the course of the simulated operating period were used to evaluate the effectiveness of operations under the three detector configurations. Two statistical tests were conducted in evaluating the performance of the detector configurations and the time-of-day plan. Hartley's Test for Homogeneity of Population Variances was performed to insure that the assumption forming the basis of a one-way analysis of variance (ANOVA) was appropriate. An ANOVA was used to determine whether or not differences in sample means were statistically significant by comparing them to variations within samples (*12*).

SIMULATION RESULTS

Total System Delay

As mentioned above, the primary measure of effectiveness chosen for the purpose of evaluating detector configuration performance was total system delay. TRAF-NETSIM reported this measure, whose units are vehicle-hours, on a cumulative basis at one minute intervals. The total system delay data was obtained from each of the 60 runs at a point 1.5 hours after the beginning of the simulation.

Tables III-6, III-7, and III-8 show the total system delay observed for the three detector configurations in each of the simulation periods. Figure III-9 shows, by simulation period, the average total system delay resulting under each detector configuration. As evidenced by these tables and figure, there was a significant amount of variation among the delay data attributable to the different simulation runs; however, the differences among the means of each group was not particularly large. Statistical tests were performed to evaluate the differences among each of the detector configurations. The following sections discuss the results of these tests.

55

TABLE III-6. Total System Delay for Each Detector Configuration
(7:00 AM - 8:30 AM)

Detector Configuration	Total System Delay (Veh-Hrs)						
	Run #1	Run #2	Run #3	Run #4	Run #5	Average	St. Dev.
#1	889.87	947.62	607.28	798.76	638.50	776.41	150.26
#2	830.13	650.52	868.14	675.84	592.33	723.39	160.01
#3	988.79	598.55	632.97	796.93	845.38	772.52	119.47

TABLE III-7. Total System Delay for Each Detector Configuration
(9:00 AM - 10:30 AM)

Detector Configuration	Total System Delay (Veh-Hrs)						
	Run #1	Run #2	Run #3	Run #4	Run #5	Average	St. Dev.
#1	408.98	407.29	434.10	397.30	292.16	387.97	55.25
#2	452.66	485.81	381.21	397.20	450.09	433.39	43.10
#3	348.07	393.74	428.37	634.81	315.45	424.09	125.42

TABLE III-8. Total System Delay for Each Detector Configuration
(12:00 Noon - 1:30 PM)

Detector Configuration	Total System Delay (Veh-Hrs)						
	Run #1	Run #2	Run #3	Run #4	Run #5	Average	St. Dev.
#1	1168.58	1298.42	1046.63	1120.50	1174.88	1161.80	91.99
#2	1176.35	974.22	1149.92	987.18	1132.34	1084.00	95.70
#3	1356.14	1173.59	1001.67	1090.6	1132.34	1105.33	172.34
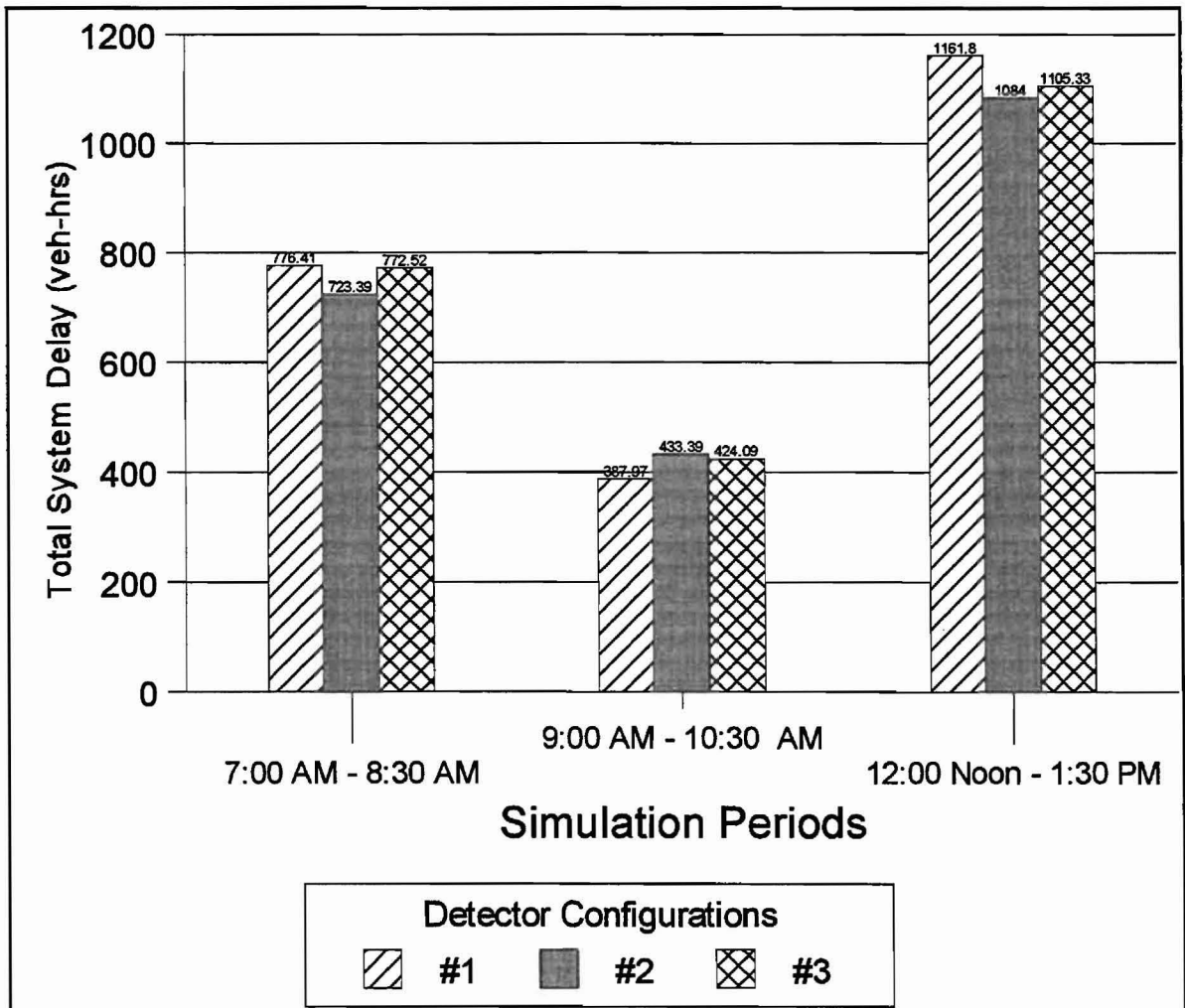
FIGURE III-7. Comparison of System Delay for Each Detector Configuration

