Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA/TX-01/1439-10		

4. Title and Subtitle
DEVELOPMENT OF AN ACTUATED TRAFFIC CONTROL PROCESS UTILIZING REAL-TIME ESTIMATED VOLUME FEEDBACK

5. Report Date
September 2000

6. Performing Organization Code

7. Author(s)
Michael J. Pacelli, Carroll J. Messer P.E., and Thomas Urbanik II P.E.

8. Performing Organization Report No.
Report 1439-10

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.
Project No. 0-1439

12. Sponsoring Agency Name and Address
Texas Department of Transportation
Construction Division
Research and Technology Transfer Section
P. O. Box 5080
Austin, Texas 78763-5080

13. Type of Report and Period Covered
Letter Report:
September 1999 – August 2000

14. Sponsoring Agency Code

15. Supplementary Notes
Research performed in cooperation with Texas Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration.
Research Project Title: TxDOT Support for the Texas A&M IVHS Research Center of Excellence

16. Abstract

The goal of this research was to develop an actuated traffic control process that could use estimated volumes in order to optimally operate the traffic signal in real-time in response to actual traffic demands, or a reasonable estimate of demand. A further goal of this research was to establish the relationship between the traditional control parameter, passage gap and key operating parameters, in order to allow changes in signal operations to be made by means of passage gap adjustments. The relationships between passage gap and cycle length, green splits, and interval length were studied, and the cycle length relationship was formalized mathematically. Results indicated that the volume estimation methodology could be readily calibrated to provide good estimates of traffic volumes by movement. The scope of the research dealt with a 60 foot stop-line detector configuration. The overall study results suggest that this configuration is operationally very efficient for minimizing delay, but provides little "dilemma zone" protection for arriving motorists at low-volumes and high-speeds.

17. Key Words
Signalized Intersections, Feedback Algorithm, Estimated Volumes, Gap Setting, Detector Setting

19. Security Classif.(of this report)	20. Security Classif.(of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	90	

DEVELOPMENT OF AN ACTUATED TRAFFIC CONTROL

PROCESS UTILIZING

REAL-TIME ESTIMATED VOLUME FEEDBACK

by

Michael J. Pacelli
Graduate Research Assistant

Carroll J. Messer, P.E.
Research Engineer
Texas Transportation Institute

and

Thomas Urbanik II, P.E.
Associate Director
Texas Transportation Institute

Report 1439-10
Project Number 0-1439
Research Project Title: TxDOT Support of the Texas A&M IVHS Research Center of Excellence

Sponsored by
Texas Department of Transportation
In cooperation with the
Federal Highway Administration, U.S. Department of Transportation.

September 2000

TEXAS TRANSPORTATION INSTITUTE
The Texas A&M University System
College Station, Texas 77843-3135

ABSTRACT

Actuated traffic controllers are intended to determine traffic conditions, in real-time by means of vehicle detection, and respond accordingly in order to maintain the highest reasonable level of efficiency under varying conditions. However, modern traffic controllers are essentially sophisticated electronic timers. Rather than merely responding to detector actuations, it is desirable to improve this system such that the controller has a goal to be achieved and acquires information about the present state of the system in the same units. The goal of this research was to develop an actuated traffic control process that could use estimated volumes in order to optimally operate the traffic signal in real-time in response to actual traffic demands, or a reasonable estimate of demand.

Once the present and desired states of the system are known, the control changes necessary to move from the present state to the optimal target state can be determined. A further goal of this research was to establish the relationship between the traditional control parameter, passage gap and key operating parameters, in order to allow changes in signal operations to be made by means of passage gap adjustments. The relationships between passage gap and cycle length, green splits, and interval length were studied, and the cycle length relationship was formalized mathematically. The quantified form can then be used as a tool to adjust the signal performance to approach the desired operating state.

Research was conducted with computer simulation, using both stand-alone software and a hardware-in-the-loop setup. Testing and comparison between methods validated the use of these models.

Results indicated that the volume estimation methodology could be readily calibrated to provide good estimates of traffic volumes by movement. Furthermore, simulation results quantified the relationship between passage gap and cycle length, thereby establishing a mechanism by which to directly implement signal operating changes at an actuated traffic signal.

The scope of the research dealt with only a 60-foot stop-line detector configuration. The overall study results suggest that this configuration is operationally very efficient for minimizing delay, but provides little “dilemma zone” protection for arriving motorists at low-volumes and high-speeds. The research results suggest that the operational results may have been different had other high-speed detection options been considered.

TABLE OF CONTENTS

Page

LIST OF FIGURES ix

LIST OF TABLES xi

INTRODUCTION 1

- Problem Statement 1
- Research Objectives 3
- Scope of Work 4

BACKGROUND 5

- Types of Traffic Control 5
- Other Traffic Signal Control Features 7
- Traffic Control Standards 9
- Feedback Control Systems 10
- Proposed Control System 13
- Volume Estimation 18
- Volume Forecasting 19
- Target Volume-Capacity Ratio Determination 20
- Cycle Length Estimation 22
- Simulation Technologies 24

STUDY METHODOLOGY 31

- Intersection Testbeds 31
- Experiment Structure 35
- Comparison of Hardware- and Software-Based Simulations 38
- Validation of Volume Estimation Methodology 39
- Cycle Length Adjustment 40

STUDY RESULTS 47

- Comparison of Theoretical Analyses and Simulations 47
- Comparison of Hardware- and Software-Based Simulations 49
- Check of NETSIM-Generated Volumes 51
- Validation of Volume Estimation Methodology 54
- Cycle Length Results 57

Green Split Results ... 60
Interval Length Results ... 62
Delay Results ... 63
Summary of Simulation Results ... 66
Erlang Model Results ... 66
Combining Theory and Reality ... 68
Calculation of Passage Gap ... 70
Resulting Control Process ... 74

CONCLUSIONS ... 75

REFERENCES ... 77

LIST OF FIGURES

FIGURE Page

1 Volume-Density Control. 8
2 Basic Continuous Feedback Control System. 12
3 Proposed Continuous Feedback Traffic Control System. 13
4 Diagram of the Actuated Control System with Estimated Volume Feedback. 17
5 Basic Queueing Theory Model. 18
6 TTI REL CID and Eagle Controller. 28
7 TransLink® Roadside Equipment Laboratory. 30
8 Basic Isolated, Actuated Intersection Testbed. 32
9 Advanced Isolated, Actuated Intersection Testbed. 33
10 Eight-Phase, Dual-Ring, Quad-Left, NEMA Signal Phasing Plan. 33
11 Queued Vehicle Trajectories at Green Onset. 35
12 Expected Cycle Length Determination Process. 42
13 Software-Only Simulation vs. Theoretical Analysis Results. 48
14 Hardware-in-the-Loop Simulation vs. Theoretical Analysis Results. 49
15 Software-Only vs. Hardware-in-the-Loop Simulation Results. 50
16 TSIS Target Volumes vs. Measured Volumes (Phases 1-4). 52
17 TSIS Target Volumes vs. Measured Volumes (Phases 5-8). 53
18 Volume Estimation Methodology (Phase 1-4). 55
19 Volume Estimation Methodology (Phase 5-8). 56
20 Software-Based, Basic Testbed Cycle Lengths. 58
21 Software-Based, Advanced Testbed Cycle Lengths. 59
22 Hardware-Based, Advanced Testbed Cycle Lengths. 60
23 Software-Only Average Delay vs. Passage Gap. 65
24 Hardware-in-the-Loop Average Delay vs. Passage Gap. 65
25 Erlang Model Cycle Length vs. Volume-Capacity Ratio. 67
26 Erlang Model Cycle Length vs. Critical Intersection Flow Ratio. 68
27 Combined Cycle Length vs. Volume-Capacity Ratio. 69

28 Combined Cycle Length vs. Flow Ratio. 70
29 C(1-Y) Transform vs. Passage Gap. 72

LIST OF TABLES

TABLE Page

1 Summary of Simulation Results 66

1. INTRODUCTION

Traffic signals are intended to provide for the safe and efficient assignment of conflicting rights-of-way at at-grade roadway intersections. Safety can be achieved by consistently following established guidelines and standards based on driver expectancy, vehicle stopping distance requirements, and pedestrian crossing times. Achieving the most efficient operation of a signal can be significantly more challenging when faced with the widely varying traffic conditions that can be present at an intersection over time.

Early pretimed signals allocated green time to movements according to a fixed schedule or timing plan input by a traffic engineer or technician. This plan was developed based on historical traffic volumes, and, although several plans may exist for each day, the signal would blindly perform the timing plan regardless of the presence or absence of traffic on a given cycle. In order to most effectively allocate the green time available at a signalized intersection, actuated signals have replaced many pretimed signals. Actuated signals use detectors to determine if vehicles are present, or calling, and serve movements only if and when demand exists.

PROBLEM STATEMENT

Actuated traffic controllers are intended to determine actual traffic conditions, in real-time by means of vehicle detection, and respond accordingly in order to maintain the highest reasonable level of efficiency under varying traffic conditions. However, modern traffic controllers are essentially sophisticated electronic timers that are calibrated to respond in a certain manner when a sensor's switch (or detector) is actuated. The traffic controller does not know the actual state of the traffic, beyond the mere fact that a call for service exists, and it does not have a defined goal to achieve by objectively modifying its output. Output is controlled solely on a stimulus-response basis without consideration of the intersection as a whole.

Given the existing control system structure in use by actuated controllers today, it is desirable to improve this system such that the controller has a goal to be achieved and acquires information about the present state of the system in the same units. This improvement would allow for more direct comparisons and subsequent determination of the appropriate control decisions to be made.

The problem with implementing this philosophy today is that goals in traffic engineering are set in terms of using the available green time to the highest reasonable level of efficiency. This setting typically means, for an actuated signal, a volume-to-capacity ratio of approximately 0.90 is indirectly used. Modifying control parameters to achieve this goal requires knowledge of the demand volumes present for each phase at any given time, but the prevalent form of detection today (inductive loops) merely reports vehicle presence. Secondary loop detectors used for counting volumes are inherently problematic, as they must be placed a great distance upstream of the stopline to avoid being covered by queue spillback, and this placement is frequently not possible in congested and highly developed areas. Other new technologies are emerging, such as video image processing, but such systems are costly, relatively unproven, and not currently in common use.

The goal of this research was to address the above problem, in part, by validating a methodology that can estimate traffic volumes based on time measurements that are readily obtainable from detectors in widespread use throughout the country.

Once the present and desired states of the system are known in the same units, the control changes necessary to move from the present operating state to the goal or target state can be determined. There are several key control parameters that can alter the operation of an actuated signal. Most notable are minimum green time, maximum green time, and passage gap (or unit extension). Minimum and maximum green times are not flexible enough to allow for cycle-to-cycle variation because they are limited by outside factors. Minimum green times are commonly restricted due to safety concerns associated with driver expectancy and pedestrian crossing times, and maximum green times are based on a reasonable upper limit on the tolerable wait time that can be experienced by motorists. Passage gap is the parameter that is used to extend a phase from its minimum green time toward its maximum green time. Consequently, passage gap is the preferred variable to be adjusted, if possible, in order to alter the operating state of the actuated signal. This use is a logical extension of the present use of passage gap, with the use of variable gap settings representing a another degree of freedom.

However, the relationships between the cycle length, green splits, and passage gap setting were not well established. It also was not clear whether or not altering the volume-capacity ratio may adversely affect other intersection parameters, and, consequently, increase overall vehicle delay. This research also addressed this issue, as described later in this report.

RESEARCH OBJECTIVES

The basic goal of this research was to develop a new traffic control strategy that could more efficiently operate an isolated, actuated traffic signal in real-time with respect to overall traffic conditions present at the intersection. The specific objectives of this research were to:

- validate the volume estimation methodology developed by Messer et al. (*1*);
- establish and quantify any relationships that may exist between variable passage gap and the resulting cycle length;
- establish any relationships that may exist between variable passage gap and the resulting green splits;
- establish any relationships that may exist between volume-capacity ratio and average intersection delay; and
- combine the above relationships in the development of a process for actuated traffic signal control utilizing real-time estimated volume feedback.

SCOPE OF WORK

Simulation studies and data reduction were performed for a series of varying traffic volume and loading conditions, as described in further detail in this report. The primary research tool was hardware-in-the-loop simulations, with validation and testing of more far-ranging scenarios conducted using software-only simulation and theoretical analysis. The details of these techniques and the subsequent analyses and results are described in detail later in this report.

2. BACKGROUND

This chapter begins with a brief overview of the current state-of-the-practice regarding the operation of isolated, actuated traffic signals, and a discussion of basic feedback control system theory. Also included are discussions and derivations relating to some of the fundamental concepts on which this research is based, including volume estimation, volume forecasting, and methods for the determination of a target volume-capacity ratio. Finally, the methodology used in the estimation of cycle lengths is explored in detail.

TYPES OF TRAFFIC CONTROL

When traffic volumes or safety concerns warrant the placement of a form of traffic control at the intersection of two or more roadways, there is a wide spectrum of available control measures. The appropriate traffic control measure for any given situation is selected according to warrants prescribed in the *Manual on Uniform Traffic Control Devices* (*2*). A number of factors are considered by these warrants, including various combinations of traffic volumes, pedestrian volumes, accident history, etc. The various levels or degrees of traffic control are listed below.

Unsignalized Traffic Control

Unsignalized control includes "no control," or those cases in which the rules of right-of-way control. It also includes various possible configurations of yield and stop sign control.

Signalized Traffic Control

Signalized traffic control uses traffic signals to alternately assign right-of-way to conflicting movements. There are various forms of signalized control, as described in the following sections.

Pretimed Time Control

In fixed time, or pretimed operation, all operating parameters of the signal are preset in the controller, which repeatedly executes the predefined pattern regardless of traffic conditions. Early pretimed controllers executed the same pattern continuously, while later, more advanced devices may have several different patterns corresponding to different times of the day, such as for morning, noon, and evening traffic flow conditions *(3)*. The traffic engineer or technician defines the patterns based on historical data and experience.

Semi-Actuated Control

In semi-actuated control, the minor street receives green only when there is traffic present. Some form of detection determines when vehicles are present on the minor street and provides a variable amount of green time to that movement depending on the status of the detector. When a call for service is no longer detected, or the maximum allowable service time has been reached, the green indication is returned to the main street. The green indication always returns to and rests in the main street phases when no minor street demand is detected *(3)*. This form of signalization is best suited to sites where a major street is intersected by a relatively low volume cross street and the volume on the main street is such that minor street vehicles would have difficulty finding acceptable gaps in traffic.

Fully Actuated Control

In fully actuated control, all phases are controlled through the use of detection. Control parameters that must be set include the minimum and maximum service time to be provided to each phase. This form of control is more efficient than the previous forms in cases where traffic demands vary over time. In addition, phases may be skipped if demand does not exist during a particular cycle *(3)*.

OTHER TRAFFIC SIGNAL CONTROL FEATURES

In addition to the use of actuated traffic signal control, other features have been selectively incorporated into traffic signals in order to improve the efficiency of their operations. This section will briefly highlight some important features.

Protected or Permitted Turn Phases

Vehicles turning left across opposing traffic may be handled in three primary ways at a signalized intersection. These three options are available to the traffic engineer regardless of whether the signal is actuated, semi-actuated, or pretimed.

Under low volume conditions, a single phase serving the combined through and turn movements is generally sufficient. Left turning vehicles yield the right-of-way to opposing through vehicles, and they may make a permitted turn through gaps in the opposing traffic stream.

As volumes begin to increase, and as the number of opposing lanes increases, it becomes more difficult for motorists to find acceptable gaps in the opposing traffic in order to safely complete a left-turn maneuver. Under these conditions, left-turn capacity is not sufficient to meet demand, and an exclusive, protected turn phase may be warranted. This phase allows turning maneuvers to proceed without interference from opposing traffic.

In order to further increase turn capacity at signalized intersections, some traffic engineers have incorporated protected/permitted phasing, which allows permitted left turns during the through phase, but it also provides an exclusive turn phase without opposing traffic. Often the protected turn phase is actuated and lags the associated through phase, so that it is only displayed if demand continues to exist following the permitted phase.

Volume-Density Adjustment

A further enhancement of actuated control is the volume-density feature available on many signal controllers. This feature reduces the passage gap setting, which extends an actuated phase by a specific amount of time whenever a vehicle is detected by a specified extension detector. By extending in this way, the length of the phase is increased when continued demand is detected. When the controller detects a gap in traffic that is longer than the passage gap, the phase terminates to provide service to a competing movement. By systematically reducing the gap size under volume-density operation, it becomes easier to terminate the phase as the time into the phase increases, thus allowing the phase times to be kept shorter than would otherwise occur if the gap setting were fixed at its initial setting. Figure 1 illustrates this concept (*4*).

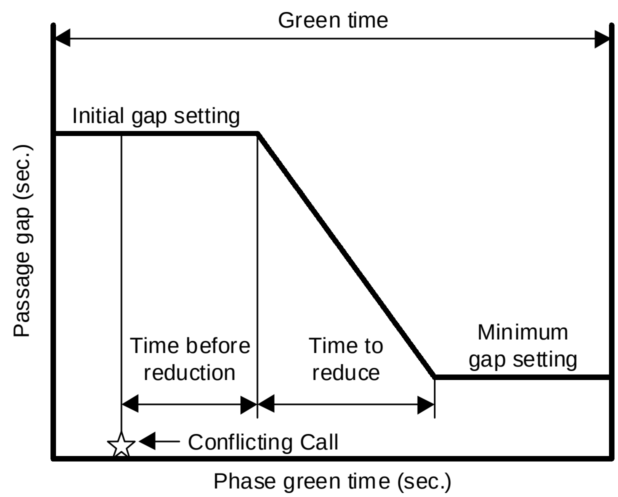

Figure 1. Volume-Density Control (*4*).

Another type of volume-density control uses detectors upstream of the stopline to count vehicles as they enter the queue, and then adjusts the length of a variable initial green period to satisfy the discharge requirements of the number of vehicles known to be in queue. Following the variable initial period, the gap can be set to zero, to terminate the phase immediately after queue discharge, or the gap can be reduced at some given rate from its initial setting.

Although this form of volume-density control does incorporate volume counting, the proper placement of the detectors is difficult. The loops must be located far enough upstream to capture the entire length of the maximum expected queue. However, as the setback distance increases, the likelihood increases that an intermediate access point, such as a driveway, will allow vehicles to enter the queue storage area and bypass the count detector. In addition, this function is unable to work with more complicated detector configurations that tend to count vehicles multiple times.

Real-Time Traffic Adaptive Control System

Another more recent development in traffic control strategy is the Real-Time Traffic Adaptive Control System (RT-TRACS), developed by PB Farradyne, Inc. under the direction of the Federal Highway Administration (FHWA). Recognizing that current traffic control strategies are not sufficient to deal with many of today's more complex traffic problems, the goal of the program was to develop a system that could compare several possible signal control strategies, decide on which is most appropriate given the prevailing traffic conditions, and implement the best strategy without operator intervention. Five signal control strategies, each with strengths and weaknesses under certain types of conditions, were commissioned as prototype strategies for RT-TRACS (*5*).

RT-TRACS is essentially a process for selecting among competing traffic control strategies for the purpose of choosing the predefined strategy that is most appropriate at that time. In that way, it encourages the development and testing of new traffic control strategies that show benefits under some, but not all, sets of conditions. Since RT-TRACS can switch to the most appropriate strategy as conditions change, new strategies developed do not need to compromise the level of performance under all conditions, but can achieve a higher level of performance under a narrower set of conditions. As such, it encourages research efforts such as this one and facilitates new approaches to traffic control.

TRAFFIC CONTROL STANDARDS

In order to ensure compatibility between traffic control equipment produced by different manufacturers, various hardware and communications standards have been established in order to facilitate the interconnection of components. The recognized authority for setting traffic control standards is the National Electrical Manufacturers Association (NEMA).

NEMA has established two primary sets of standards. The older TS1 standard was put into practice in the late 70's, and was subsequently upgraded throughout the 80's. With increasing requirements for communications and other more advanced features, the TS2 standard was proposed in the late 80's and has come into common use in the 90's.

The introduction of the TS2 standard brought about standardization for a number of features now commonly used in actuated traffic signal control. The most noteworthy of these features include:

- conditional service,
- added detector inputs,
- detector delay/extension and switching,
- dual entry,
- alternate phase sequencing,
- automatic flash, and
- standardized coordination (*6*).

FEEDBACK CONTROL SYSTEMS

Given that the actuated traffic control practices described above are variations of feedback control systems, it is important to define the terminology and explore the fundamental concepts behind basic feedback control systems. This section will serve only as a brief introduction to the subject area, with examples referring specifically to the envisioned use of feedback control in traffic signal systems.

First, in the most general terms, the following definition can be used:

> A control system is an arrangement of physical components connected or related in such a manner as to command, direct, or regulate itself or another system (*7*).

A control system interacts with other systems and the surrounding environment by one of two means. Input is defined as any "stimulus, excitation, or command applied *to* a control system," while output is defined as the "actual response obtained *from* a control system" (*7*).

There are two fundamental types of control systems: open-loop and closed-loop. The difference is rooted in the nature of the input to the control action, or the source that is responsible for activating the control system to produce output. If the control action is independent of the output produced, then the system is open-loop. If the control action is dependent on the output, then the system is said to be closed-loop (*7*).

The calibration of the control system determines the quality and accuracy of the output from an open-loop system. Consequently, such systems are generally less susceptible to variations and instabilities in the environment. Closed-loop systems utilize feedback, which allows the control system to generate its control action as a function of both the input and the output. The use of feedback in a closed-loop system has a number of advantages, including increased accuracy and reduced effects of external disturbances and variations. Since actuated traffic control systems are intended to adjust their operations (control actions) in response to variations in the traffic conditions (external disturbances), a closed-loop system is clearly required. As such, the remainder of this report will focus solely on closed-loop feedback control systems (*7*).

In this type of system, a reference input is provided to the processor, as well as a feedback signal indicating some measure of the current state of the system. The processor then generates a control action or actuating signal as a function of both the input and the feedback. The control elements implement this control action, and, via a control signal, alter the state of the plant or process under control. The effects of external disturbances also impact the plant or process directly. Measurements of the resulting state of the system are returned as feedback by feedback elements to the processor to be used as inputs in the determination of the next control action. Figure 2 illustrates, in block diagram format, the structure of a nonspecific continuous feedback control system.

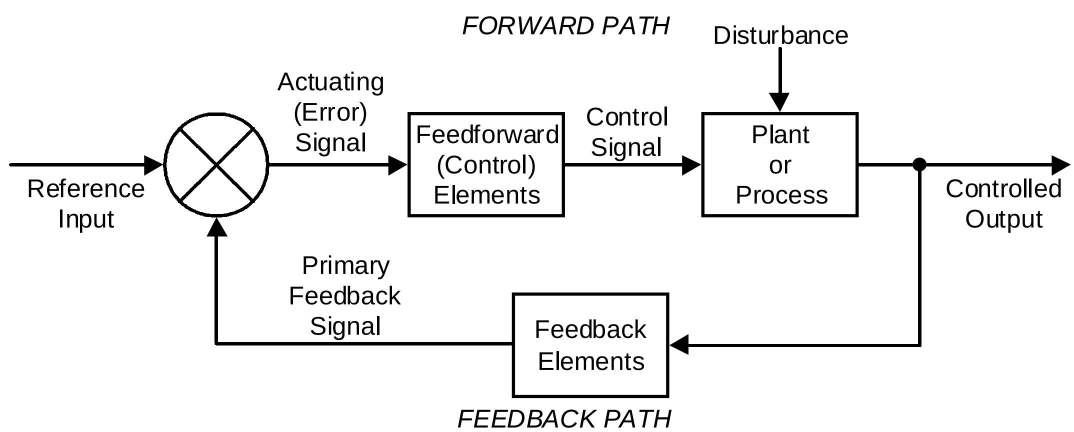

Figure 2. Basic Continuous Feedback Control System (*7*).

The above procedure could be used in a conventional, traffic-actuated signal controller. Figure 3 presents a block diagram illustrating this envisioned feedback process with the elements identified as they would be in the case of a traffic signal controller.

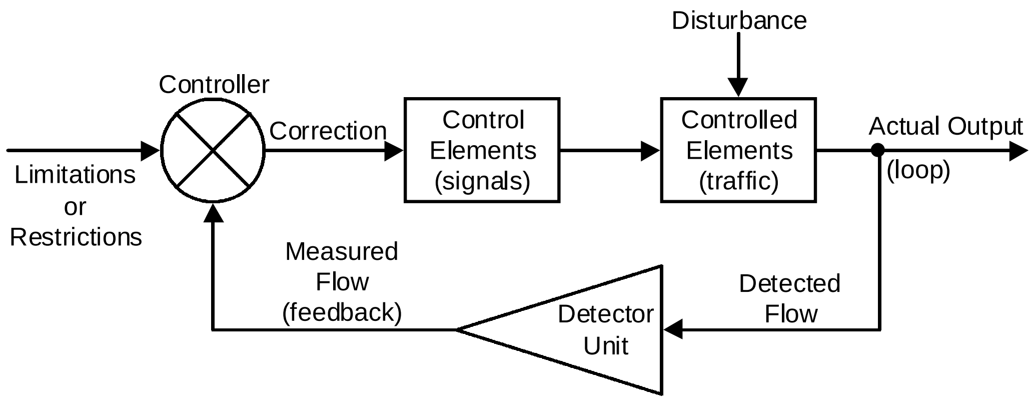

Figure 3. Proposed Continuous Feedback Traffic Control System *(1)*.

PROPOSED CONTROL SYSTEM

The control system proposed in this research is a modification of the basic feedback control system commonly used in traffic signal control, as described above. We propose that several additional steps be added, and validity of each proposed step was examined as part of this research.

It should be noted here that the control system described below is proposed for an intersection with certain assumed characteristics. For example, this research deals exclusively with long (60-ft.) stopline detectors, the details of which are discussed in the next chapter. No consideration is made for the use of extension detectors setback from the stopline, or for the use of multiple detectors. It is also assumed that the intersection is isolated and is subjected to randomly distributed traffic patterns. Therefore, the discussions, statements, and analysis contained herein refer to those intersections with these characteristics, and may or may not apply to other conditions.

Volume Estimation

The first addition in the proposed control system is a volume estimation function. Traditional traffic controllers receive all of their information regarding the current state of the traffic system in the form of detector calls, typically from inductive loop detectors cut into the pavement surface. These calls, or actuations, simply indicate that something is in a specific location, demanding service for a particular movement. The controller cannot determine if that call is due to a single vehicle or a large platoon of vehicles.

For example, consider the following hypothetical case. Assume that an isolated, actuated intersection is servicing the main street. A call for service is registered on the cross street. If the minimum green time has elapsed on the main street, the controller would terminate its current phase and provide service to the cross street as soon as it is safe to do so (all extensions elapsed). If there is a significant queue on the cross street, then the minor street phase may run until it reaches its maximum green time, and then terminate and return service to the main street. Any remaining queue will be forced to wait through a long cycle before service comes back to the cross street. Still detecting a call on the cross street, the controller would return to that phase after the minimum green time and any required main street extensions (or minimum pedestrian times) have elapsed. The controller would not have any memory of the last time that it served the cross street, would not know that there was a long queue on the previous cycle, and would attempt to terminate the cross-street phase by searching for a given gap size just as if it had never served that movement before. Consequently, a slow-moving vehicle or inattentive driver could produce such a gap, and the phase may terminate prematurely without providing a reasonable amount of service to the cross street for the volume conditions present.

If the controller had known that on the previous cycle, or perhaps several cycles, there was significant demand, it could have adjusted its critical gap setting in order to extend the likely phase length provided to the cross-street, given that it was experiencing an unusually high demand. In this manner, the controller would be more "forgiving" in terms of critical gap selection in order to ensure that the high volume movements are receiving service commensurate with the demand volumes.

If a phase does terminate prematurely, then a residual queue would remain and a call for service would be registered for that phase on the next cycle. Assuming that the cause of the problem has cleared the intersection and that the phase does not terminate prematurely on the next cycle, then this proposed control process would recognize through volume estimation that a high volume exists for the given movement. As such, the gap setting for that movement would be increased and subsequent cycles would be more likely to extend, and the control system would be more “forgiving” of the causes of premature gap out for that phase until the volumes had been reduced again.

Similarly, a call for service that follows a long period of little or no demand may indicate that the critical gap setting could be shortened, thus allowing the phase to terminate more readily, reducing the likelihood of wasting green time with an excessively long extension period.

The ability to estimate volumes, rather than simply register calls, would allow the controller to adjust the critical passage gap setting as described in the example above. The proposed addition to the control system would use stopline detector actuations to estimate the demand volume for each respective movement, using the methodology described in the next chapter of in this report.

Controller Processing

Although not significant in terms of major conceptual changes to the control system, the incorporation of volume estimation and passage gap variability would require new calculations to be made within the controller.

First, the availability of volume estimations allows controllers to work in units of measurement other than time. Traditionally, controllers were programmed with strictly time-based measures, such as minimum and maximum greens, cycle length, and critical gap. The goal was generally to apply these time-based inputs while trying to achieve a cycle length that was previously calculated by a traffic engineer to be optimal. However, outside of a coordinated network, the true goal is not to provide a cycle of some given length, but to provide the most efficient signal operation for the current set of traffic conditions within a range of reasonable (safe and tolerable) cycle lengths.

Therefore, the control system we propose in this research operates, at least in part, in terms of volume-capacity ratio, since this term is a more reasonable measure of efficient operation than cycle length. It is possible to calculate a volume-capacity ratio using the estimated flow ratio, measured cycle length, and an assumed value for total intersection lost time. Although the process of assuming values does introduce some uncertainty into the process, it is only important if the lost time changes notably with time or by movement.

The value calculated for the volume-capacity ratio can then be compared against a target value input by the traffic engineer. The methods available for determining the desired volume-capacity ratio are described in a later chapter.

Once the controller knows whether it needs to increase or decrease the capacity of each critical movement, the necessary adjustments in terms of cycle length and green splits among movements can be determined and implemented using the techniques described in the following section.

Passage Gap Adjustment

The second major concept incorporated into the proposed control system is the use of variable passage gap or critical gap settings to change the signal operating parameters, particularly cycle length and green splits. This research sought to establish the relationship between passage gap and cycle length, and between passage gap and green splits, for the purpose of defining specifically how adjustments are to be accomplished in this way. The specific procedure to be used is described in a later chapter.

It is also important to note that, since it was not feasible to study the effects of variable passage gap on every parameter of intersection operation, an overall summary measure, average delay per vehicle, was selected. This measure was checked against varying passage gap settings to ensure that varying this setting for the purpose of altering signal operation did not adversely effect other performance measures.

Combined Control System

Figure 4 shows how the two major components of this research effort (shown shaded) fit together to support a new actuated control system using estimated volume feedback and passage gap adjustment.

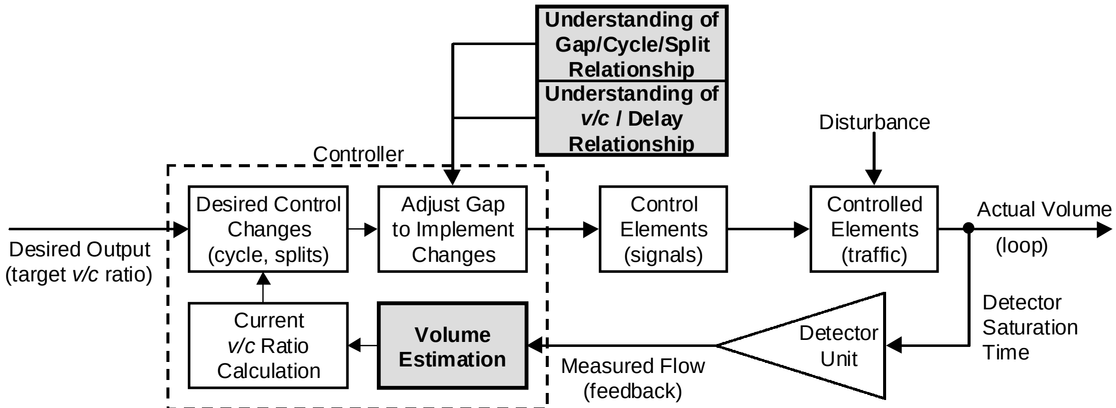

Figure 4. Diagram of the Actuated Control System with Estimated Volume Feedback.

VOLUME ESTIMATION

Messer et al. developed the methodology in order to estimate the flow ratio loading of any given signal phase based on stopline detector actuation timing measurements and fundamental queuing theory (*1*). The flow ratio is estimated as follows:

$$y_i = \frac{V}{S} \approx \frac{g_p}{r + g_p} \qquad \textbf{(1)}$$

where

y_i = flow ratio of study approach *i*;

V = average arrival volume on approach, vph;

S = saturation flow rate at stopline on approach, vphg;

g_p = measured duration of platoon clearance (saturated green) time, sec; and

r = measured duration of red interval preceding measured green, sec (*1*).

Figure 5 illustrates the derivation of this equation.

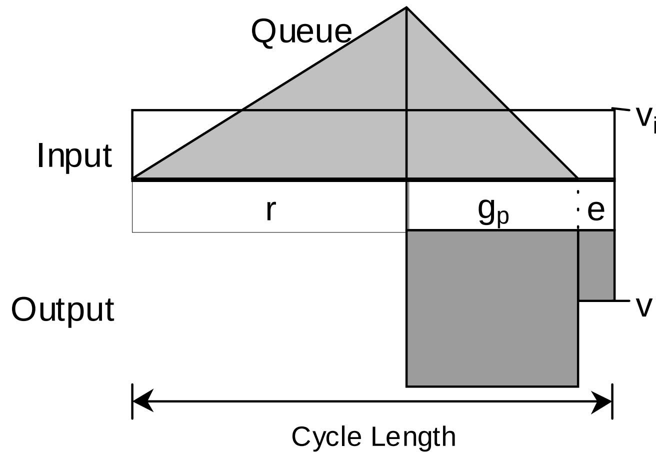

Figure 5. Basic Queueing Theory Model.

Assuming undersaturated conditions, the total number of vehicles entering the queue (shown in the upper half of Figure 5) must equal the total number of vehicles output through the signal (shown in the lower half of Figure 5) for any given cycle. For the purposes of analysis, an isolated intersection is assumed and input vehicles are assumed to arrive at a uniform arrival rate (v_i). Since no vehicles are assumed to be discharged during the effective red (r) portion of the cycle, arriving vehicles are stored in queue during this period. Vehicles are discharged at a uniform rate (s), or the saturation flow rate, during the saturated portion of the effective green. Any remaining effective green (e), or extension, outputs vehicles at the arrival rate (v_i), since no queue exists during that part of the cycle. Consequently, the flow in equals the flow out when no initial queue is present. If the signal is undersaturated, then the total flow in also equals the total flow out over the entire cycle length.

Using the variables defined previously, the volume balance described above can be written mathematically as follows:

$$v_i(r + g_p + e) \approx s(g_p) + v_i(e) \qquad \textbf{(2)}$$

Volume flow during the extension period simply passes through the system without joining the queue, and so it is represented identically on both sides of the equation and can be cancelled out. The remaining terms can be rearranged such that, if the saturated portion of green (g_p) and the effective red (r) are known, and a saturation flow rate can be estimated, then the input volume (v_i) can be found as follows:

$$v_i \approx s \frac{g_p}{g_p + r} \qquad \textbf{(3)}$$

The approach flow ratio (v_i/s) can also be readily found using time-based measures of key cycle parameters and stopline detector actuations without the need for estimating the saturation flow rate.

This methodology allows estimation of the volume processed for any given movement using existing stopline detection, rather than requiring additional detection means upstream of the intersection area. Estimated volumes could be utilized on a cycle-by-cycle basis, or averaged over several cycles to provide smoother transitions.

VOLUME FORECASTING

The volume estimation method described in the previous section relies on time-based measurements of events that occurred during the previous cycle. Therefore, the estimation of volume represents traffic that has already passed through the signal. This could mean the input volume to the signal for the last full cycle, or it could be an average value taken over several preceding cycles in order to smooth random variations. In either case, however, an estimate of the immediate past could be used to adjust signal timing.

Although beyond the scope of the studies conducted for this research, it would be desirable to use the volume estimations from the past to project future volumes for each movement at the intersection, and to adjust the signal timing parameters for future cycles to accommodate the anticipated traffic rather than the past traffic.

TARGET VOLUME-CAPACITY RATIO DETERMINATION

The proposed control system described in the previous sections determines the current operating volume-capacity ratio and then adjusts the key control parameters in order to move the signal operation closer to the target volume-capacity ratio. Although the target volume-capacity ratio is more representative of the overall operating efficiency of the signal than the target cycle length, this process still requires a specific target value to be selected. Clearly, development of a control system for the purpose of achieving a specific operating volume-capacity ratio is not beneficial to the flow of traffic if selection of the target value is arbitrary. This section describes two methods for selecting the volume-capacity ratio that could serve as the target value for the controller.

Maximum Reasonable Volume-Capacity Ratio

Typically, when traffic engineers are selecting a volume-capacity ratio in order to calculate an optimal cycle length, the largest reasonable value is chosen. A larger value implies better utilization of the available service time and, therefore, more efficient operation of the signal. There is a disadvantage to this high level of efficiency, however, as the ratio approaches a value of one. In this case, there is little or no reserve capacity to absorb the natural volume fluctuations that occur from cycle to cycle, and a larger number of cycles will be oversaturated, leaving a residual queue and significantly increasing delay. Consequently, it is widely accepted that the largest reasonable ratio that can be reliably used is approximately 0.90. This usage leaves 10 percent of the capacity unutilized, on average, to account for future cycles having above average arrival volumes due to a random (Poisson) traffic distribution.

Minimum Delay Volume-Capacity Ratio

There is also another, perhaps superior, method of selecting a target volume-capacity ratio. Another method of improving operating efficiency is to minimize the total delay experienced at the intersection. A volume-capacity ratio that results in the minimum delay can be calculated as described in the following paragraphs and used as the target value for the controller to compare against measured field conditions.

First, Webster has shown that the minimum delay cycle length for fixed time signals can be computed as follows:

$$C_O = \frac{1.5L + 5}{1 - Y} \tag{4}$$

where

C_o = minimum delay (optimal) cycle length, sec;

L = total intersection lost time, sec; and

Y = sum of critical flow ratios (*8*).