Hartley's Test for Homogeneity of Population Variances

Before the total system delay for each detector configuration could be directly compared, it was first necessary to determine whether the variances of the samples were the same. To confirm the assumption of equal variance, a Hartley's Test for Homogeneity of Population Variances (Hartley's Test) was performed (*12*). For this test, the null hypothesis was that the variances of the samples for each detector configuration were equal (as shown below):

$$H_0: (\sigma_1)^2 = (\sigma_2)^2 = (\sigma_3)^2.$$

57

The alternate hypothesis was as follows:

H_a: At least one of the sample variances were not the same.

Acceptance of the null hypothesis would imply that there was no statistical difference among the sample variances and thus supports the assumption made in the ANOVA tests. Rejection of the null hypothesis would imply that there was a statistical difference among the sample variance and would have undermined the assumption made in the ANOVA tests. The analysis was performed using a 95% confidence level.

Table III-9 summarizes the results of Hartley's Test. The results of Hartley's test showed that there was not any statistical evidence to reject the hypothesis that equal variances in the data were equal. The results also indicated that an ANOVA could be performed on the data without violating the necessary assumption of equal variances between samples.

TABLE III-9. Results of Hartley's Test for Homogeneity of Population Variances

Test Statistic	7:00 AM - 8:30 AM	9:00 AM - 10:30 AM	12:00 Noon - 1:30 PM
F_{MAX}	1.79	8.47	3.51
F_{RR}, $\alpha=.05$	15.5	15.5	15.5
F_{MAX} vs. F_{RR}	1.79 <15.5	8.47 <15.5	3.51<15.5
Reject H_o?	No	No	No

Analysis of Variance

An ANOVA was also performed on the mean total system delays for each detector configuration. The null hypothesis for the ANOVA was that the mean total system delay for each detector configuration were equal. The alternative hypothesis was that the total system delay for at least one of the detector configurations was not equal to the rest. A 95% confidence level was used to test the equality of the total system delay of the different detector configurations. The tests were performed using Statistical Analysis System (SAS) software.

Tables III-10, III-11, and III-12 depict the ANOVA tables that correspond to each of the three operational periods. The results of the ANOVA showed that no statistical difference existed among the mean system delay incurred under the various detector configurations. The results of the ANOVA implied that, for all practical purposes, each of the detector

configurations resulted in timing plans being implemented that caused the same amount of total system delay.

Number of Timing Plan Changes

As described above, a minimum of two control periods (120 seconds) was considered to be the transition period after a new timing plan was implemented. After those two control periods passed, the control logic would permit the timing plan to change. In some cases, especially where the cycle length and offset selection parameters were close to the established thresholds, it was found that there was a tendency to oscillate between two timing plans. This continuous shifting of timing plans was considered to be a potential source of increased system delay.

While the minimum elapsed time between transitions was kept short to simulate TxDOT's current operational practice, there have been documented cases of increased delay associated with arterial signal transitions. One study of the UTCS first generation traffic responsive control systems found that, during increasing traffic demand conditions, the detrimental effects caused by transitioning between timing plans increases as the number of timing plan changes increases (13). Furthermore, the authors of the study stated that the percent increase in total network delay caused by transitioning increased with an increase in network saturation. It was reported that 30 minute periods between transitions is too short to allow the benefits of implementing new timing plans to be realized.

TABLE III-10. Results of ANOVA Test of Differences in Total System Delay for Each Detector Configuration (7:00 AM - 8:30 AM)

Source	Sum of Squares	Degrees of Freedom	Mean Square	F Test	Prob. >F
Between Samples	8732.51	2	4366.25	0.21	0.8137
Within Samples	249820.34	12	20818.36		
Totals	258552.85	19			

TABLE III-11. Results of ANOVA Test of Differences in Total System Delay for Each Detector Configuration (9:00 AM - 10:30 AM)

Source	Sum of Squares	Degrees of Freedom	Mean Square	F Test	Prob >F
Between Samples	5758.51	2	2879.25	0.42	0.6673
Within Samples	82563.86	12	6880.32		
Totals	88322.37	14			

TABLE III-12. Results of ANOVA Test of Differences in Total System Delay for Each Detector Configuration (12:00 Noon - 1:30 PM)

Source	Sum of Squares	Degrees of Freedom	Mean Square	F Test	Prob >F
Between Samples	16161.35	2	8080.67	0.51	0.6117
Within Samples	189288.30	12	15774.02		
Totals	205449.65	14			

Table III-13 summarizes the average number of timing plans that were implemented during each of the simulations compared to the number of timing plans that would be implemented if the system operated in a time-of-day mode. This table shows that Detector Configuration #3 resulted in the most timing plan changes under the two peak period conditions (7:00 AM and 12:00 PM operational periods). In contrast, Detector Configuration #1 consistently resulted in the lowest number of timing plan changes in response to detected traffic pattern changes. In lieu of the previously cited study suggesting that signal transitions be kept to a minimum, this would seem to indicate an advantage of Detector Configuration #1. Detector Configuration #2 tied Detector Configuration #1 for the minimum number of timing plan changes during the two peak periods.