The cycle length for a volume-capacity ratio of any value can be calculated as follows:

$$C_x = \frac{L}{1 - Y/X} \tag{5}$$

where

C_x = calculated cycle length, sec; and

X = given volume-capacity ratio (*8*).

Combination of the two previous equations allows us to determine the volume-capacity ratio at which the minimum delay fixed time cycle length occurs. The specific form based on the data presented by Webster is as follows:

$$X_o = \frac{Y(1.5L + 5)}{L(Y + 1/2) + 5} \tag{6}$$

Another form of this equation, generalized to allow for calibration to local conditions, is as follows:

$$X_o = \frac{Y(\alpha L + \beta)}{\alpha L + \beta - L(1 - Y)} \tag{7}$$

where

α = optimal cycle length calculation coefficient (i.e., Webster's $\alpha = 1.5$) and

β = optimal cycle length calculation constant (i.e., Webster's $\beta = 5$).

Note that this formula provides a more scientifically defensible method of selecting a target volume-capacity ratio, but it assumes that the startup lost times on all critical movements are approximately equal. This assumption is generally reasonable, but it should not be overlooked (*8*).

CYCLE LENGTH ESTIMATION

An additional component of this research involved the development of a procedure for estimating the cycle length that would result from a given set of geometric, traffic, and timing parameters. This procedure allows direct comparisons of the current operating cycle length with Webster's minimum delay cycle length for fixed time signals. Additionally, the operating cycle length was compared against the minimum saturated cycle length. The methodology utilized is described in detail in the following chapter.

The degree to which the traditional theoretical models for determining the expected actuated cycle length correspond to the performance of actual field hardware was analyzed, and a calibrated model was developed. This section briefly presents the models considered for the determination of the expected extension time of an actuated traffic signal. Note that the minimum green times and clearance times for the critical phases would be added to this expected extension value in order to ascertain the expected cycle length for later comparisons. The basic models are presented in this section, but the complete procedure followed is described in the following chapter.

Erlang Model

The Erlang model used in this research calculates an expected value of extension based on an Erlang distribution of arrival headways. This arrival distribution was initially developed by A.K. Erlang, who did extensive work in queuing theory for application to congested telephone networks (9). The same distribution is frequently used in vehicular traffic modeling as well, and this report will show its application to traffic signal operations.

The general form of the Erlang extension model is as follows:

$$\mu(w)_a = \frac{e^{aqT} - \sum_{i=0}^{a} \frac{(aqT)^i}{i!}}{q \sum_{i=0}^{a-1} \frac{(aqT)^i}{i!}} \tag{8}$$

where

$\mu(w)_a$	=	expected mean extension value, sec;
q	=	flow rate, vehicles/sec;
T	=	study time period, sec; and
a	=	shape factor (*10*).

Typically, values of "*a*" considered range from one to four. The resulting forms of the Erlang equation for a = 1, 2, 3, and 4, respectively, are as follows: (*10*)

$$\begin{aligned} \mu(t)_1 &= q^{-1}(e^{qT} - 1 - qT) \\ \mu(t)_2 &= \frac{e^{2qT} - 1 - 2qT - 2(qT)^2}{q(1 + 2qT)} \\ \mu(t)_3 &= \frac{e^{3qT} - 1 - 3qT - 4.5(qT)^2 - 4.5(qT)^3}{q[1 + 3qT + 4.5(qT)^2]} \\ \mu(t)_4 &= \frac{e^{4qT} - 1 - 4qT - 8(qT)^2 - 10.7(qT)^3 - 10.7(qT)^4}{q[1 + 4qT + 8(qT)^2 + 10.7(qT)^3]} \end{aligned} \tag{9}$$

The minimum value of a is one, which represents a negative exponential distribution. Increasing values of a represent distributions with increasing symmetry and concentration about the mean value (9).

Akcelik-Rouphail Model

Other distributions can also be used to represent the arrival headway distribution, thus altering the calculation of the expected extension time. The Akcelik-Rouphail model is based on the assumption that, for single-lane approaches, headways conform to a shifted negative exponential distribution *(11)*.

The general form of the Akcelik-Rouphail model is written as follows:

$$\mu(w) = he^{\frac{UE-\Delta}{h-\Delta}} + \Delta - h \qquad \textbf{(10)}$$

where

$\mu(w)$	=	expected mean extension value, sec;
h	=	average arrival headway, sec/veh;
UE	=	unit extension or passage gap, sec; and
Δ	=	minimum headway, or negative exponential shift, sec/veh *(11)*.

Although an accepted model for some traffic engineering applications, the assumption of single-lane approaches does not allow the shifted negative exponential distribution to be appropriately applied to the primary testbed used in this research.

SIMULATION TECHNOLOGIES

In order to allow for feasible and cost effective testing of a variety of intersection configurations and test scenarios, computer traffic simulation was used in this research. Actual field implementation would have been impractical and cost prohibitive, and simulation allows other factors and environmental variables that are not being studied to be held constant, which is not possible in field trials.

Two types of traffic simulation were used in order to conduct this research. The primary characteristics, benefits, and disadvantages of each are described in this section.

Traffic Software Integrated System (TSIS)

The Traffic Software Integrated System (TSIS) package was used as the simulation tool for each type of computer traffic simulation conducted. This section briefly describes the TSIS package, with its specific application to each type of simulation described in the following two sections.

TSIS has been developed by ITT Systems and Sciences Corporation for the U.S. FHWA, and is a combined suite of utilities designed to aid both practicing traffic engineers and researchers. The main component of TSIS is CORSIM, which is a microscopic computer simulation program capable of modeling freeways (FRESIM) and surface streets (NETSIM) (*12*). NETSIM is the component of interest in this research, as it can model signalized intersections. NETSIM was initially released in 1971 and was subsequently updated throughout the 1970s. In the early 1980s, NETSIM was incorporated into the integrated traffic simulation system called TRAF, marking a significant milestone in the development of its features and capabilities (*13*).

TRAF-NETSIM models the operational performance of individual vehicles in detail, uniquely determining the state of each vehicle on a second-by-second basis. It determines the performance of a given vehicle as a function of several parameters, including vehicle category (auto, carpool, bus, or truck), type (specific operating and performance characteristics), and driver behavior (passive, normal, or aggressive). Each vehicle is represented relative to a network of unidirectional links and nodes, with all changes in direction occurring only at the nodes (*13*).

The simulation collects a wide variety of measures of performance, including speed, volume, density, delay, spillback, queueing, and turning movements, as well as estimates of fuel consumption and emissions (*13*).

Software-Based Simulation

First, accelerated-time simulations were performed using software only, specifically using the TSIS package. This software-only package has the ability to run as quickly as the processing capacity of the host computer permits. For the simple intersection testbeds that were studied, a common personal computer (PC) is able to simulate 15 minutes of traffic flow and signal operation, while collecting the relevant operational data and measures of effectiveness, in just a couple of minutes. This capability allows for a great deal of flexibility, as multiple simulation runs can be processed in a reasonable time. Additional scenarios could be tested, and additional runs of each scenario could be simulated because of the speed with which the software could carry out the simulations. Multiple runs improved the reliability of the results due to the smoothing of random variations in the simulated traffic patterns.

The most notable disadvantage of software-only simulation is that the entire system being tested, including the vehicle characteristics, driver behavior, signal operation, and controller processes, are operated by the same software running on a multifunction computer processor. This software is not directly related to the software that runs the highly customized traffic controllers used in the field. There are some advanced features that are built into the specialized traffic control hardware that are difficult to replicate in software on a PC, and there are some processes that can be executed more rapidly on a PC than on the standard field traffic controller. This difference means that traffic controller performance and behavior that are seen in software-only simulation may not directly translate to hardware controllers.

Hardware-in-the-Loop Simulation

A second and presumably more realistic simulation option is the hardware-in-the-loop simulation system. This system incorporates a real controller with software-simulated traffic and signals. The software simulation generates traffic, modeling the movement of the simulated vehicles through the network, and reporting the actuation of the simulated detectors to the controller. The controller reads the detector inputs and operates the signals according to the control parameters set within the controller. The signal indications set by the controller are returned to the computer simulation, where they are displayed to the simulated traffic.

Although the TSIS software package is commercially available, hardware-in-the-loop systems are found only in select research environments. The system used in this study has been developed by the Texas Transportation Institute's (TTI) TransLink® Roadside Equipment Laboratory (REL) (*14*).

The TTI REL hardware-in-the-loop system used in this research utilizes the TRAF-NETSIM model for the software aspects of the system, allowing for relatively easy and reliable transfer of files from software-only to hardware-in-the-loop simulation. This common platform also permits direct comparisons of simulation results between simulators.

The controller interface device (CID) serves as the bridge linking the computer with the controller, allowing the controller to behave as though it is hard-wired into a traffic signal cabinet. Several variations of CIDs have been developed by researchers throughout the country. The system used in this research was developed and built by the TTI REL and has been used extensively for a variety of research applications. Figure 6 shows a REL CID and NEMA controller setup. The I/O interface board shown in the bottom right corner of the picture is inserted into a personal computer, which can then run the simulation software.

As was done in this research, it is preferable for the I/O card to operate on one computer, termed the "client," while the simulation software operates on another machine, termed the "server." The client and server are linked by a network to permit the transfer of data. This arrangement allows more controllers to be connected to a simulation than would be possible if each required the installation of an I/O card into an expansion slot, as personal computers have limited numbers of available expansion slots. It also separates the I/O processing and simulation processing, which ensures that the computer is capable of executing all necessary functions at once. This capability is important in hardware-in-the-loop simulation, as all aspects of the process must be executed in real-time in order for the controller and simulation to remain in sync.

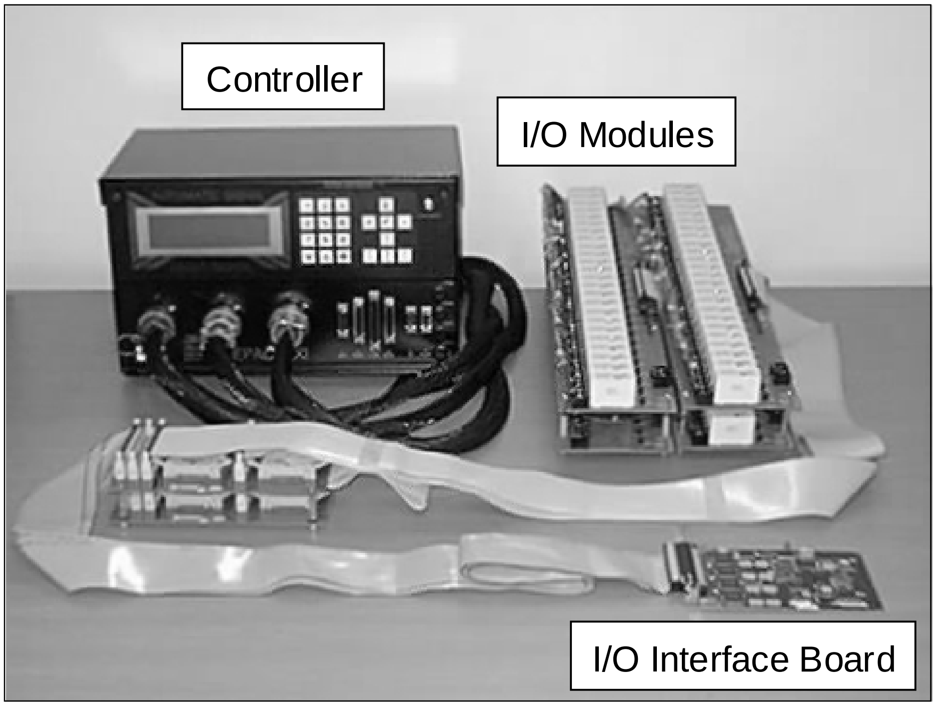

Figure 6. TTI REL CID and Eagle Controller *(15)*.

The complete hardware-in-the-loop setup available in the TransLink® Roadside Equipment Laboratory at the time of this research is shown in Figure 7. Note that the number and variety of controllers available allows for simulation of larger networks than was used for this research. A NazTec Series 900 NEMA TS2 signal controller (shown in the center of Figure 7, to the immediate left of the monitor) was used for all hardware-in-the-loop simulations performed as part of this research.

Hardware-in-the-loop simulation has the advantage of better reflecting the features and operating characteristics of a live controller in the field. Since the actual field hardware is used in the laboratory as a component of the simulation, scenarios tested in simulation can be immediately and directly implemented, provided that the other associated hardware, such as communications and detection, are available. The use of a real controller also enhances the realism of the simulation process and allows for more certainty in the results produced.

A major disadvantage of this arrangement is that, since a real controller is used, simulations are restricted to running in real time. This restriction significantly limits the number of possible scenarios that can be practically simulated, and so greater care is required in selecting the configurations and parameters that will produce the most useful results.

Figure 7. TransLink® Roadside Equipment Laboratory.

3. STUDY METHODOLOGY

This chapter describes the methodology that was followed in the development of an actuated traffic control process utilizing real-time estimated volume feedback. The methods used for simulation and testing of the fundamental concepts, the specific test scenarios that were constructed, and the procedures used to compare and evaluate the results are also described in this chapter.

INTERSECTION TESTBEDS

Two hypothetical testbed intersections were developed for simulation study using the hardware and software described in the preceding chapter. Each intersection was simulated under a variety of traffic loading conditions and passage gap settings, and multiple runs were conducted for each configuration, with the results averaged to mitigate the random variations inherent in the microscopic, stochastic simulation process. The testbed intersections and conditions simulated are described below.

Basic Intersection Geometry

The basic study intersection served to validate the underlying theories only and was not intended to be representative of real-world conditions. This simple testbed had two conflicting approaches crossing at a right angle, with balanced approach volumes and no turning movements. There was one approach lane on each leg, and traffic was controlled by simple two-phase actuated signal operation. Detection for this configuration was accomplished by use of a single inductive loop (6 x 60 ft) located at the stopline on both approaches. The reasons for the selection of a 60-foot loop detector are explained in "Detector Section" later in this chapter. A sketch of this basic intersection is shown in Figure 8.

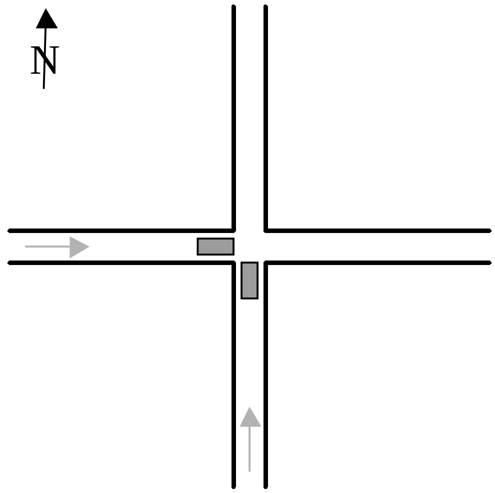

Figure 8. Basic Isolated, Actuated Intersection Testbed.

Advanced Intersection Geometry

The other hypothetical intersection was a more advanced isolated, actuated intersection. It consisted of four, symmetric approaches, with balanced approach volumes and equal turning movements. The traffic entering each approach was divided such that 70 percent went through, 15 percent turned left, and 15 percent turned right. A combined through and right-turn lane, a dedicated through lane, and an exclusive left-turn pocket of adequate storage length were provided on all approaches.

Detection for this intersection was also accomplished by means of inductive loop detectors (6 x 60 ft) located at the stopline in each lane. Detectors in the right and middle lanes were considered as a single detector calling the through and right-turn phase, while the detector in the left-turn pocket operated independently and drove the left-turn phase. A sketch of this intersection is shown in Figure 9.

The signal followed the traditional NEMA eight-phase, dual-ring phasing pattern with leading exclusive left-turn phases on all approaches. This common phasing plan is illustrated in Figure 10.

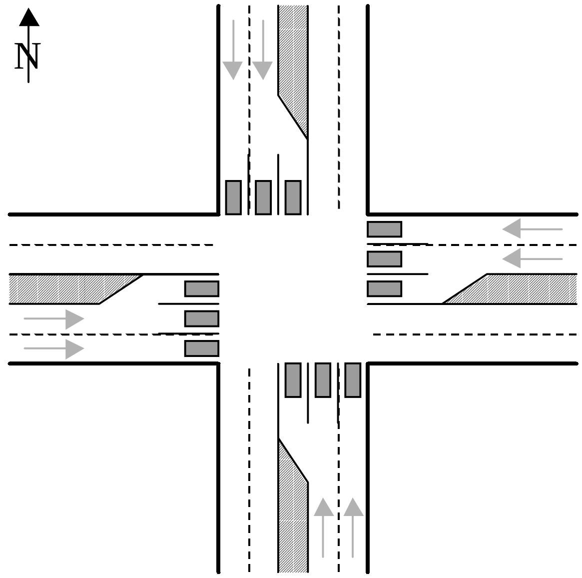

Figure 9. Advanced Isolated, Actuated Intersection Testbed.

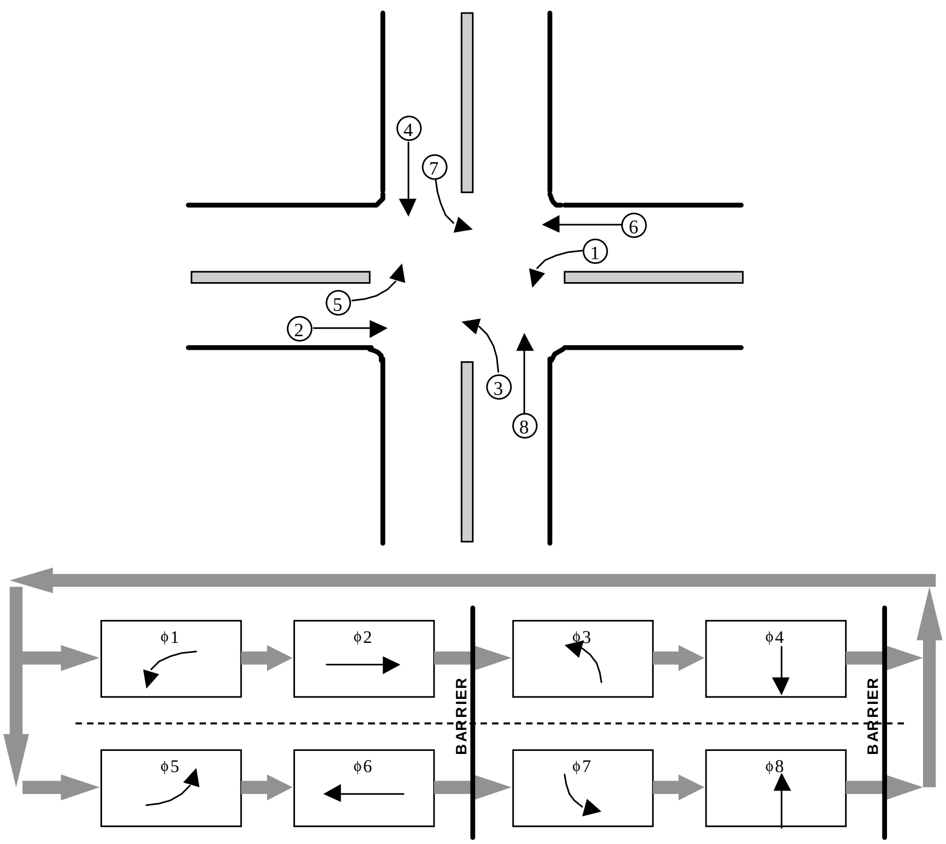

Figure 10. Eight-Phase, Dual-Ring, Quad-Left, NEMA Signal Phasing Plan.

Detector Selection

Both testbed intersections used inductive loop detectors (6 x 60 ft) at the stopline across all approach lanes. Stopline detectors, which are in widespread use across the country, were the only source of input for the simulations beyond the provision of reasonable minimum and maximum phase times. For the approach speeds simulated, which were 35 mph, some form of extension detector would generally be included in a field implementation. However, these detectors were not included in order to simplify the design and prove the fundamental volume estimation concepts being researched.

Note that, in order to validate the volume estimation equation described previously, it was necessary to maintain continuous actuation of the stopline detector while the queue was discharging. Analysis of vehicle dynamics showed that 60-foot long loops were needed in order to ensure that the loop in CORSIM is continuously occupied while a queue is discharging. This is illustrated graphically in Figure 11 and described in detail below.

The solid, curved lines represent the (front bumper) paths of queued vehicles following green onset. The increasing delay in the start of forward motion of the queued vehicles represents the combination of an initial perception-reaction time and the effects of the shock wave as it propagates backward through the queue.

The vertical dotted lines represent the upstream and downstream edges of the detection zone created by a 60-foot inductive loop detector located at the stopline, including the length of the vehicle on the downstream end. The length of the vehicle was also included on the downstream end, as detector actuation would continue until the entire length of the vehicle was completely off the detector. The horizontal dotted lines indicate the time at which each queued vehicle, stopped upstream of the detector, enters the detection zone. If a vehicle enters the detection zone before the previous vehicle has left, then both vehicles would be detected and there would be continuous actuation of the detector. This would be represented by a horizontal dotted line that begins when a given vehicle enters the detection zone and intersects the path of the previous vehicle in the queue.

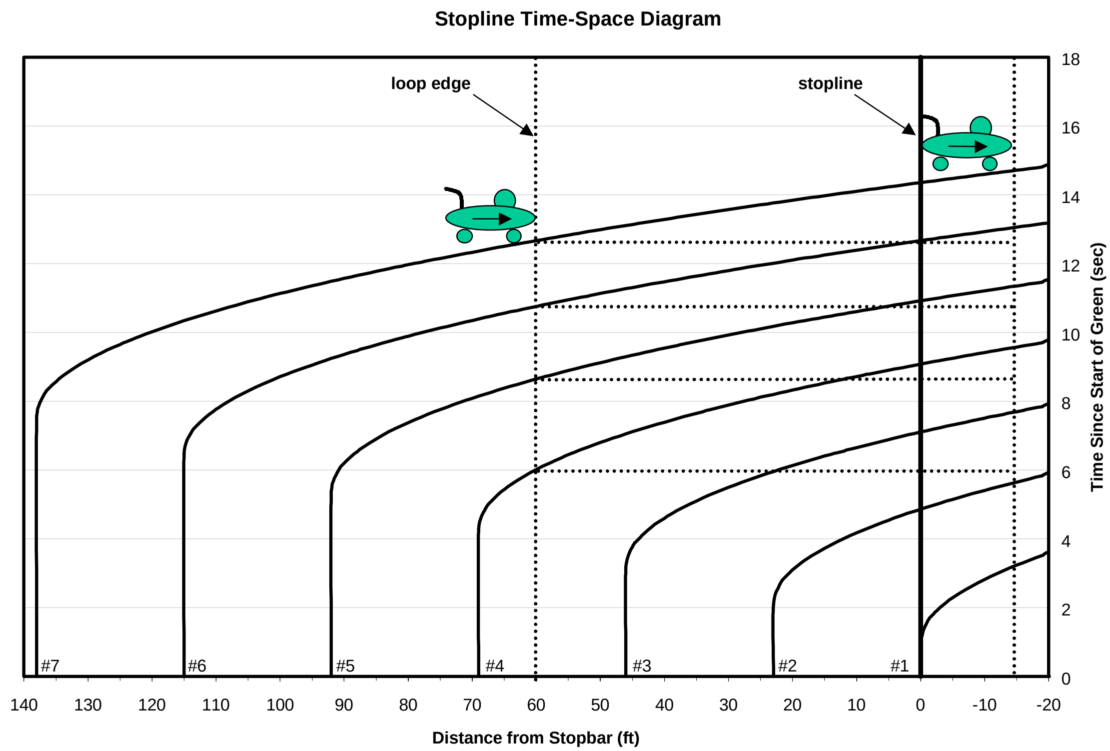

Figure 11. Queued Vehicle Trajectories at Green Onset.

Note that visual inspection of the simulation operations confirmed that the simulated detector was under continuous actuation while the queue was discharging.

EXPERIMENT STRUCTURE

This section describes the structure and organization of the experiments performed for this research, as well as the process by which simulation output data was reduced.

Range of Scenarios

Since the software-based simulation method allowed for faster execution of scenarios, additional runs were made beyond those conducted using the hardware-in-the-loop technique.

Scenarios were simulated using the TSIS software only for passage gap settings of 1.1, 2.0, 3.0, 4.0, and 5.0 seconds. Each gap setting was tested under five unique volume conditions. For three volume conditions, ten replications of each scenario were conducted with different random number seeds. In addition, four replications of the remaining two volume levels were conducted, using random number seeds that match those used in the hardware-in-the-loop replications of the same conditions.

Under hardware-in-the-loop simulation, three volume levels were tested, using the same range of gap settings described previously. Four replications of each scenario were performed using different random number seeds.

Organization of Data

Due to the large number of simulations performed under various methods, organization of the data, both input and output, was an important concern. This section describes the naming convention that was developed and used consistently throughout this research.

First, it is important to note that TSIS uses several types of electronic files. The extension of each file indicates its type and purpose, as follows:

.*TRF* = TRAF-NETSIM input file;
.*OUT* = TRAF-NETSIM output file; and
.*SDI* = computer-controller interface setup file.

One of each of these files is required for each scenario and each replication tested, with the third file type (.*SDI*) used only for hardware-in-the-loop simulation.

All files associated with a particular simulation run were identically named, with the extension indicating the file type. Each character, or group of characters, in the filename identifies a characteristic of the particular run, as described in the following paragraphs.

The first character indicates whether it is a software-only (*s*) or hardware-in-the-loop (*h*) run.

The second character indicates whether the intersection conforms to the basic (*b*) or advanced (*a* or *s*) geometry, as described previously.

The third character represents the volume conditions, as well as whether or not the gap adjustment is balanced. For both hardware and software runs, the following key was applied:

a = unbalanced, high volume (V=1300 vph per approach);
b = balanced, high volume (V=1300 vph per approach);
c = unbalanced, medium volume (V=1000 vph per approach);
d = balanced, medium volume (V=1000 vph per approach);
e = unbalanced, low volume (V=600 vph per approach);
f = balanced, low volume (V=600 vph per approach);
1 = balanced, high volume (V=1300 vph per approach);
3 = balanced, low volume (V=300 vph per approach); and
8 = balanced, medium volume (V=800 vph per approach).

The slight variation in naming scheme is the result of changes in experiment strategy that were made while the research was underway in order to ensure that the widest possible range of meaningful conditions was tested.

The fourth and fifth characters represent the passage gap setting (x10) of the run.

The sixth, seventh, and eight characters represent the replication number of the simulation run. All runs with the same replication number use the same random number seed for the generation of traffic.

For example, the file *ha330r04.TRF* is an input file for a hardware-in-the-loop simulation of the advanced geometry intersection, with a volume of 300 vph per approach and a passage gap setting of 3.0 seconds on all phases. This would be the fourth replication of this scenario, and, as such, would have identical traffic patterns to the run *sa330r04.TRF*, thereby allowing direct comparison of results between the two simulation methods.

Output Processing

All of the measures of performance that TSIS generates during the course of a simulation run are output in a single text file (*.OUT*), along with a listing of the input file that can be used for verification and reference. Also, several of the parameters of interest in this research, including phase length, cycle length, and green splits are not explicitly output by the software. In order to obtain these values, a second-by-second listing of the signal indications for all phases is recorded in the output file. For a typical 15 minute run, this produces 900 line entries, composed of a series of ones and zeroes representing the signal indications and permissive movements, and it also includes occasional identification lines and notices of spillback and spillback clearance. Consequently, this is a somewhat cumbersome file from which to extract useful results data.

In order to facilitate the reduction of this data, a spreadsheet was created to sort and interpret the signal indication data, and to collect the other pertinent data. A unique copy of this large spreadsheet was generated for each run conducted throughout this research, and then summaries were generated from these files. These summaries provided the data needed for the analyses described in the remainder of this chapter.

COMPARISON OF HARDWARE- AND SOFTWARE-BASED SIMULATIONS

A large set of scenarios was simulated using software-only simulation, and a subset of these scenarios was tested using the hardware-in-the-loop simulation equipment. Since both simulation methods rely on TSIS for the generation of traffic patterns, it is reasonable to compare the hardware and software results to determine the realism of the software-only simulation. The accuracy with which the software replicated the true field hardware is an important consideration in determining the usefulness of this technology for applications of this type, which depends heavily on controller behavior.

Statistical comparisons between hardware- and software-based results were made, despite differences in sample size, by evaluation of the standard deviation of the difference in the means. This was calculated as follows:

$$\hat{s} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad \textbf{(11)}$$

where

$\hat{s}$ = standard deviation of the difference of the means;

s_1, n_1 = standard deviation and number of observations in study 1; and

s_2, n_2 = standard deviation and number of observations in study 2 (*16*).

The statistical significance of any difference between the two methods of simulation was then established using a standard z-test for statistical significance, with the test statistic calculated as follows:

$$z = \frac{\mu_1 - \mu_2}{\hat{s}} \quad \textbf{(12)}$$

where

z = normal distribution test statistic;

μ_1 = mean value of observations in study 1; and

μ_2 = mean value of observations in study 2 (*17*).

Simulation output statistics were tested at a 95 percent confidence level (α = 0.05). The results of this test procedure are discussed in the following section.

VALIDATION OF VOLUME ESTIMATION METHODOLOGY

A primary goal of this research was to validate the volume estimation equation developed by Messer et al. Described in the preceding chapter. The accuracy with which this equation can estimate the actual traffic volumes is very important, as this is the primary source of input for the signal control process developed in this research.

One would expect that, if the methodology holds true, the estimated values would match the actual traffic volume values over the range of scenarios tested. Clearly, however, it is unlikely that the methodology can exactly estimate the demand volume in all cases, but it should be reasonably close for the majority of the time. The statistical equivalence of the estimated and actual values was tested to a 95 percent confidence level using a traditional z-test. As such, it was assumed that volume levels would be normally distributed about the mean.

It is also important to note that the actual traffic volumes used were output from the simulation software for all scenarios. The input volume levels were not used as the simulation randomly assigns traffic to each approach, and it is unlikely that the exact volume established as input would be generated during the study period on all approaches.

CYCLE LENGTH ADJUSTMENT

The other primary goal of this research was to develop a process for using passage gap setting to adjust the operation of a traffic signal to achieve the desired volume-capacity ratio. Since volume-capacity ratio cannot be directly controlled, the ability to use passage gap to alter the operating cycle length and relative green splits among phases must first be established. This section describes the process used in determining how the cycle length was adjusted in order to approach the target operating state.

Since it is unrealistic to expect to reliably operate an actuated signal at an exact cycle length under constantly changing volume conditions and randomly distributed traffic patterns, a range of desirable operating cycle lengths was defined.

For efficient operation, the lower bound of the desired operating range was selected to be the saturated cycle length. This means that one would never want to operate a signal with a volume-capacity ratio greater than one. The upper bound of the desired operating range was selected to be the minimum delay cycle length, as defined by Webster *(8)* and described in Chapter 2, which would represent the minimum overall delay at the intersection if the signal were operated under a fixed timing plan. Consequently, the goal is for the operating cycle length for any given cycle to be between these upper and lower bounds, and to operate as close as possible to the minimum delay cycle length.

Average Operating Cycle Length

The average operating cycle length can be obtained by direct measurement from the controller. Again, the value for the previous cycle could be used, or several cycles could be averaged to produce smoother transitions. This value establishes a starting point for any needed adjustments in the operation of the signal.

Saturation Cycle Length

The saturation cycle length represents the lower bound of the desired operating range, and it is the minimum cycle length needed in order to process all of the demand volume at an intersection with no wasted service time. Since no slack time is available, the slightest increase in demand would oversaturate the intersection and result in cycle failures and increased delay to motorists. As such, this value is used as a lower bound, but it is probably not desirable to operate the signal at the saturated cycle length.

Webster's Minimum Delay Cycle Length

The minimum delay cycle length for fixed time operation represents an upper bound of the desired operating range. The derivation of this value was explored in the preceding chapter.

Target Cycle Length

At this point, the appropriate passage gap setting must be found that, based on the results of this research, could be expected to produce a target cycle length within the desired range as described above. Therefore, it was necessary to have a reservoir of data correlating cycle lengths and passage gap settings for a variety of possible flow ratios. This section describes the iterative procedure that was used to calculate an expected cycle length using the models described in the previous chapter. The results are described in the following chapter.

A variety of models were processed for comparison, but the Erlang (a=2) model was selected for calibration to reflect the actual controller performance. This procedure was implemented in a large, interconnected spreadsheet form that allowed for easy alteration of input values while providing for consistent execution of the iterative process. Figure 12 illustrates the process for determining the expected cycle length.

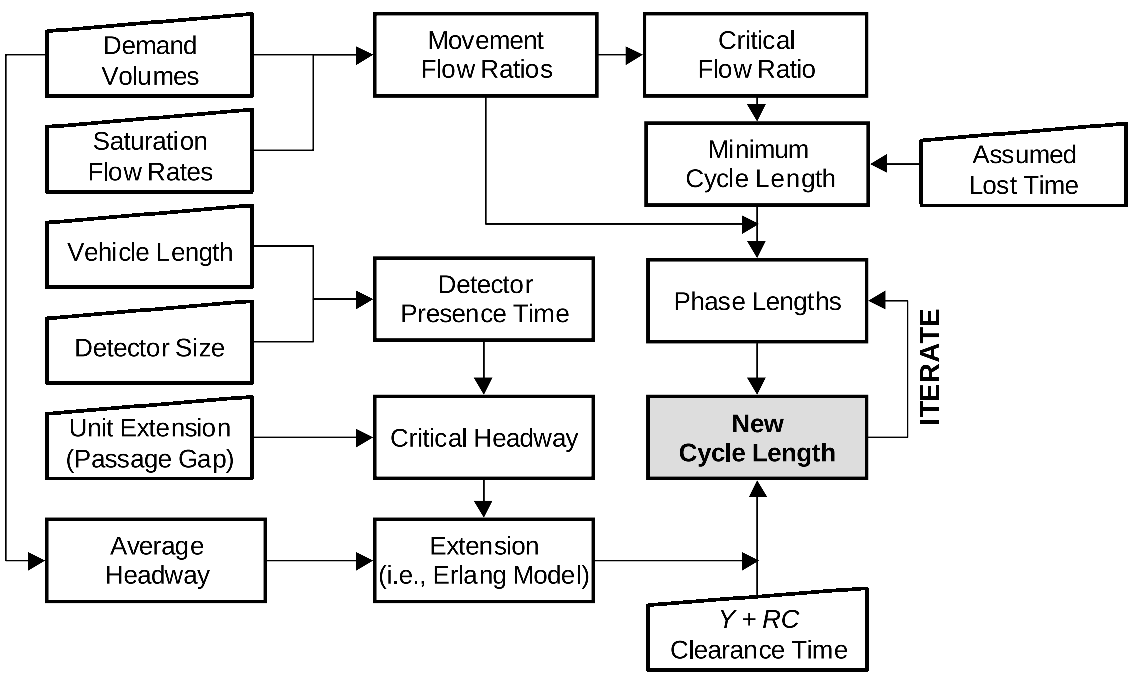

Figure 12. Expected Cycle Length Determination Process.

The primary inputs consisted of the demand volume and number of lanes for each NEMA phase (numbered 1 through 8), as well as an estimated saturation flow rate. In the examples used in this research, a saturation flow rate of 1800 vphgpl was input for all through movements. This saturation flow rate was reduced by a factor of 0.95 for left-turn movements to account for the reduced operating speed and capacity of left-turn maneuvers. A typical vehicle length of 15 feet was also assumed as input and later used to calculate loop presence time.

Using the volumes and saturation flow rates provided, a flow ratio for each movement could be calculated. This flow ratio was used throughout the procedure in order to maintain service times proportional to the demand volumes. The calculation of flow ratios for all movements also allows the critical movements in each phase pair to be determined, and for the sum of critical flow ratios to be calculated.

The sum of startup and clearance lost times was assumed to equal the total clearance time (yellow plus all-red). Therefore, any startup lost time caused by driver perception-reaction time and vehicle acceleration at green onset was assumed to be equally offset by effective service time gained during the yellow interval. Consequently, the effective green time approximately equals the displayed green time. The total lost time for the critical phases was then summed in order to find the total lost time per cycle.

Having established the sum of critical flow ratios and the total lost time per cycle, the minimum cycle length needed to serve the demand volume could be found using the following formula:

$$C = \frac{L}{1-Y} \tag{13}$$

where

C = minimum cycle length, sec;

L = total lost time per cycle, sec; and

Y = sum of the critical flow ratios.

This provided a cycle length that served as a base point for further iterations.

Each subsequent iteration then consisted of rationing the total available cycle time in proportion to the previously calculated flow ratios. Minimum green thresholds of 10 seconds for through movements and four seconds for left turn movements were also set to reflect realistic values that would be likely in the field.

The presence time of a vehicle of known length travelling at constant speed over the loop detector was found as follows:

$$P = \frac{l_d + l_v}{v_{85}} \qquad \textbf{(14)}$$

where

P = vehicle presence time on loop detector, sec;

l_d = length of detector, feet;

l_v = length of vehicle, feet; and

v_{85} = 85^{th} percentile approach speed, fps.

The sum of this presence time and the unit extension applied to the phase is the critical headway, as shown below.

$$H = P + UE \qquad \textbf{(15)}$$

where

H = critical headway, sec;

P = vehicle presence time on detector, sec; and

UE = unit extension (passage gap), sec.

This critical headway can be compared against the average headway, which was found by dividing the number of vehicles arriving per hour into the number of seconds in an hour (3600 sec/hr).

A new phase length could then be calculated as the sum of the green time given to the phase, the clearance time (Y + RC), and the expected extension. A new cycle length can then be found as the sum of the critical phase lengths and used as a starting point in the next iteration. Following this procedure, the calculations converge to a stable, expected cycle length in less than ten iterations for all reasonable input values tested. The results of this process are described in the following chapter.

4. STUDY RESULTS

This chapter presents the results of the simulations and subsequent analyses conducted as described in the previous chapter. Comparisons between the theoretical analyses and the simulation studies conducted are provided and explained. A direct comparison of software-only and hardware-in-the-loop simulations is also provided, followed by results validating the volume estimation methodology described previously. The resulting cycle length and green splits produced through simulation are discussed, and the models formulated are described.