**TABLE III-13. Average Number of Timing Plans Implemented During
Each Simulation Period**

Detector Configuration	Average Number of Timing Plans Implemented During Simulation Period		
	7:00 AM-8:30 AM	9:00 AM-10:30AM	12:00 Noon-1:30PM
#1	6	3	6
#2	6	9	6
#3	8	8	10
Time-of-Day	2	2	2

DISCUSSION AND INTERPRETATION OF RESULTS

One potential explanation for the lack of statistical differences in the sample means rests in the number of replications of the operational periods. Five random seeds were tested for each configuration. This number was chosen based upon time limitations imposed by the nature of TRAF-NETSIM. Because each TRAF-NETSIM file had to be rerun every time a new timing plan was introduced, it typically required three to four hours to simulate one complete hour of TRAF-NETSIM data. It is conceivable that further replications of the operational periods through the use of additional random seeds might have made a statistical difference among the various configurations tested apparent.

In addition, the traffic operations in the system were simulated for 1.5 hours only. It is conceivable that limiting the simulation period to 1.5 hours did not permit the full impact of operating the signal system in traffic responsive mode based with the different detector configurations to be realized. Had longer simulation periods been possible, more apparent differences among the detector configurations may have been realized.

It should also be noted that some of the threshold ranges for cycle length and offset level selection were extremely close. The lack of significant overlap between these threshold ranges was found to result in excessive oscillation between timing plans under certain traffic conditions. The problems were especially apparent when operating the 110 and 120 second cycle lengths using Detector Configuration #1 and Detector Configuration #3. Frequent timing plan changes were also observed to occur between the inbound and outbound offset conditions of Detector Configuration #2. These three problem areas were a direct result of artificially adjusting the thresholds in the cases where the calculated thresholds originally did not overlap.

61

CONCLUSIONS AND RECOMMENDATIONS

Based upon the statistical analysis of the delay data, one can conclude that there was no statistical difference among the performance of the three detector configurations in terms of when to implement timing plans that minimized the total amount of delay in the system. The performance of each of the detector configurations was the same for all practical purposes; therefore, other measures of effectiveness were considered. Since Detector Configuration #1 had the fewest number of system detectors, it was judged to be the best for this network.

In considering the potential applicability of using the Detector Configuration #1 for operations during low volume conditions, several considerations should be made. Paramount among these is that such a configuration may not work for all scenarios of traffic responsive signalization. If an arterial system's critical intersection shifted by time of day, there would be obvious limitations to the use of Detector Configuration #1. Also, it was noted in viewing the animation of the various files that Detector Configuration #1 did not always offer a rapid response to major traffic shifts. For example, consider the impact of a major employer releasing its employees from a location several intersections upstream of the critical intersection. It is conceivable that the control system would continue to run a timing plan that favored the opposite direction's movement until such time that enough vehicles from the employer's location reached the system detectors and thus established a need to shift timing plans. Similar large scale traffic movements downstream of the critical intersection that never even entered the critical intersection could conceivably occur and remain unnoted by the traffic responsive system.

Another shortcoming of Detector Configuration #1 involves the nature of operations during periods when one or more of the system detectors fail. In a system where there are multiple system detectors in use, redundancy might not be a critical issue; however, failure of one or more detectors when only a relatively small number of system detectors are in use could prove disastrous to operational efficiency. In the event of multiple detector failures, it is conceivable that a reversion to time-of-day operations would more satisfactorily supplement operations until such time that the defective detectors could be repaired or replaced.

CHAPTER IV.
TIME-OF-DAY VERSUS TRAFFIC RESPONSIVE MODE

Most closed-loop signal systems today are capable of being operated in either a time-of-day or traffic responsive mode. With time-of-day mode, traffic signal timing plans are automatically selected from a library of timing plans on a time-of-day and day-of-week basis. In a time-of-day mode, the timing plans are selected without regard to the current traffic conditions that exist in the network. Under traffic responsive mode, however, the signal system automatically selects a timing plan from a stored library of timing plans that provides the best control for the measured traffic conditions. Each timing plan in the library has a unique volume and occupancy signature. Recent volume and occupancy measurements from the system detectors are compared to the signatures of the stored timing plans. The timing plan that best matches the measured traffic conditions is then automatically implemented by the signal system.

Despite the added flexibility offered by operating signal systems in a traffic responsive mode, most closed-loop signal systems today operate in a time-of-day mode. In part, this is because comprehensive guidelines do not exist for determining when to operate signal systems in a time-of-day or a traffic responsive mode. An objective of this research project is to develop and test guidelines for identifying when to operate a signal system in time-of-day or traffic responsive mode.

PAST RESEARCH

Historically, traffic engineers use the traffic responsive mode in two different ways: 1) to pinpoint when time-of-day timing plan changes need to be made, and 2) to provide better control during atypical events (such as a sporting event, concert, or incident) that cause *major shifts* in traffic patterns in a control area. With the first application, the traffic responsive mode is used to monitor changing traffic patterns throughout the day and implement new timing plans as conditions warrant. As a result, the time that a specific timing plan is implemented can vary as traffic demands vary on the network, instead of being implemented at a specified time.