COMPARISON OF THEORETICAL ANALYSES AND SIMULATIONS

As described previously, fundamental concepts of traffic science were used to generate expected actuated cycle lengths through an iterative process. This section compares the expected cycle lengths found against the results of both software-only and hardware-in-the-loop simulations conducted for the same volume levels and passage gap settings.

For those conditions that produce short to moderate cycle lengths, the simulated and theoretical cycle lengths were shown to match well, with the majority of points falling within the bounds of a 95th percentile confidence interval. The results of these comparisons are illustrated in Figure 13 and Figure 14 for software-only and hardware-in-the-loop simulation, respectively.

Those conditions that would be expected to produce longer cycle lengths, such as higher volumes and longer passage gap settings, illustrated a key difference between the two methods. The simulation results level out as the cycle length approaches 200 seconds, which is the maximum possible cycle length given the maximum green times allowed for the individual phases. However, since the iterative process did not include such limiting outside factors, the theoretical or expected cycle length continues to climb well beyond 200 seconds, resulting in the sharp curve shown in the figures.

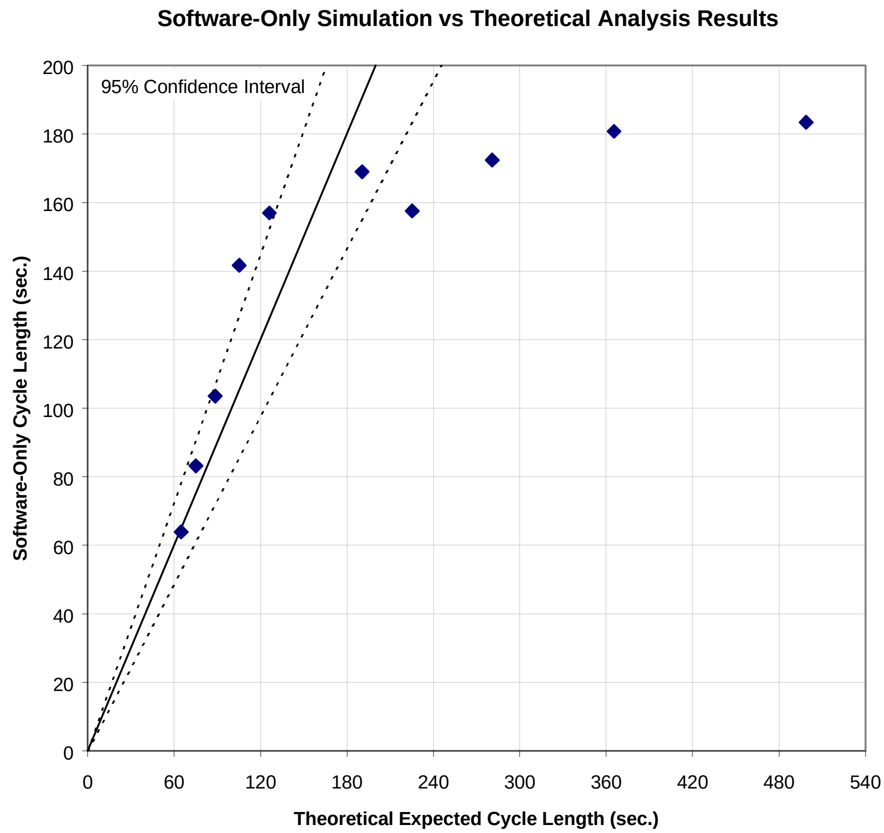

Figure 13. Software-Only Simulation vs Theoretical Analysis Results.

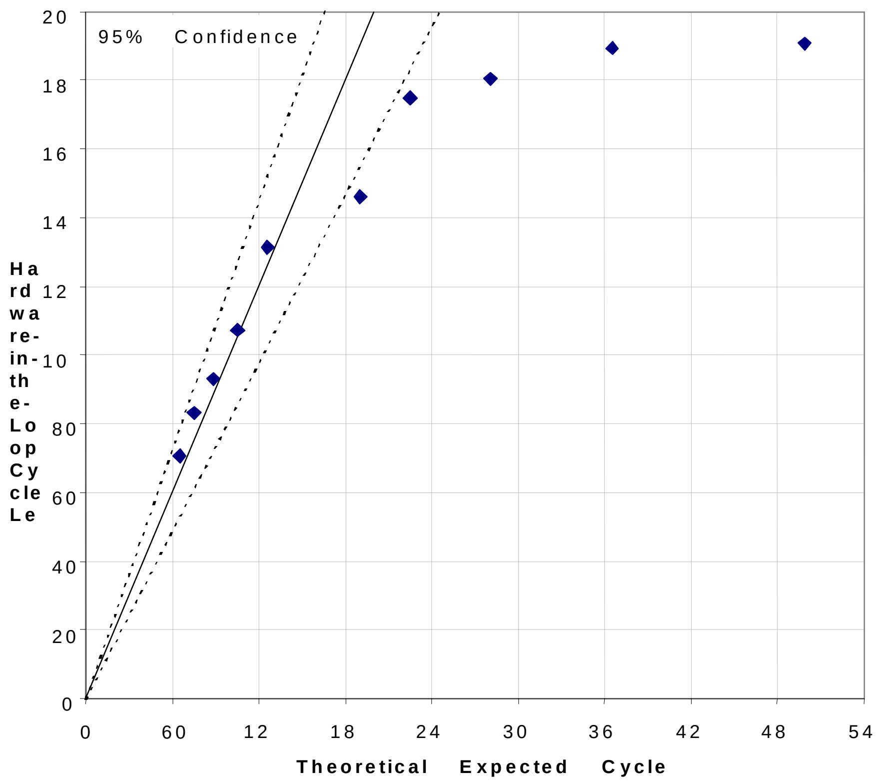

Figure 14. Hardware-in-the-Loop Simulation vs Theoretical Analysis Results.

COMPARISON OF HARDWARE- AND SOFTWARE-BASED SIMULATIONS

Due to the random variations inherent in the simulation process, it is improbable that the cycle lengths produced through software-only simulation would be identical to those produced through hardware-in-the-loop simulation, but it would be expected that the values would be statistically similar for the majority of runs. Additionally, the potential for discrepancies is amplified given the fact that the hardware-in-the-loop simulations were run using a controller that operates continuously and may therefore begin a simulation run at any point in a cycle, whereas the software goes through an initialization process prior to every run.

Given this potential for discrepancy, as described in the previous section, the data points from hardware- and software-based runs with identical geometric and traffic conditions were compared. It is also important to note that, in order to minimize randomness in traffic patterns, each pair of runs used the same initial random number seed for the generation of arrival volumes and distribution patterns.

The results are plotted graphically in Figure 15, with the hardware-in-the-loop cycle length plotted on the y-axis against the software-only cycle length on the x-axis. The solid, diagonal line represents equivalent values, while the dashed lines to either side mark the boundaries of the 95 percent confidence interval about the equivalency line. A regression line placed through the data points is also shown in gray.

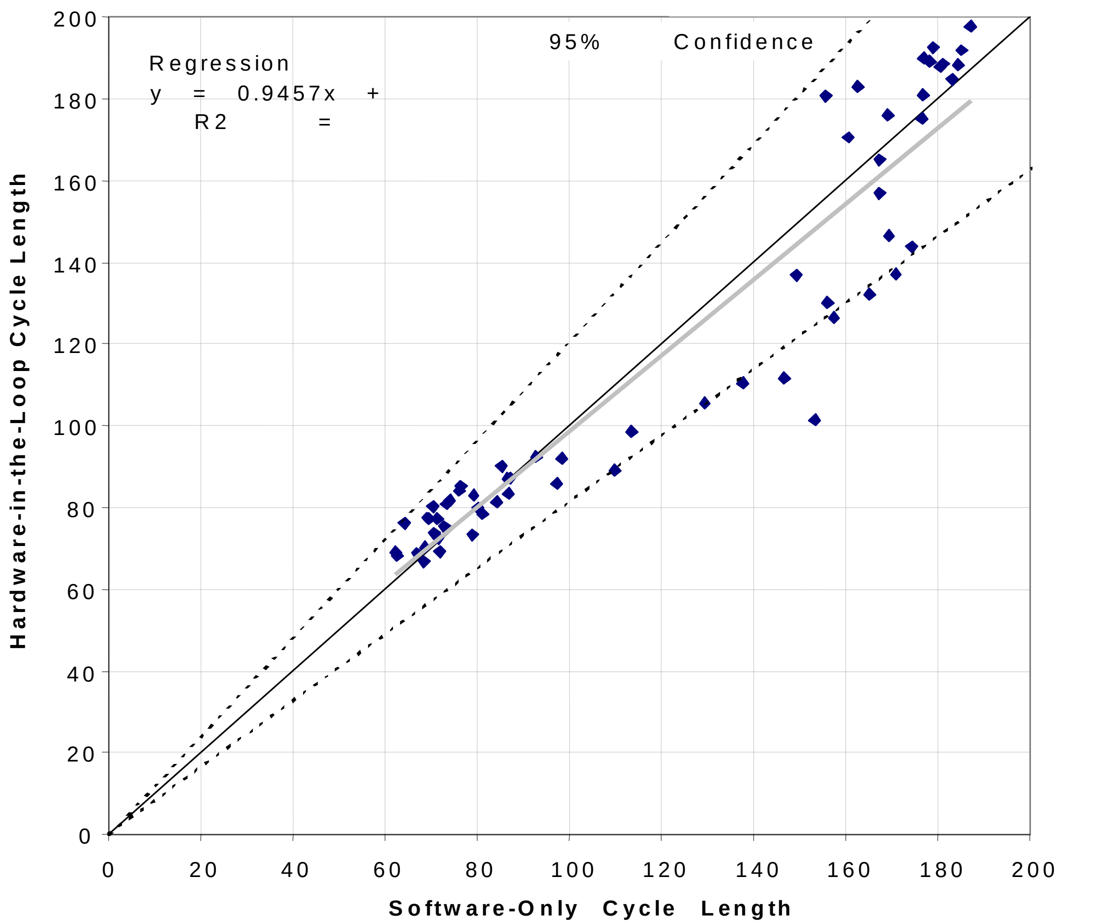

Figure 15. Software-Only vs. Hardware-in-the-Loop Simulation Results.

Although not every pair of simulation runs falls within the boundaries of the 95 percent confidence interval, it is clear that the majority of points do, and that the remainder are quite close. In order to further illustrate the precision of the relationship between the software-only and hardware-in-the-loop cycle lengths, a regression line was plotted through the points representing each pair of runs. This line very closely matches the equivalency line, and is certainly within the 95 percent confidence interval limits.

It is important to note that the full range of configurations that was tested under software-only simulation were not processed using hardware-in-the-loop simulation, and that the statistical equivalence of these two techniques can only be proved for those conditions that were examined using both methods. It is clear, however, that the two simulation methods perform nearly identically in terms of the cycle lengths produced for the range of conditions tested under hardware-in-the-loop simulation.

CHECK OF NETSIM-GENERATED VOLUMES

As discussed previously, TRAF-NETSIM uses random number seeds to generate randomness in the traffic patterns loaded onto the network, in order to more accurately reflect real-world conditions. As such, it is unlikely that the exact demand volume that was specified for an approach in the input file will be produced. One would expect that, on average, the volumes produced should equal the demand volume specified, however.

In order to verify this expectation, the TSIS target volumes, as specified in the input files for each run, were plotted against the actual volumes produced for that run, as reported in the output file. Again, a regression line was fitted through the data points, and the overlay of this regression line on the equivalency line indicates that the volumes produced do represent the volumes specified, as expected. Note that such plots were constructed for each signal phase, to ensure that this held true for the lower volume left-turn movements as well as for the higher volume through movements. Phases one through four are shown in Figure 16, and phases five through eight are shown in Figure 17. Recall that under the standard NEMA nomenclature, odd numbered phases are for left-turn movements and even numbered phases are for combined through and right-turn movements.

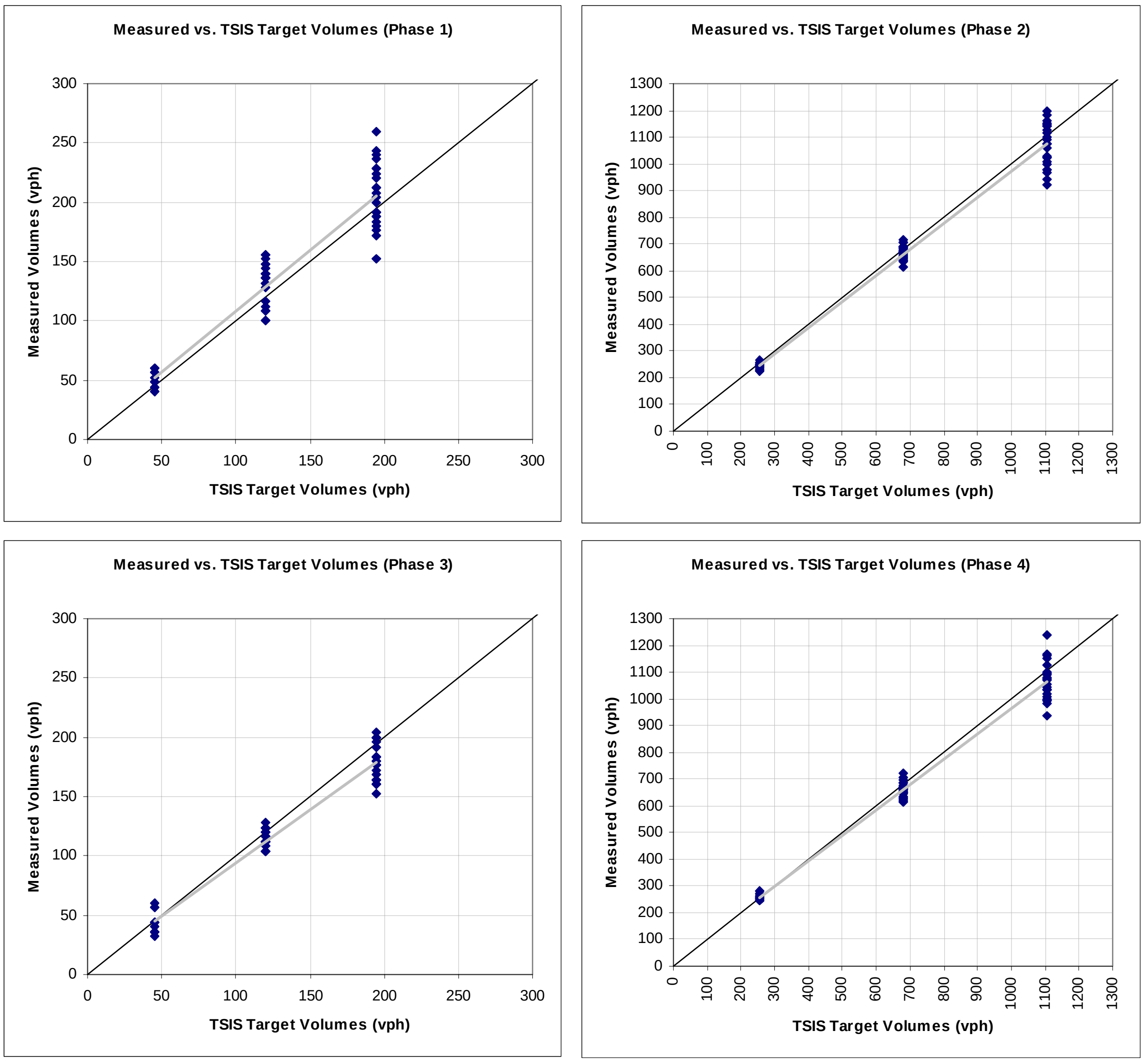

Figure 16. TSIS Target Volumes vs. Measured Volumes (Phases 1-4).

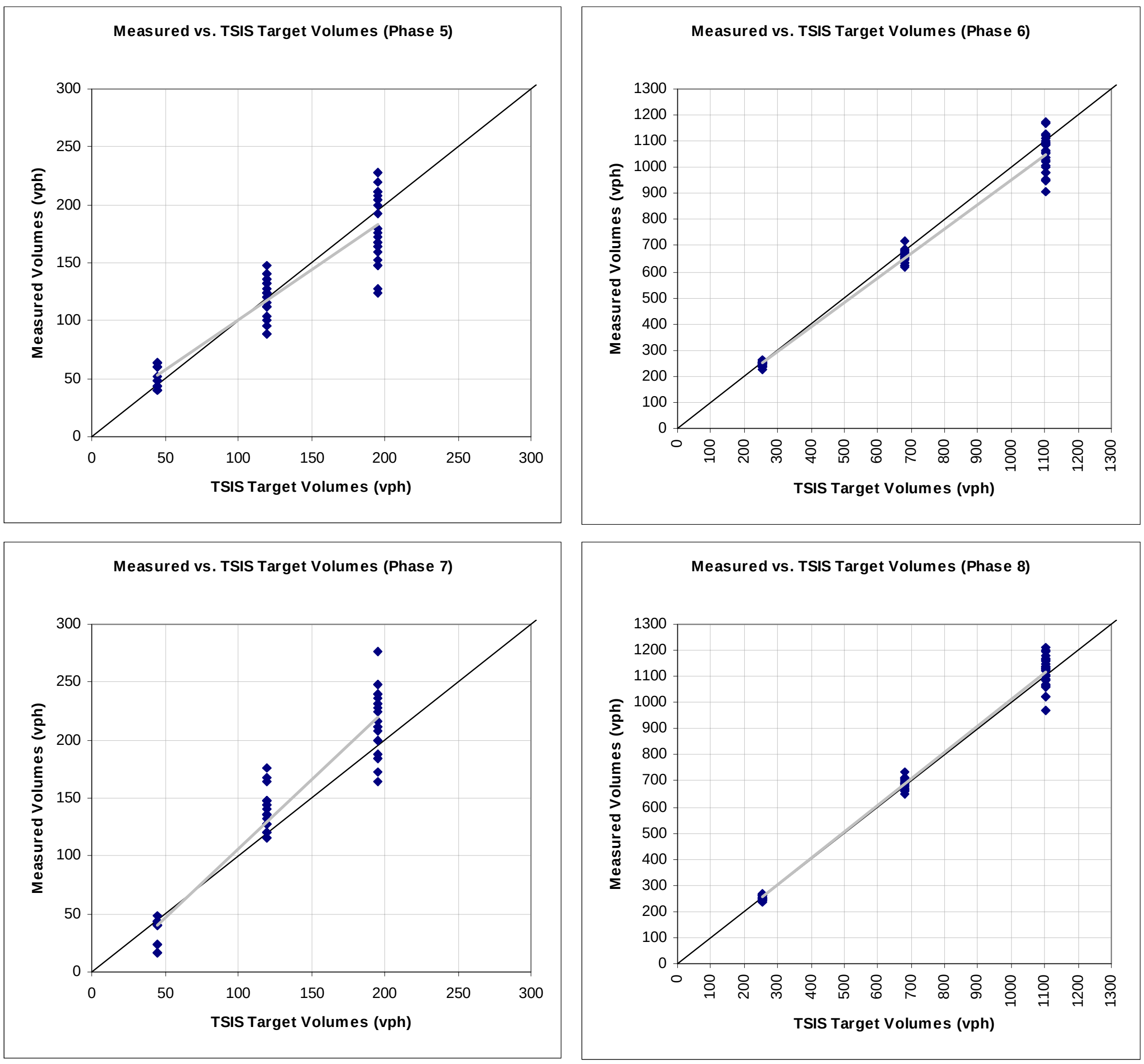

Figure 17. TSIS Target Volumes vs. Measured Volumes (Phases 5-8).

VALIDATION OF VOLUME ESTIMATION METHODOLOGY

Having established the proper operation of the simulation models, the volume estimation methodology described previously can be validated by comparison against the demand volumes reported by NETSIM. The volumes were estimated using the technique described previously, with all analyses performed on a phase-by-phase basis.

In order to calculate an estimated demand volume for a given phase, it was necessary to make two assumptions: saturation flow rate and startup lost time. As in the mathematical calculations for expected cycle length, a saturation flow rate of 1800 vphgpl for the through phases was chosen. This was reduced by five percent to 1710 vphgpl for the left-turn phases for the reasons described previously.

It was necessary to subtract a startup lost time because of the manner by which the saturated green time was measured. The software utility that was used to measure saturated green recorded the length of time from green onset until the stopline detector for a given phase was no longer actuated. However, this overestimates demand volume, as no traffic flow is occurring immediately following green onset. Therefore, a value for startup lost time was assumed and was subtracted from the average saturated green time prior to the calculation of the estimated demand volumes for each phase. This startup lost time was assumed to be 3.0 seconds for all through phases and 1.5 seconds for all left-turn phases, and they were applied equally across all simulation runs. The values chosen, although realistic in magnitude, were selected to serve dual purposes. They reasonably represent the startup lost time, which must be subtracted from the saturated green, as described above, and also are calibration factors in the adjustment of the volume estimation methodology to the specific conditions of the site.

A regression line was plotted through the points representing each individual run in order to show the overall trend. The volumes estimated on the through phases clearly match the actual volumes recorded very well. The volume estimates for the left-turn phases appear to follow a slightly flatter slope than the equivalency line, although the majority of points are still quite close to the actual volumes. This slight discrepancy is likely due to the difficulty of estimating at very low volume levels. Since the left-turn volumes are so low, the cycle-to-cycle variations in arrival volume, saturation flow rate, and startup lost time could be expected to have a detrimental impact on the ability to estimate volumes, particularly when one set of assumed values is applied across many simulation runs.

Phases one through four are shown in Figure 18, and phases five through eight are shown in Figure 19.

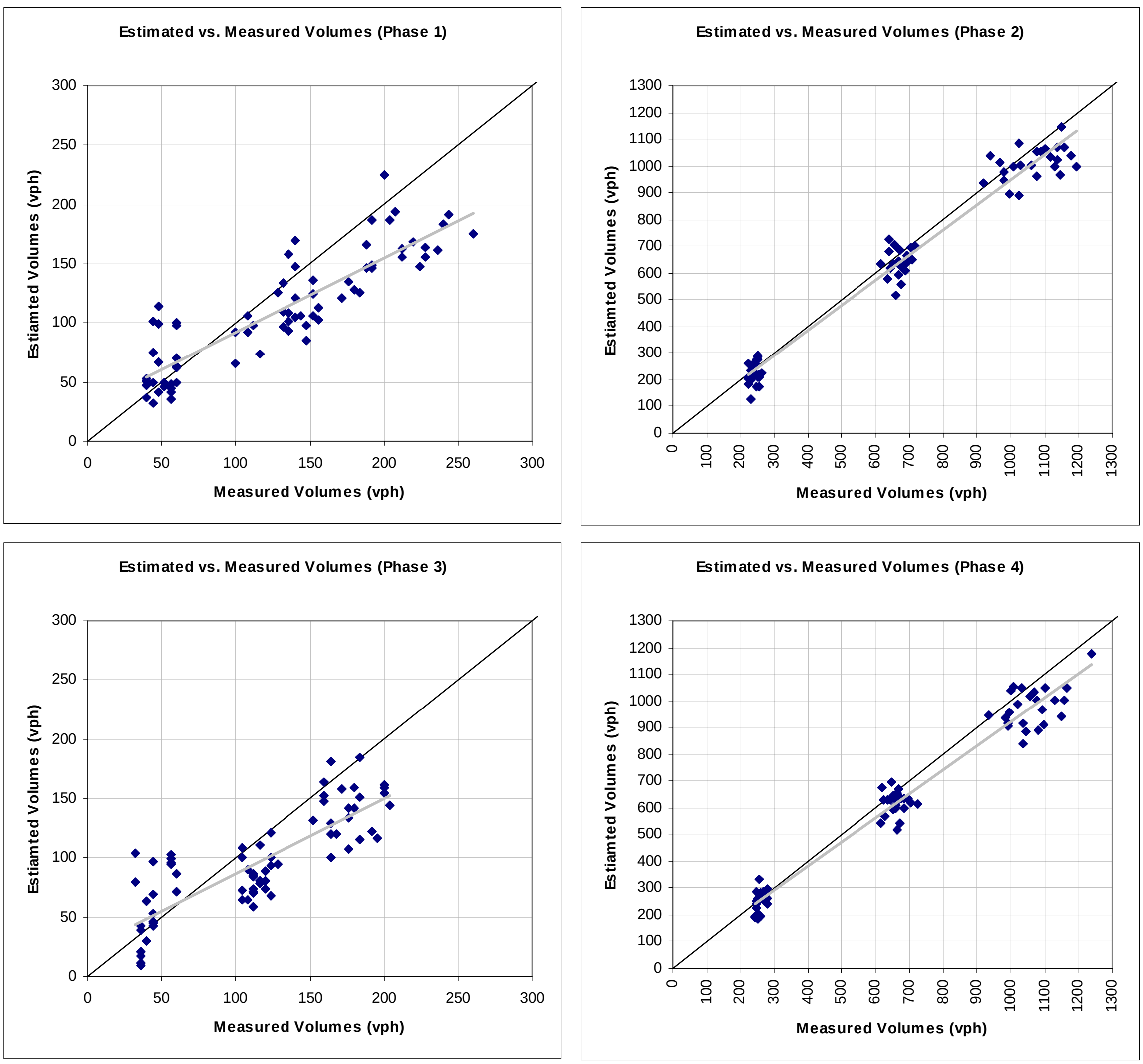

Figure 18. Volume Estimation Methodology (Phases 1-4).

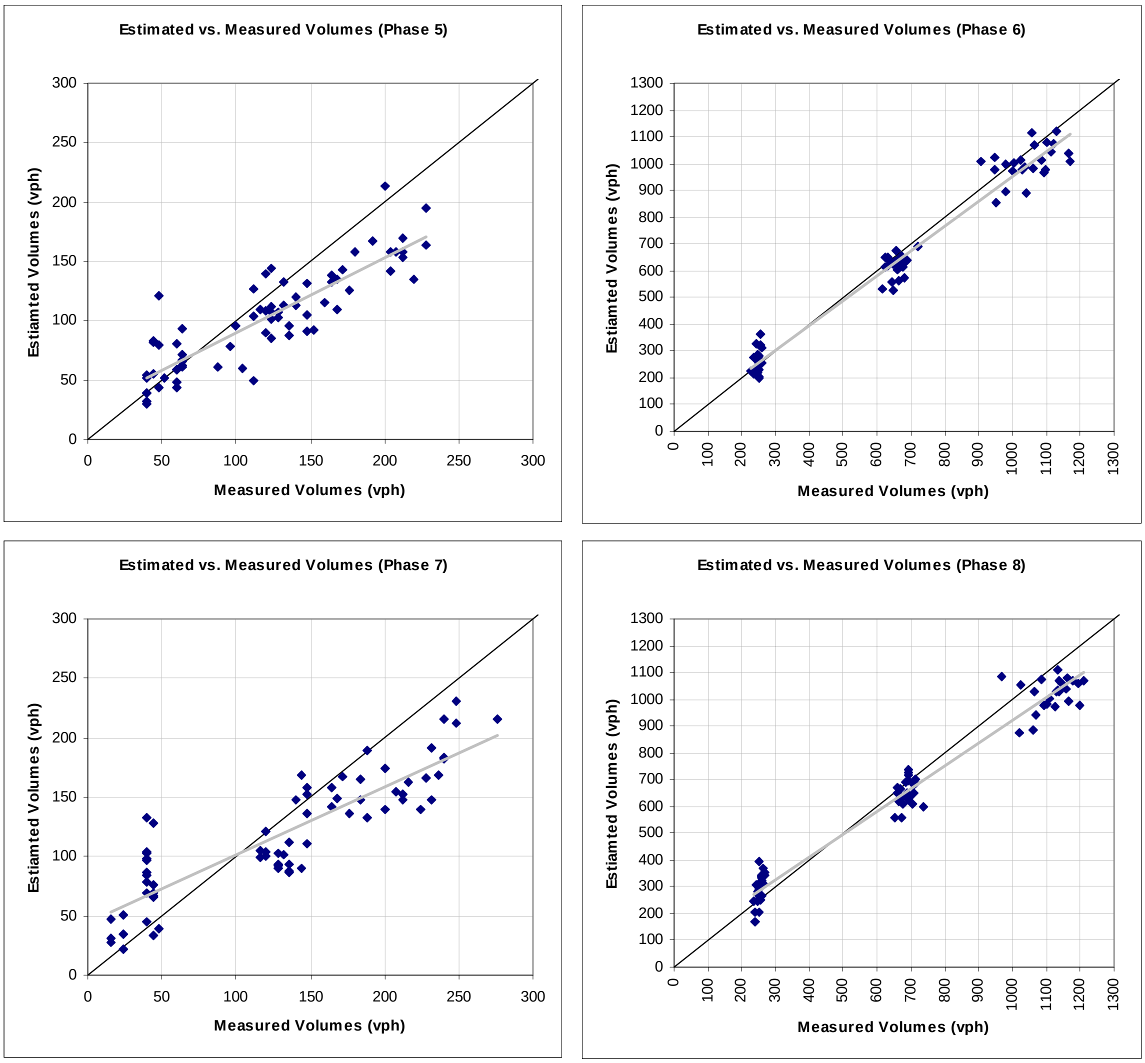

Figure 19. Volume Estimation Methodology (Phases 5-8).

CYCLE LENGTH RESULTS

This section describes the effects of variable passage gap on the resulting actuated cycle length for each of the scenarios tested. As discussed in the third Chapter, both balanced and unbalanced variations of passage gap were conducted. During the balanced simulations, passage gap settings were varied equally on all approaches and all movements. During the unbalanced simulations, passage gap settings were varied only on one approach in the basic testbed, and only on one pair of opposing approaches in the advanced testbed. The results of both sets are described below.

Software-Based Simulation – Basic Testbed

Several trends in cycle length were found in the results of the simulations performed using the basic testbed. Without exception, cycle length increased as the passage gap settings were increased, although, as expected, this increase was more profound for the balanced conditions than the unbalanced conditions. The results are shown in Figure 20.

The most notable feature was a significant increase in cycle length found for the balanced high volume scenario as passage gap was increased. This particular scenario produced unrealistically high cycle lengths, but this result is likely due to the heavily oversaturated conditions and long maximum green values for simple two-phase operation. The capacity of this basic testbed was notably reduced because of the single-lane approaches, as compared to the greater capacity of the two-lane approaches with turn pockets in the advanced testbed. Therefore, the application of such high volumes to a limited capacity facility could not be expected to produce reliable results. However, the other, lower volume scenarios do illustrate the relationships of interest in this research.

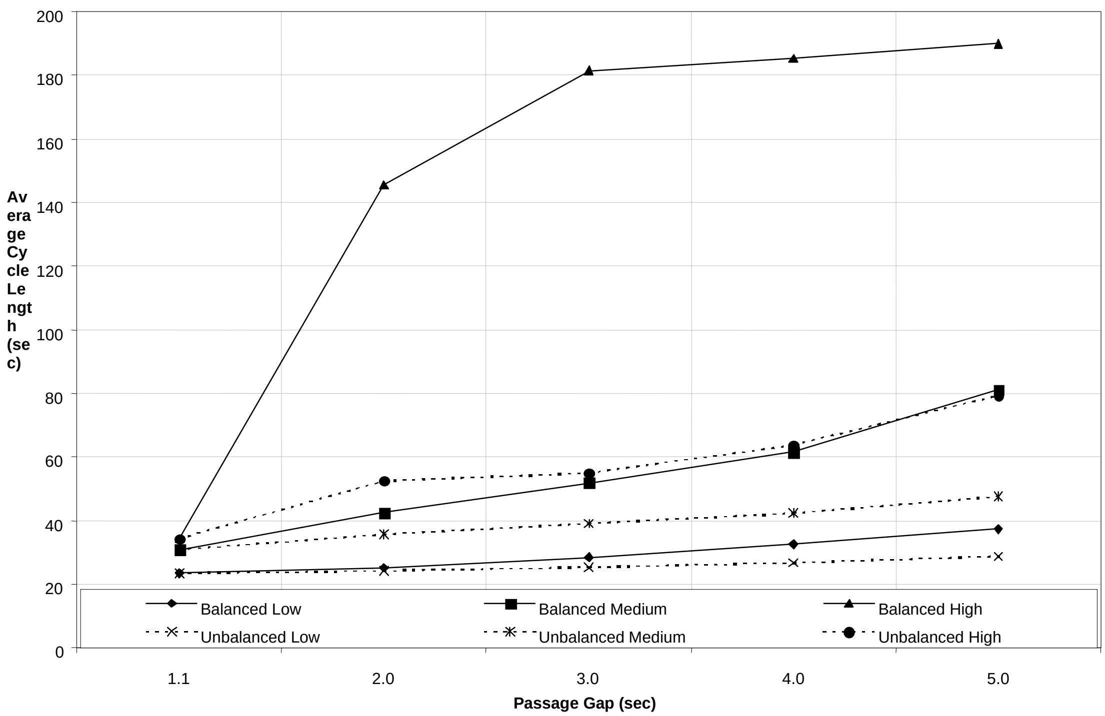

Figure 20. Software-Based, Basic Testbed Cycle Lengths.

Software-Based Simulation – Advanced Testbed

As with the basic testbed, there was a clear increase in cycle length with increasing passage gap for all configurations tested. Again, the balanced simulations consistently increased more sharply than unbalanced simulations. Overall, the range of cycle lengths that resulted was reasonable, ranging from a low of about 65 seconds for the lower volume with a 1.1 second passage gap, to a high of just over 180 seconds for a high volume condition and a passage gap of 5.0 seconds. Figure 21 shows a summary chart of the resulting cycle lengths.

Note that the lower end of cycle lengths produced in the basic testbed was in the range of 20 to 40 seconds, while the advanced testbed produced cycle lengths around 60 to 80 seconds. This difference is likely due to two primary factors. First, left-turn phases are included in the advanced testbed, and so the cycle length is automatically increased by two minimum greens and two clearance intervals. Second, the advanced testbed has two-way traffic on all legs, whereas the basic testbed has one-way traffic only. Therefore, the likelihood of extending and increasing cycle length for through movements is much greater with twice the traffic volume available to extend those phases that run simultaneously.

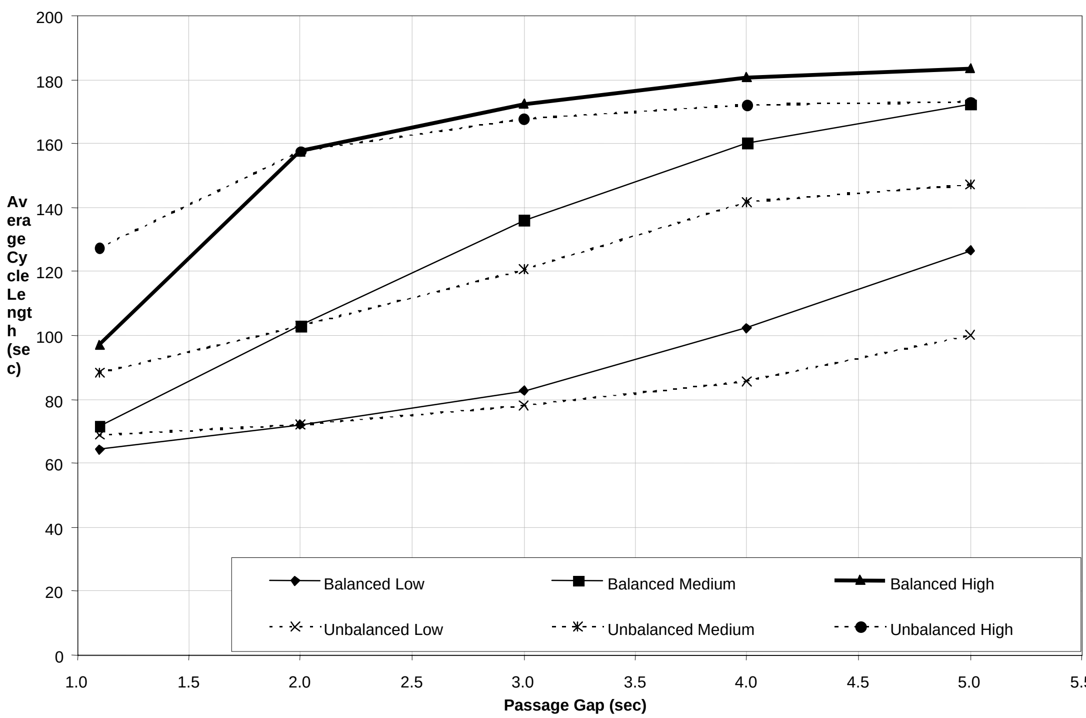

Figure 21. Software-Based, Advanced Testbed Cycle Lengths.

Hardware-in-the-Loop Simulation – Advanced Testbed

As expected, the hardware-in-the-loop simulation produced results that followed the same trends as the software simulation. The actuated cycle lengths produced increased consistently with passage gap, and the range of variation was similar to that found for software-only simulation. The cycle lengths ranged from a low of 60 seconds to a high of 190 seconds for gap settings of 1.1 to 5.0 seconds in the high volume

(V=1300 vph per approach) scenarios. A series of lower volume (V=800 vph per approach) scenarios produced more reasonable cycle lengths, ranging from 65 to 130 seconds. As seen with the software simulation as well, the increase in cycle length is more pronounced for higher volume scenarios. The results are shown in Figure 22.