Research has shown that using traffic responsive mode to pinpoint when to implement time-of-day plans results in only marginal benefits over properly designed time-of-day mode (*14,15,16*). This is because minor fluctuations in traffic demand can cause frequent timing plan changes. Other research has shown that frequent timing plan changes over a short period can have a deleterious effect on the performance of a signal system (*13*). Frequent timing plan changes can actually impede traffic operations because of the transition that is required between timing plans. Therefore, the benefits achieved by implementing a new timing plan to pinpoint

when time-of-day changes occur may often be offset by the delays associated with transitioning between timing plans.

Another way that engineers have used the traffic responsive mode is to provide control when unexpected major shifts in traffic patterns occur in the network. Unexpected major shifts are usually caused by atypical events (such as sporting events, concerts, incident conditions, and holidays) occurring in or near the control network. Usually, the size of the change in traffic patterns associated with these events is known (e.g., the amount of traffic at a sporting event is dictated by the size of the sporting arena). What is often not known in many situations, however, is the exact time the traffic demand on the network will change. For example, although an engineer may know the exact time that a special event (like a football game or concert) begins, the exact time that the event ends varies. Because the exact ending time is not known, it is difficult to implement a time-of-day plan that can accommodate the demand from these events. Furthermore, the amount of traffic (and thus the duration of the increased demand) may vary from event to event. With a signal system in a traffic responsive mode, conditions in the control area can be monitored to detect when significant changes in traffic occur in the control area.

SIMULATION STUDIES

Simulation studies were conducted to quantify the benefits of operating a traffic signal system in a traffic responsive versus a time-of-day mode. The same network and methodology used to evaluate various system detector configurations was also used in this evaluation. This section summarizes the methodology used in this evaluation and presents the results of the analysis.

Study Site

Eight intersections on NASA Rd. 1 in Houston, Texas, served as the network used in this evaluation. The portion of NASA Rd. 1 under investigation will be placed under control of a closed-loop system in the near future. This site was chosen because of the availability of traffic volume/vehicle turning movement and signal timing data, and the existing presence of systems detectors located throughout the network.

Study Methodology

A process similar to the one used to examine the different configurations of system detectors was used to compare the performance of the traffic signal system on NASA Rd. 1 operating in a time-of-day mode to the performance of the traffic signal system operating in a

traffic responsive mode. As with the detector configuration, TRAF-NETSIM was used to simulate traffic conditions on NASA Rd. 1. The time-of-day plan that was used in this simulation was developed by TxDOT for use on NASA Rd. 1. As with the simulation studies to examine different system detector configurations, traffic conditions in three periods were simulated: 7:00 - 8:30 AM, 9:00-10:30 AM, and 12:00 Noon to 1:30 PM. A total of five simulation runs were performed with different random seed values with the signals operating in a time-of-day mode. The amount of total system delay that resulted when the signal system was operating in a time-of-day mode was compared to the amount of total system delay that occurred in the system when the traffic signals were operated in a traffic responsive mode.

Simulation Results

Table IV-1 shows the amount of total system delay that resulted from simulating the traffic condition on NASA Rd. 1 with the traffic signal system operating in a time-of-day mode during each of the three simulation periods. The table also shows the average total system delay and the standard deviation of the total system delay for each simulation period.

TABLE IV-1. Total System Delay on NASA Rd. 1 Under Time-of-Day Control

Time Period	Total System Delay (Veh-Hrs)						
	1	2	3	4	5	Average	St. Dev.
7:00 - 8:30 AM	858.50	816.64	607.28	767.03	684.10	746.71	101.44
9:00-10:30 AM	465.02	425.92	418.56	439.36	300.65	409.90	63.59
12:00 Noon - 1:30 PM	1418.59	1245.04	1042.00	804.68	807.74	1063.61	270.14

As discussed in Chapter III, a simulation study was performed to examine the effects of different system detector configurations on the ability of the system to operate in a traffic responsive mode. The study showed that there was no statistical difference in the amount of total system delay that occurred under any detector configuration with the signal system operating in a traffic responsive mode. As a result of this finding, the amount of total system delay that was generated using each of the system detector configurations was averaged across all the detector configurations to produce a total system delay that occurred during each period on NASA Rd. 1 under traffic responsive control.