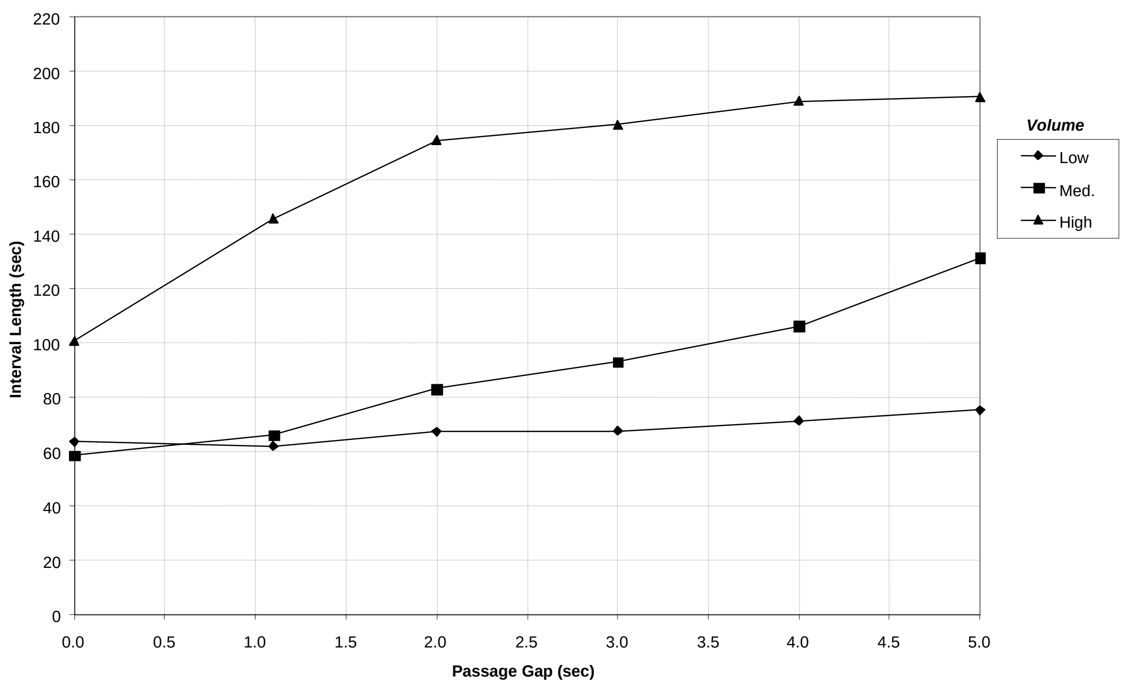

Figure 22. Hardware-Based, Advanced Testbed Cycle Lengths.

GREEN SPLIT RESULTS

The effects of variable passage gap on the green splits, or the proportion of the total available green time that is assigned to each phase, are described in this section for each of the scenarios tested.

Software-Based Simulation – Basic Testbed

When the passage gap is varied under balanced conditions, and the approach volumes are identical on both approaches, it would not be expected to find a change in the green splits between phases. This result was shown to be the case, with the two phases each receiving approximately 50 percent of the available green time over all passage gap settings. It was noted, however, that the variability and oscillation in green splits about the mean increased as the approach volumes increased, although the results still remained approximately evenly split.

Under unbalanced conditions, the variation of passage gap on one approach while holding the other constant should clearly illustrate the effect of variable passage gap on phase length and green allocation. This effect also proved to be the case with the proportion of green time allocated to the phase increased sharply as the passage gap setting for that phase was increased from 1.1 to 5.0 seconds. The magnitude of this increase was even more pronounced at higher volume levels. While both phases received nearly equal shares of green time at a gap setting of 1.1 seconds for all volumes, the ratio between phases reached as high as 80 percent to 20 percent for high volumes exposed to a gap setting of 5.0 seconds on only one approach.

Software-Based Simulation – Advanced Testbed

The advanced testbed, due to the presence of the exclusive left-turn phases, does not show the same equal splits of green time as the basic testbed. Due to the lower demand volumes for the left-turn phases, these movements receive a lower proportion of the total green time, as expected. Each left-turn phase accounts for roughly five percent of the available green time under low volumes and short gap settings. Note that, although volumes were varied, there was no change in turning vehicles as a percentage of the total approach volume. Across all passage gap settings, the percentage of green time allotted to left-turns tends to remain stable as volume increases.

The unbalanced simulations clearly show the effects of variable passage gap through direct contrast. For the approaches that were not varied, the gap setting was held constant at 2.0 seconds for all runs, and this is evident as a crossover point in the results. For gap settings lower than 2.0 seconds, the variable approaches clearly received a lower percentage of the available green time than the non-variable approaches. This decrease reversed as the variable passage gap was raised above 2.0 seconds. Variation in the splits among through phases was actually less dramatic at higher volumes, presumably because phase lengths at these high volumes are constrained by the maximum green time.

Hardware-in-the-Loop Simulation – Advanced Testbed

The hardware-in-the-loop simulation of the advanced testbed exhibited the same performance characteristics as the software-based simulation. The green splits produced remained stable and constant across all passage gap settings and volume levels tested, with only slight variation about a mean value. This consistency was expected, as the volume of traffic was increased by an equal percentage regardless of movement, and so the relative percentages of traffic demand on each movement remained constant as the overall volumes increased.

Due to the constraints of simulating in real-time, no unbalanced scenarios were tested using hardware-in-the-loop.

INTERVAL LENGTH RESULTS

As part of the analysis process, the resulting interval lengths produced by simulation were also reviewed. This analysis combines the review of cycle length and green splits into a single measure of the impact of variable passage gap on signal operations.

Software-Based Simulation – Basic Testbed

Under balanced simulations, the consistency of the green splits over variable passage gaps translates directly to nearly equal phase lengths. Given that cycle lengths increased with increasing passage gap, the lengths of the phases also increased accordingly, but competing phases remained evenly balanced.

However, the combined effects of increasing cycle length with passage gap and changes in green splits further exaggerate the difference in phase lengths associated with the unbalanced variation of passage gap. The increase in interval length for the phase being varied is quite dramatic compared to the competing phase with passage gap held steady.

Software-Based Simulation – Advanced Testbed

Under both balanced and unbalanced simulations, the changes in phase lengths mirrored the changes, or lack of changes, in green splits. Slight overall increases with increasing passage gap, as expected given the slightly increasing cycle lengths, were noted for all phases in all scenarios.

Hardware-in-the-Loop Simulation – Advanced Testbed

As with the software-only simulations, the interval lengths produced tracked the changes seen in cycle length as the passage gap setting was increased. Since all phases were treated equally in terms of passage gap, no clear distinctions between phases were noted, although the random variations were somewhat more apparent at higher volumes than lower volumes.

DELAY RESULTS

As described previously, average delay per vehicle was used as a measure of performance to ensure that the variation of passage gap did not have an unexpected detrimental effect on intersection operations. As discussed in the preceding chapter, the ultimate goal of most signal optimization strategies is to minimize delay. Therefore, it would ideally be possible to select the passage gap that minimizes delay by simply plotting delay versus passage gap for any given set of traffic and geometric conditions.

This selection was not possible in this experiment, however, due to limitations imposed by the detector configuration. Both intersection testbeds utilized very long stopline detectors, and extension detectors were omitted in order to improve the measurement of saturated green, and, therefore, the estimation of demand volumes. However, although testing of a range of passage gaps established the relationship between passage gap and phase and cycle lengths, these extensions were added to the presence time of the stopline detector. Therefore, any extension value automatically increases delay, as the vehicle has cleared the stopline just as the extension is beginning. In fact, the most optimal extension in terms of delay minimization would be a negative value. Therefore, all delay values obtained climbed steadily from those obtained with a passage gap of zero, illustrating that the detector configuration used, although appropriate for the purposes of this research, is inherently most efficient without the use of any extension.

For both the software-only and hardware-in-the-loop simulation runs, delay values for all runs conducted were plotted, and regression lines were fitted through all points conducted for runs with the same volumes. Software-only delay results are shown in Figure 23, and hardware-in-the-loop results are presented in Figure 24. It would not be valid to compare delay across runs with different volumes, even if the passage gap were the same, as one would expect that the average delay incurred per vehicle would increase as volume increased.

Both simulation methods exhibited the same trends, with delay for a given volume level rising slowly with passage gap. In addition, the difference in average delay between volume levels was quite pronounced.

Methods for addressing the delay minimization issue encountered in this research are discussed in the following section.

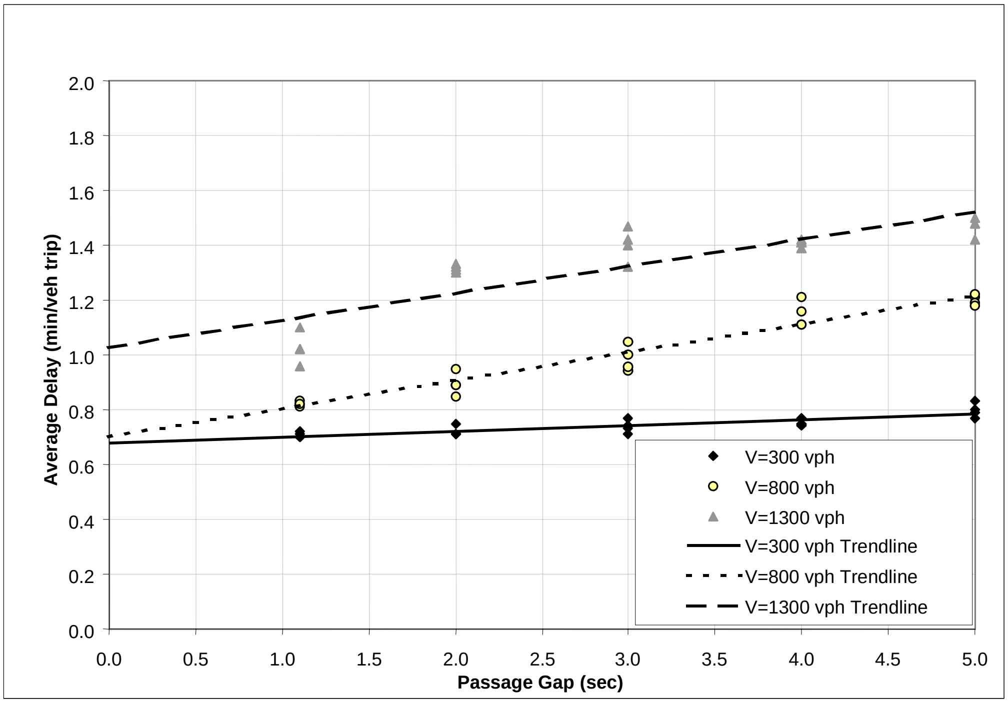

Figure 23. Software-Only Average Delay vs. Passage Gap.

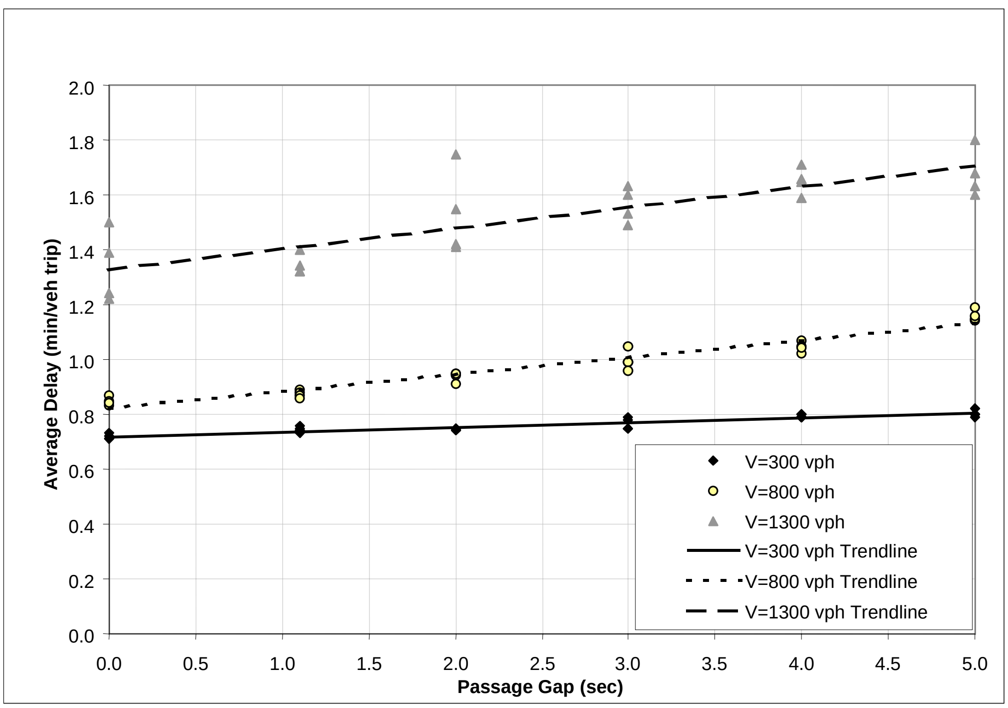

Figure 24. Hardware-in-the-Loop Average Delay vs. Passage Gap.

SUMMARY OF SIMULATION RESULTS

The general trends found as a result of simulation of the various scenarios are summarized in Table 1.

Table 1. Summary of Simulation Results.

Effects of Increasing Passage Gap					
		Cycle Length	Green Splits	Interval Length	Average Delay
	Software-Only Simulation				
Basic Testbed	Balanced	N/A**	None	Increase	N/A**
Basic Testbed	Unbalanced*	N/A**	Increase	Increase	N/A**
Advanced Testbed	Balanced	Increase	None	Increase	Slight Increase
Advanced Testbed	Unbalanced*	N/A**	Increase	Increase	N/A**
	Hardware-in-the-Loop Simulation				
Advanced Testbed	Balanced	Increase	None	Increase	Slight Increase

* NOTE: *For unbalanced scenarios, effects on approach(es) with variable passage gap are shown.*

** NOTE: *Since the fundamental relationships have been established, inclusion of these charts was not deemed necessary, and so they are omitted from this summary table.*

ERLANG MODEL RESULTS

The results of the iterative process to determine the expected cycle length are described in this section. As discussed in the previous section, an Erlang (a=2) headway distribution was assumed and, based on the input volumes, an expected extension was found. This, in turn, was used to find the expected cycle length.

Figure 25 shows the relationship between the expected cycle length under the Erlang model and the volume-capacity ratio for a range of passage gap settings. Note that this gap setting is the value that would be set in a controller and does not include the presence time of the vehicle on the detector. The detector presence time was included in the critical headway in the calculations, however.

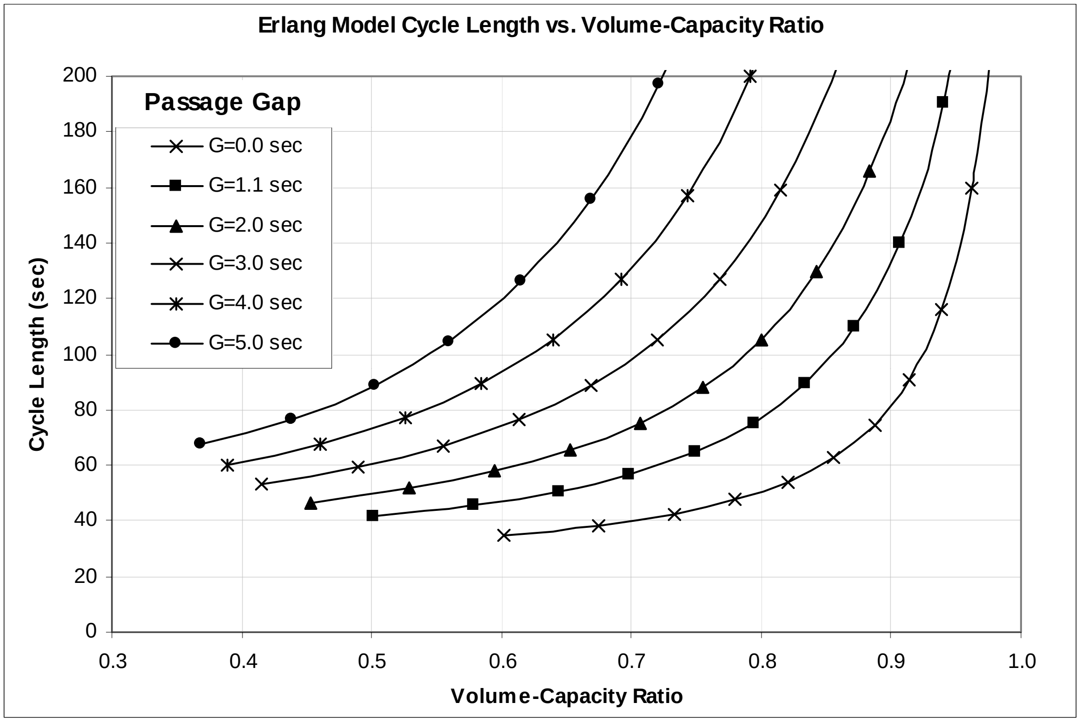

Figure 25. Erlang Model Cycle Length vs. Volume-Capacity Ratio.

Figure 26 shows the relationship between the expected cycle length and the critical intersection flow ratio. As described in the previous section, the flow ratio is based on a saturation flow rate of 1,800 vphgpl for through and right-turn lanes and 95 percent of 1,800 vphgpl (1,710 vphgpl) for the left-turn lanes.

The relationships shown are as anticipated, but two crucial factors are still missing if one is to use these charts to select the passage gap that will produce the desirable cycle. The boundaries of the range of desirable cycle lengths must be known, and the theoretical curves shown in these figures must be calibrated to match the true performance of the actual signal controller. These issues are addressed in the following sections.

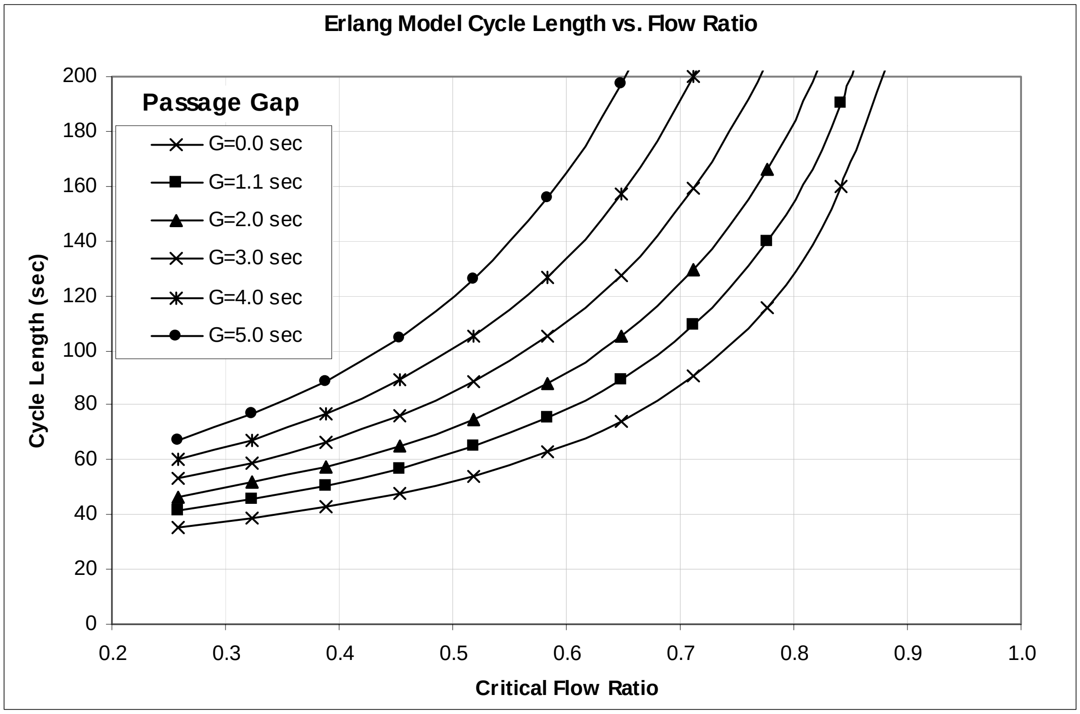

Figure 26. Erlang Model Cycle Length vs. Critical Intersection Flow Ratio.

COMBINING THEORY AND REALITY

The results of the theoretical model for predicting actuated cycle lengths were presented in the previous section. This section will build upon those two graphs, cycle length versus volume-capacity ratio and cycle length versus flow ratio, by incorporating real controller results and the boundary cycle lengths discussed previously.

Figure 27 and Figure 28 show the Erlang model curves, as well as points representing the performance found using hardware-in-the-loop simulation. Each point represents an average of at least four independent simulation runs, to account for random variations. The symbol used for each point corresponds to the passage gap setting used in the controller for those runs and matches the symbols shown along the Erlang curve representing the same passage gap value.

Also shown in Figure 27 and Figure 28 are the boundary conditions, as described previously. Recall that the saturated cycle length (v/c = 1.0) serves as a lower bound and that Webster's minimum delay cycle length for fixed time operation serves as the upper bound of the range of desirable cycle lengths. The goal is to operate at a cycle length within this range and presumably as close as possible to the minimum delay cycle length boundary.

Note that few passage gap settings fall within the desirable cycle length range. Again, this is due to the detector configuration in the intersection testbed. If extension detectors were included, one would expect to see all of the measured and theoretical curves presented shift downward, placing more points within the desirable range.

For example, at any given passage gap setting, a detector at the stopline, as was the case in these experiments, produces a given cycle length. If that same detector were placed upstream of the stopline, and the same gap setting were used, one would logically expect that the cycle length would be shorter, as the last vehicle to “make a phase” would trigger the extension detector earlier in the cycle, given that the detector would be further upstream. Consequently, less time would be wasted and the resulting cycle length would be shorter.

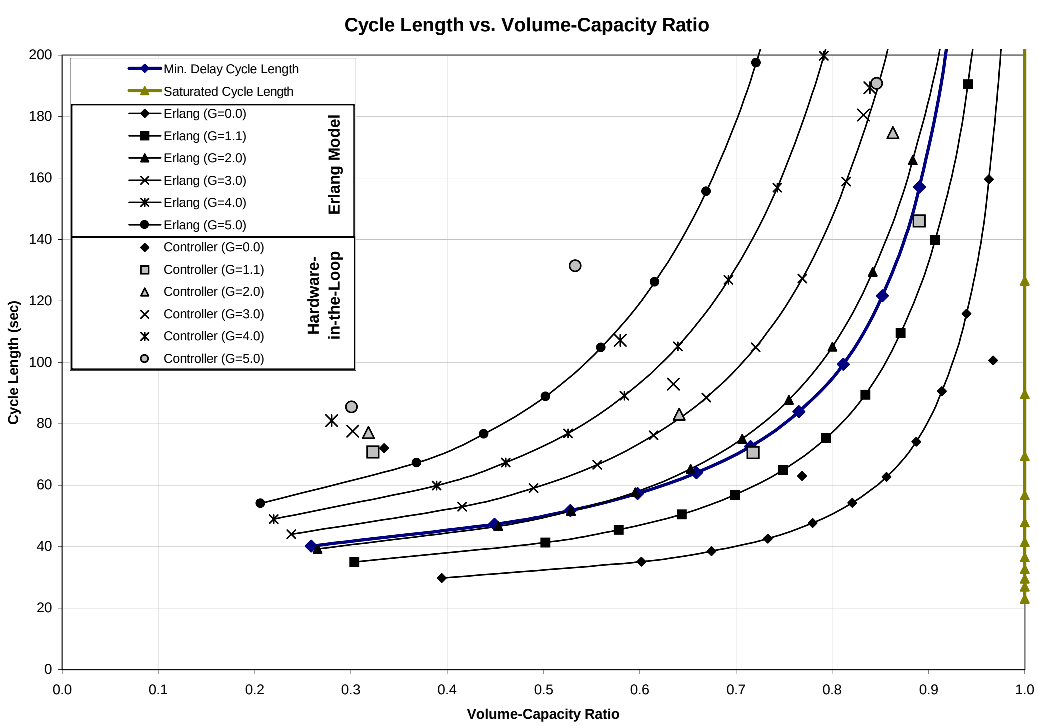

Figure 27. Combined Cycle Length vs. Volume-Capacity Ratio.

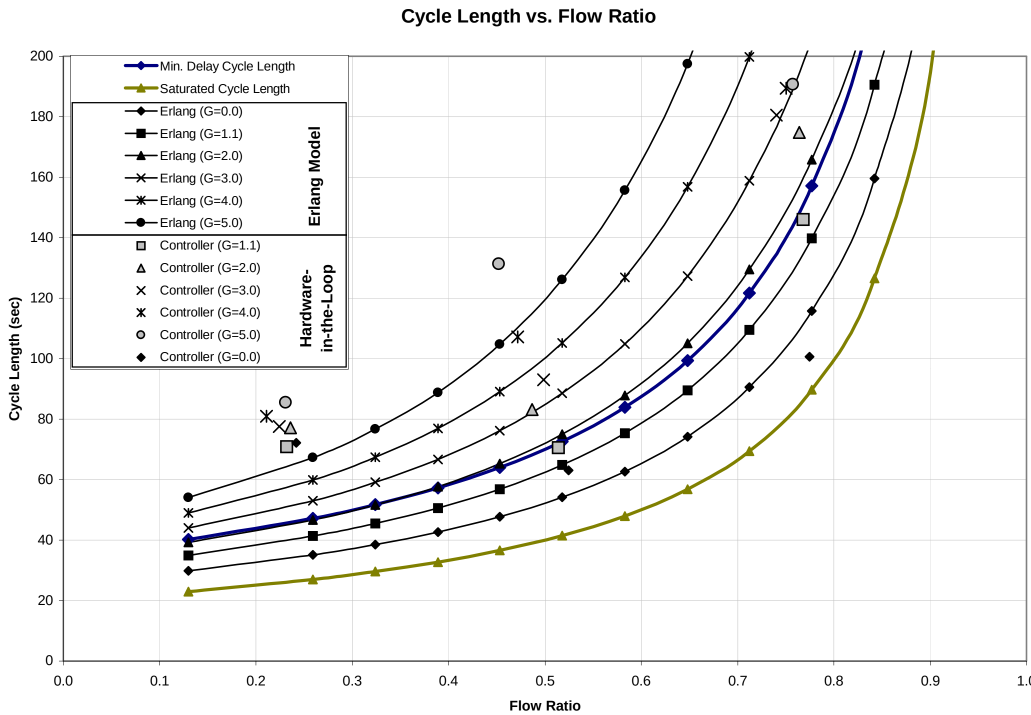

Figure 28. Combined Cycle Length vs. Flow Ratio.

CALCULATION OF PASSAGE GAP

A clear relationship between passage gap and cycle length has been established, and it has been shown that the theoretical model for predicting expected values performs similarly to the actual controller. Therefore, the next step is to formalize a method for determining the specific passage gap to be implemented under a given set of conditions.

Using the volume estimation methodology validated previously, a flow ratio and/or volume-capacity ratio can be determined for the current traffic conditions in real-time. A desirable target cycle length can be selected within the boundaries of the previously established range. Knowing the target cycle length and the current flow ratio loading of the intersection, the passage gap could be readily calculated.

However, there are several ways to derive the exact formula by which to calculate passage gap. Linear regression could be performed on the data resulting from the Erlang model iterations, or from the hardware-in-the-loop or software-only simulations. Regression, although possibly representative of the data generated under the test conditions at the time, lacks the theoretical foundation to explain why the parameters used in the function behave as they do. Therefore, an alternative method of representing the data was desired.

Many equations in traffic science take the form of Webster's equation, denoted earlier in this report as follows:

$$C_O = \frac{1.5L + 5}{1 - Y} \tag{16}$$

where

C_o = minimum delay (optimal) cycle length, sec;

L = total intersection lost time, sec; and

Y = sum of critical flow ratios (*8*)

In this equation, the parameters in the numerator were found through regression, but the structure of the equation represents the relationship that these parameters have to flow ratio and cycle length. It is this sort of transformation that was applied to the data generated in this research.

The above equation was rearranged to isolate the known parameters, cycle length and flow ratio, and the remaining unknown expression was determined by linear regression for each of the three techniques. Figure 29 presents the results graphically for the Erlang model, the software-only simulation, and the hardware-in-the-loop simulation.

Note that all three lines are quite close to one another, and that they behave in a similar manner. Considering the radically different techniques by which the data were arrived at that generated each set of data, the slight discrepancies are insignificant.

Use of any of the three equations to calculate passage gap would have a valid foundation, but the hardware-in-the-loop data was of greatest interest in this research as it most closely represents the operation of real-world intersections using actual field equipment.

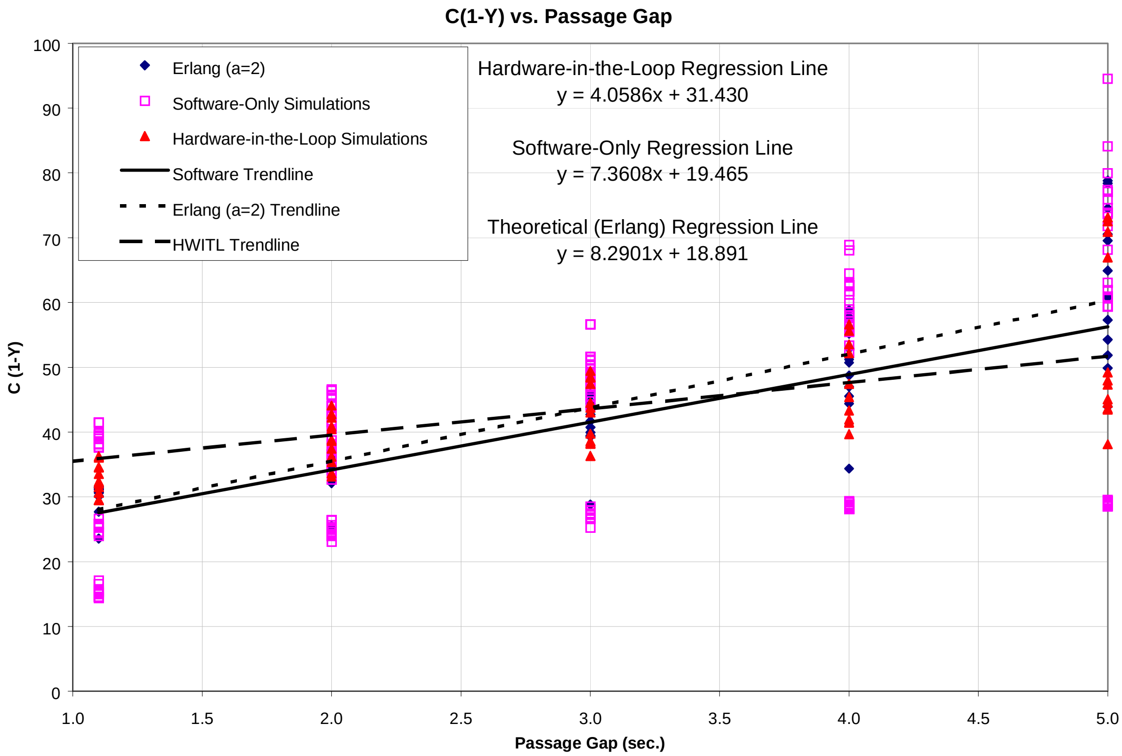

Figure 29. C(1-Y) Transform vs. Passage Gap.

Using the values obtained from the hardware-in-the-loop method, and given the geometric conditions, traffic pattern, and controller type used in these experiments, the cycle length for any given passage gap can be written as follows:

$$C = \frac{4.0586T + 31.43}{1 - Y} \qquad \textbf{(17)}$$

where

C = resulting actuated cycle length, sec;

T = passage gap, sec; and

Y = sum of critical flow ratios.

Consequently, one can solve directly for passage gap, knowing the desired cycle length and current flow ratio, as follows:

$$T = \frac{C(1-Y) - 31.43}{4.0586} \quad \textbf{(18)}$$

For comparison to the results found using Webster's formula, the total intersection lost time can be readily calculated if the passage gap were set to zero. The value for lost time would be 17.6 seconds, which is certainly reasonable and quite close to the value of 20 seconds assumed throughout this process. Using the assumed lost time value of 20 seconds, the passage gap setting that would produce equal results from the equation derived above and Webster's equation would be approximately 0.9 seconds. Again, the reasonableness of this gap value provides a "sanity check" on the derived equation presented above.

Similarly, the passage gap relationship can be quantified for software-only simulation as well, with the cycle length written as follows:

$$C = \frac{8.2901T + 18.891}{1-Y} \quad \textbf{(19)}$$

In addition, the software-only simulations showed that passage gap can be found as follows:

$$T = \frac{C(1-Y) - 18.891}{8.2901} \quad \textbf{(20)}$$

The results theoretical analysis showed that cycle length can be found as follows:

$$C = \frac{7.3608T + 19.465}{1-Y} \quad \textbf{(21)}$$

Rearranging the above equation, the passage gap can be solved for as follows:

$$T = \frac{C(1-Y)-19.465}{7.3608} \quad \textbf{(22)}$$

Note that although the mathematical formulations derived from all three analysis procedures are presented above for completeness, the hardware-in-the-loop results are recommended for use for the reasons discussed previously.

RESULTING CONTROL PROCESS

Having established the validity of the volume estimation procedure, and formalized the relationship between passage gap and cycle length, a spreadsheet was developed to recommend optimal passage gap settings in real-time. This format would allow direct time-based measures, including operating cycle length and saturated green, to be translated into estimated flow ratios and compared against the desired cycle length range. In turn, the most appropriate passage gap setting could be directly calculated using the formula provided above and implemented into the controller.

The spreadsheet, although functional, is limited in its ability to optimize signal performance by the detector configuration, which requires negative gap settings in order to truly minimize delay. However, the real-time procedure was tested in several trial runs to see that the process was workable. This proved to be the case to the extent that it could be tested with the simulation configuration used in these experiments. Recommendations on how to best implement this work are described in the following chapter.

5. CONCLUSIONS

This chapter describes the conclusions reached as a result of this research. Since the specific results found with regard to each component of the traffic control process proposed have been described in the previous chapter, only a final summary is presented here.

Since the evaluation tool used in this research was computer simulation, a number of aspects of the simulation environment were evaluated in order to substantiate the validity of the tool itself. The reliability with which the TRAF-NETSIM package produced the demand volumes that were requested of it in the input file was studied. It was concluded that the desired and measured volumes were quite close, and they were well within the limits of the random variations expected in the traffic flow. It was also concluded that the software-only simulation behaved in a similar manner to the hardware-in-the-loop simulation. Again, this conclusion substantiates the process that was followed during the course of this research effort and lends credence to the results produced. Furthermore, it was shown that, for low to moderate cycle lengths, the theoretical analysis procedure also matched the results produced by both forms of simulation and that the limits imposed by maximum green times restricted simulation cycle lengths when, in theory, such cycle lengths would have continued to increase sharply.

This research sought to develop a traffic control process capable of utilizing real-time estimated volume feedback and passage gap adjustment. The results presented previously show that the volume estimation methodology has been validated, and that, with calibration for site-specific conditions, it can be applied on a phase-by-phase basis. Assumptions for startup lost time and saturation flow rate were found to be useful in calibrating the equation when the true traffic volumes are known. The reliability with which the estimated volumes matched the measured volumes shows that this method can be relied upon to independently estimate the demand volumes for any given phase, in real-time, from stopline detector measurements. Overall, it can be concluded that, with calibration for local traffic and geometric conditions, the volume estimation methodology presented in this report can reliably estimate demand volumes, by phase, from stopline detectors with reasonable accuracy.

The use of passage gap to adjust signal operating parameters was also quantified. The relationship between passage gap and cycle length was established in mathematical form for the primary study intersection, and the derivation and calibration processes were described to allow application to other intersections as well. The relationship between passage gap and green splits, and, by extension, between passage gap and interval length, was also established qualitatively for some test scenarios. Overall, it can be concluded that increasing the passage gap for a particular phase increases the length of that phase. If the passage gap is increased equally on all phases, then the cycle length increases, and if it is increased on one phase only, then the relative green split to that phase increases.

REFERENCES

1. Messer, C.J. and Anwar, M. *Development and Demonstration of Sustainable Transportation Control Concepts: Actuated Control Optimizer Design*. Research Report SWUTC/98/472840-00013-1. Texas Transportation Institute, College Station, Texas, 1998.

2. *Manual on Uniform Traffic Control Devices, 1988 Edition*. U.S. Department of Transportation, Washington, DC, 1988.

3. McShane, W.R. and Roess, R.P. *Traffic Engineering*. Prentice Hall, Englewood Cliffs, New Jersey, 1990.

4. Kell, J.H. and Fullerton, I.J. *Manual of Traffic Signal Design, Second Edition*. Institute of Transportation Engineers, Washington, DC, 1998.

5. Owen, L.E., Stallard, C.M., and Glitz, D.M. *An Evaluation of Real-Time Traffic Adaptive Control Prototypes*. Prepared for Presentation at the 76th Annual Meeting of the Transportation Research Board, Washington, DC, 1997.

6. Bell, P.R. and Duncan, G.F. The New TS2 Standard: An Update. *ITE Journal* 59, August 1989.

7. DiStefano, J.J., Stubberud, A.R., and Williams, I.J. *Feedback and Control Systems, Second Edition*. Schaum's Outline Series. McGraw-Hill, Inc., New York, 1990.

8. Webster, and F.V., Cobbe, B.M. *Traffic Signals*. Technical Paper 56. Road Research Laboratory, London, 1966.

9. Tanner, M. *Practical Queueing Analysis*. McGraw-Hill, London, 1995.

10. Drew, D.R. *Traffic Flow Theory and Control*. McGraw-Hill, New York, 1968.

11. Li, J., Rouphail and N.M., Akcelik, R. Overflow Delay Estimation for a Simple Intersection with Fully Actuated Signal Control. In *Transportation Research Record 1457*, Transportation Research Board, National Research Council, Washington, DC, 1994.

12. ITT Systems and Sciences Corporation. *Traffic Software Integrated System Version 4.2 CORSIM Run-Time Extension*. FHWA, U.S. Department of Transportation, 1998.

13. Rathi, A.K. and Santiago, A.J. Urban Traffic Simulation: TRAF-NETSIM Program. *Journal of Transportation Engineering*, American Society of Civil Engineers, December 1990.

14. Engelbrecht, R.J., Poe, and C.M., Balke, K.N. *Development of a Distributed Hardware-In-The-Loop Simulation System for Transportation Networks*. Prepared for Presentation at the 78th Annual Meeting of the Transportation Research Board, Washington, DC, 1999.

15. Photo courtesy of TransLink® Roadside Equipment Laboratory, Texas Transportation Institute, College Station, Texas, 1998.

16. May, A.D. *Traffic Flow Fundamentals*. Prentice Hall, Englewood Cliffs, New Jersey, 1990.

17. Bullock, D. and Catarella, A. *A Real-Time Simulation Environment for Evaluating Traffic Signal Systems*. Prepared for Presentation at the 77th Annual Meeting of the Transportation Research Board, Washington, DC, 1998.