Figure IV-1 compares the average total system delay of the signals operating in a traffic responsive mode to the average total system delay of the signals operating in a time-of-day
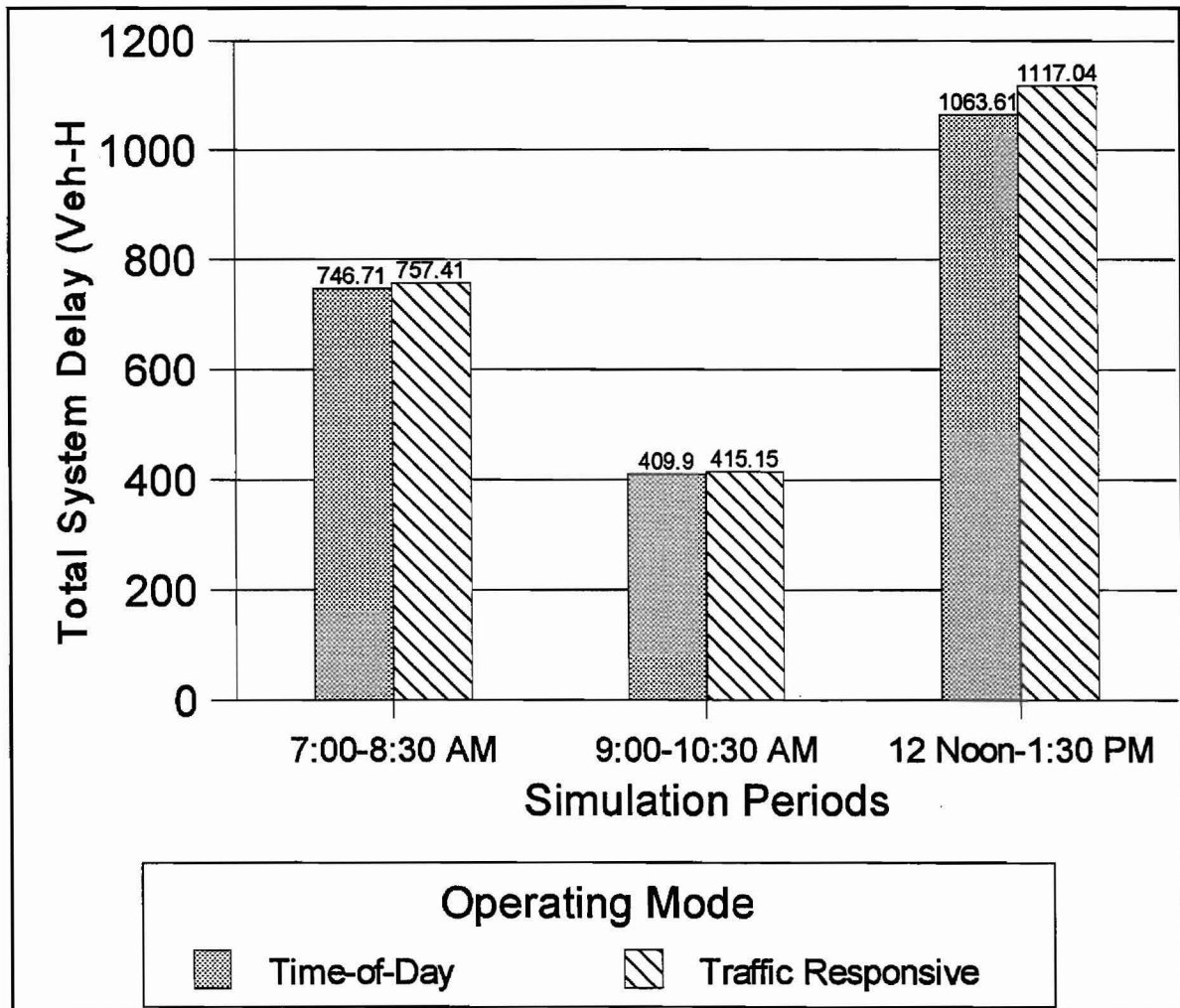
FIGURE IV-1. Total System Delay Occurring with Different Signal Operating Modes

mode for each of the three evaluation periods. From this figure, it can be seen that total system delay is slightly higher when the signal system was simulated as operating in a traffic responsive mode than when the signal system was simulated as operating in a time-of-day mode. In fact, operating the traffic signal system in a traffic responsive mode produced, on the average, total system delays that were 1.43%, 1.28% and 5.02% higher than if the signal system was operated in a time-of-day mode in each of the three time periods, respectively.

A standard t-test was used to compare the total system delays of the two operating modes in each period. Tables IV-2, IV-3, and IV-4 show the results of these analyses. Because total system delay was used in the comparison of the detector configurations in a previous analysis, the t-score defining the rejection region was adjusted to guard against falsely rejecting the null hypothesis while still maintaining a 95% confidence level.

TABLE IV-2. Results of T-Test Comparing Time-of-Day and Traffic Responsive Mode During the 7:00 - 8:30 AM Simulation Period

Operating Mode	N	Mean	St. Dev.	T	DF	Significant ?
Time-of-Day	5	746.71	101.44	-0.1607*	18.0	No
Traffic Responsive	15	757.41	135.87			

* Compared to Critical $T_{0.0125, 18}$ = 2.4768

TABLE IV-3. Results of T-Test Comparing Time-of-Day and Traffic Responsive Mode During the 9:00 - 10:30 AM Simulation Period

Operating Mode	N	Mean	St. Dev.	T	DF	Significant ?
Time-of-Day	5	409.70	63.59	-0.1334*	18.0	No
Traffic Responsive	15	415.15	79.43			

* Compared to Critical $T_{0.0125, 18}$ = 2.4768

TABLE IV-4. Results of T-Test Comparing Time-of-Day and Traffic Responsive Mode During the 12:00 Noon - 1:30 PM Simulation Period

Operating Mode	N	Mean	St. Dev.	T	DF	Significant ?
Time-of-Day	5	1063.61	270.14	-0.4282*	4.5	No
Traffic Responsive	15	1117.04	121.14			

* Compared to Critical $T_{0.0125, 4.5}$ = 3.409

Discussion of Results

The results of the comparison of means showed that there was no statistical difference in the amount of total system delay generated when the traffic signal system was simulated as operating in a traffic responsive mode and when the traffic signal system was simulated as operating in a time-of-day mode. This finding agrees with previous research evaluating the benefits of traffic responsive mode as compared to time-of-day control. These previous studies showed between 0.2% to 2.8% delay savings was achieved by operating the system in a traffic responsive mode over a time-of-day mode (*14,15,16*).

One possible explanation of why there was no significant difference in the amount of total system delay between the two operating modes has to do with the number of timing plan changes that occurred when the traffic signal system was operated in a traffic responsive mode. Table IV-5 shows the average number of timing plans that were implemented during each period using the different operating modes. This table shows that under the traffic responsive mode, over 3 times as many timing plans were implemented than if the signal system was operated in a time-of-day mode. While changing timing plans to match traffic conditions is the objective of traffic responsive control, frequent timing plan changes can also have a detrimental impact on traffic operations in a system. Whenever the timings in a system are changed, there is a period during which the signals transition from the old timing plan to the new timing plan. During the transition phase, coordination on the main stream can be lost, and excess delays on the cross-streets is common (*15*). These impacts are particularly acute when a cycle length and/or offset changes are required between two timing plans. Because of this, there is the potential that the additional delays that were incurred during the transition to the new timing plan offset any benefit that may have resulted from changing timing plans to more closely meet traffic conditions.

**TABLE IV-5. Average Number of Timing Plans Implemented During
Each Simulation Period by Operating Mode**

Operating Mode	7:00-8:30 AM	9:00-10:30 AM	12:00 Noon - 1:30 PM
Time-of-Day Mode	2	2	2
Traffic Responsive Mode	6.7	6.7	7.3

This appears to be the case in this simulation study. The majority of the timing plan changes that were implemented during the study involved either an increase or a decrease in cycle lengths. To simulate existing transition methods, changes in timing plans were forced to

occur over two cycles. As a result of the transition between timing plans, coordination was lost, and the phasing was less than optimum for at least two cycles at some intersections. As a result, any reductions in system delay that may have been achieved by changing timing plans was offset by the additional delay incurred during the transition phase.

In addition to the added delay caused by the transitions between timing plans, another possible explanation of why the traffic responsive mode did not perform better than the time-of-day mode was that the traffic responsive plan was trying to implement cycle lengths that were too close to one another to result in significant delay savings. Research has shown that delays are relatively insensitive to changes in cycle length (*17*). As illustrated in Figure IV-2, Webster showed that delay at an isolated intersection is never more than 10 to 20% when the cycle length is within a range of 0.75 to 1.5 times the optimum cycle length. This implies that significant delay savings (i.e., greater than 10 to 20%) cannot be achieved unless a new cycle length is outside the range of 0.75 to 1.5 times the optimum cycle length of the previous traffic conditions.
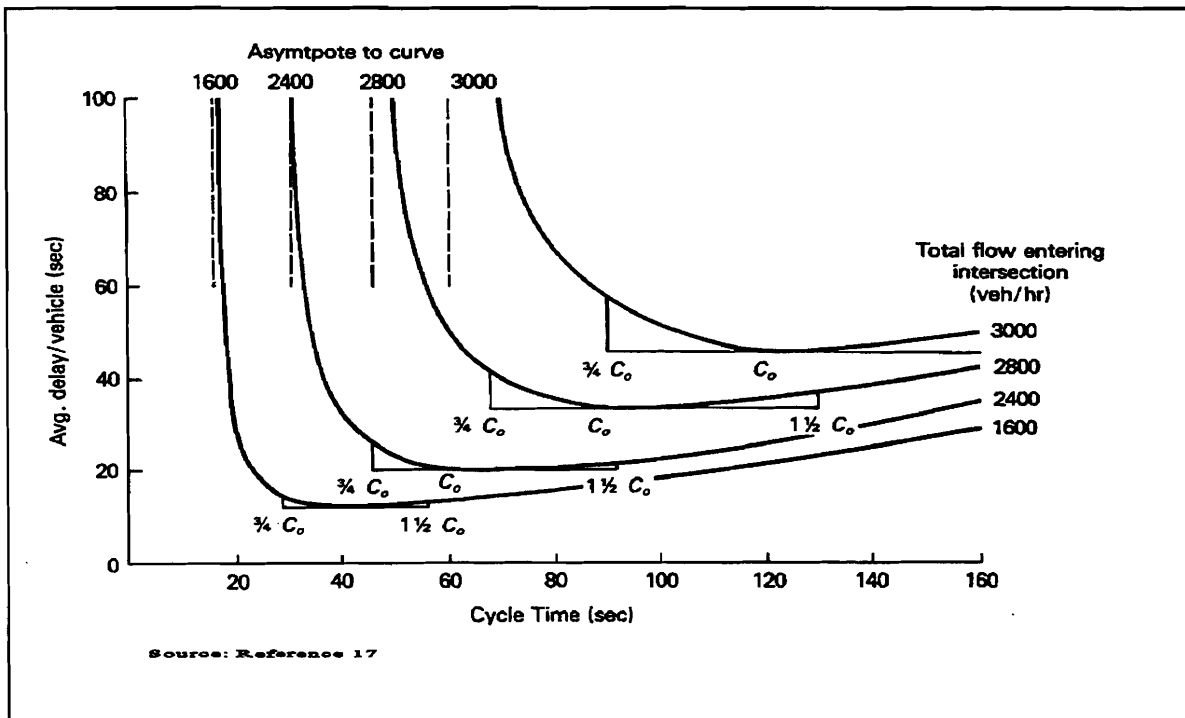

FIGURE IV-2. Effect on Delay of Variations in Cycle Length

Table IV-6 shows the lower and upper thresholds that tend to produce similar delays of some of the more commonly used cycle lengths. With the NASA Rd. 1 system, the cycle lengths used in the traffic responsive mode varied between 100 to 140 seconds. Notice that the

cycle lengths that were available in the traffic responsive mode fall within the upper and lower boundaries of the 100 second cycle. As a result, one could not expect the NASA Rd. 1 system to produce significant delay savings in the traffic responsive mode. To produce significant delay savings in a traffic responsive mode, cycle length changes on the order of magnitude from 90 to 140 seconds would be needed.

TABLE IV-6. Range of Cycle Length Producing Similar Delays

Lower Cycle Length [0.75 X C_0 (sec)]	Optimum Cycle Length [C_0 (sec)]	Upper Cycle Length [1.5 X C_0 (sec)]
45	60	90
67.5	90	135
75	100	150
82.5	110	165
90	120	180
105	140	210

RECOMMENDATIONS FOR USING TRAFFIC RESPONSIVE MODE

From the results of the simulation studies and the past literature, it is clear that traffic responsive mode should not be expected to yield significant reductions in total system delay over time-of-day control when traffic conditions do not change significantly. Instead, traffic responsive mode should be used to guard against major shifts in traffic demand (i.e., shifts that would require a cycle length outside 0.75 to 1.5 times the cycle length for a condition). The following sections discuss some of the traffic situations where traffic responsive control might produce significant delay savings over time-of-day mode.

It should be noted that not all of these conditions must exist in order to operate a signal system in a traffic responsive mode. Engineers must examine local conditions to determine which of the below conditions may apply in their specific locale.

Incidents

Incidents, by their nature, are unpredictable events and can have a dramatic impact on traffic patterns in a control area; therefore, areas that are subject to changes in traffic patterns due to incidents are likely locations for implementing traffic responsive mode. The impact of an incident on the traffic conditions varies depending on whether the incident occurs inside or outside the control area. When an incident occurs on an arterial within the control area, traffic flow upstream of the incident generally becomes more congested, while traffic flow downstream of the incident becomes less congested. Engineers may find it desirable to use traffic responsive mode to detect when these situations occur in the network and implement a timing plan that is specifically designed to accommodate traffic demands and manage queues that are associated with incidents.

Incidents that occur outside of the actual control area can also impact traffic operations within the control area. Traffic diverting from another arterial street or from a freeway can dramatically alter traffic patterns in a control area. Diverting traffic may result in a general increase in traffic demand throughout the entire network. If the signal systems are operating in a traffic responsive mode, these changes in traffic patterns can be detected, and a new timing plan that could mitigate the impacts of the incident on traffic flow in the control area can be implemented.

Special Events

One situation where traffic responsive mode may be particularly beneficial is in providing control after a special event (such as a football game, concert, etc.). The problem with providing time-of-day mode for special events is that, although the starting time of the event is known, the precise ending time is often unpredictable. Therefore, it is difficult to develop a time-of-day plan that can be implemented for special events. In a traffic responsive mode, the system detectors can be used to monitor traffic conditions to determine when the event ends. As traffic builds in the network, the signal system could then implement a plan specifically designed to accommodate traffic from the special event.

Early Exit of Time-of-Day Plan

Another situation where traffic responsive mode may be beneficial is in identifying when it may be appropriate to leave a particular time-of-day plan early. The need to exit a specific time-of-day plan early can arise when an expected traffic demand does not materialize on the network. For example, fluctuations in peak period demand may make it necessary to leave a peak period plan early. By operating the signal system in a traffic responsive mode, the signal system can implement appropriate timing plans when demands do not materialize as expected.

71

In the case where the ending point of the period remains relatively constant from day to day, a time-of-day operating mode would be more appropriate.

Adaptive Holiday Control

Another potential application for operating traffic signal systems in a traffic responsive mode is to provide for adaptive control during holiday periods. With some holidays (e.g., near Christmas), traffic patterns can be heavier than normal. With other holidays, traffic patterns can be lighter than normal. If traffic patterns are known, then the traffic responsive mode can be used to adapt timing plans to meet holiday conditions. This may include extending the peak plan past its normal time-of-day ending point, implementing a weekday peak timing plan during a weekend period, or exiting a time-of-day plan early because the normal traffic demands did not materialize due to the holiday period.

Low Volume

A final situation where traffic responsive mode might prove to be beneficial is during low volume conditions (e.g., like those occurring at night). Under these conditions, traffic volumes are generally unpredictable and can vary from cycle to cycle. As a result, it is difficult to provide good coordination during low volume conditions. The traffic responsive mode can be used to bring the signals into and out of coordination, as necessary.

CHAPTER V.
SUMMARY

Most closed-loop systems today are capable of operating in a traffic responsive mode; however, many systems still are left to operate in a time-of-day mode after installation. TxDOT believes that one reason why traffic responsive mode is not used more often is that there are no clear guidelines or procedures to help engineers set up a closed-loop system in a traffic responsive mode. Specifically, guidelines are needed to address the following issues:

- when to operate a group of traffic signals in an isolated versus a coordinated mode,

- where to locate system detectors to support traffic responsive mode, and

- when to operate a closed-loop system in a time-of-day versus a traffic responsive mode.

This report summarizes the results of several simulation studies that were performed to support the development of guidelines and procedures for setting up a closed-loop traffic signal system to operate in a traffic responsive mode.

A simulation study was performed to address the issue of when to provide isolated versus coordinated control in a group of signal-controlled intersections. A review of the literature revealed that several procedures have been developed to determine when to provide coordination between two signals. Of these procedures, the Interconnection Desirability Index was selected for further evaluation because it was the only procedure that expressly quantified the two main conditions that impact the ability to provide progression: the amount of platoon dispersion that occurs between two intersections and the amount of vehicles in the traffic stream traveling in platoons. Simulation studies showed, however, that the Interconnection Desirability Index was not a good indicator for choosing between isolated and coordinated control when mid-block volumes were high. As a result, a revised Interconnection Desirability Index was developed. Additional simulation studies were then performed to determine thresholds defining when to provide isolated and coordinated control. These studies found that when two intersections had a score of 0.4 or greater with the revised Interconnection Desirability Index, the system performed better under coordinated control. Likewise, the studies found that when two intersections had a score of 0.3 or lower, the signals produced less delay under isolated control. Between 0.3 and 0.4, system performance was approximately the same under isolated control as it was under coordinated control.

73

Simulation studies were also performed to examine where to locate system detectors to support traffic responsive mode in a closed-loop signal system. Three different configurations of system detectors were examined: one where system detectors were provided at the critical intersection only; another where system detectors were located on all approaches to the critical intersection and at the mid-point of the system; and a final configuration where, in addition to system detectors at the critical intersection and the mid-point of the arterial, system detectors were placed at the end downstream entry point of the system. Traffic conditions in the network were simulated with each detector configuration. Volume and occupancy measures from the detector configurations were used to provide input into the traffic responsive algorithm. The simulation studies showed that there was no significant difference in the amount of total system delay that was provided by the signal system operating in a traffic responsive mode under the different detector configurations. It was concluded, therefore, that as long as the system is small, system detectors placed only at the critical intersection should be sufficient to operate the system in a traffic responsive mode.

A final simulation study was performed to examine the issues of when to operate a closed-loop signal system in a traffic responsive versus a time-of-day mode. This simulation study showed that there was no statistically significant difference in the amount of total system delay when the traffic signal system was operated in a traffic responsive mode versus a time-of-day mode. The primary reasons provided for why no significant difference could be found between the two operating modes include the following:

- any reduction in the amount of total system delay that was achieved by operating the signals in a traffic responsive mode was offset by the delays associated with the transitions between timing plans and

- since previous research has shown that delays are relatively insensitive to changes in cycle lengths, the cycle lengths used to operate the system in a traffic responsive mode were too close to permit significant delay savings.

Therefore, it was recommended that when setting-up a traffic responsive system, engineers consider using timing plans that use large differences in cycle lengths.

REFERENCES

1. McShane, W.R., and R.P. Roess. *Traffic Engineering*. Prentice Hall, Englewood Cliffs, New Jersey, 1990, pp. 527-529.

2. Yagoda, H.N., E.H. Principe, C.E. Vick, and B. Leonard. Subdivision of Signal Systems into Control Areas. In *Traffic Engineering*, Vol. 43, No. 2, September 1973, pp. 42-45.

3. *Traffic Signal Applications Manual*. Department of Transportation, City of Arlington, Texas, February 1988, pp. 61-62.

4. Whitson, R.H., and C.J. Messer. *A Study of System Versus Isolated Control as Developed on the Mockingbird Pilot Study*. Prepared for the City of Dallas. Texas Transportation Institute, Texas A&M University System, College Station, Texas, February 1973.

5. Chang, E.C.P., and C.J. Messer. *Warrants for Interconnection of Isolated Traffic Signals — Final Report*. Report No. FHWA/TX-86/67+293-1F. Texas Transportation Institute, Texas A&M University System, College Station, Texas, August 1986.

6. Messer, C.J., D.B. Fambro, and D.A. Anderson. *A Study of the Effects of Design and Operational Performance of Signal Systems -- Final Report*. Report No. TTI-2-18-75-203-2F. Texas Transportation Institute, Texas A&M University System, College Station, Texas, August 1975.

7. Malakapalli, M.P. *Enhancements to the PASSER II-90 Delay Estimation Procedures*. Master of Science thesis. Texas Transportation Institute, Texas A&M University System, College Station, Texas, December 1992.

8. Kay, J., R. Henry, and S. Smith. *Locating Detectors for Advanced Traffic Control Strategies: Handbook*. Report FHWA-RD-75-91. FHWA, U.S. Department of Transportation, 1975.

9. Woods, D., and N. Rowan. *Detector Location for Computerized Arterial Street Sampling Detectors*. Report FHWA/TX-95/1392-6. Texas Transportation Institute, Texas A&M University System, College Station, Texas, 1995.

10. *921-TRM Traffic Responsive Closed Loop On-Street Master Manual Supplement, Preliminary, Revision 1*. Naztec, Inc., Sugarland, Texas, April 1992.

11.   *ASC/2M-1000 Zone Master Programming Manual.* P/N 37525, Econolite Control Products, Inc. Anaheim, California, undated.

12.   Ott, L. *An Introduction to Statistical Methods and Data Analysis.* Second Edition. Duxbury Press, Boston, Massachusetts, 1984.

13.   Basu, S. *A Study of the Factors Influencing the Number of Signal Timing Plans Required for UTCS First Generation Traffic Control Systems.* Report FHWA/RD-80/039. FHWA, U.S. Department of Transportation, July 1981.

14.   Henry, R.D., R.A. Ferlis, and J.L. Kay. *Evaluation of UTCS Control Strategies -- Technical Report.*   Report FHWA-RD-76-150.   FHWA, U.S. Department of Transportation, August 1976.

15.   Kay, J.L., J.C. Allen, and J.M. Bruggeman. *Evaluation of the First Generation UTCS/ BPS Control Strategy.*   Volume 1--Technical Report.   Report FHWA-RD-75-27. FHWA, U.S. Department of Transportation, February 1975.

16.   Henry, R.D., R.A. Ferlis, and R.M. White. *Application of UTCS First Generation Control Software in New Orleans.* Report FHWA-RD-78-3. FHWA, U.S. Department of Transportation, January 1978.

17.   Webster, F.V. Traffic Signal Settings. *Road Research Technical Report 39.* Road Research Laboratory, Her Majesty's Stationary Office, 1958: reprinted with minor amendments 1961.