Final report of ITS Center project: Signal timing algorithm

For the Center for ITS Implementation Research

A U.S. DOT University Transportation Center

"Stochastic Traffic Signal Timing Optimization"

Principal Investigators:

Anil Kamarajugadda
Department of Civil Engineering
Center for Transportation Studies
University of Virginia

Dr. Byungkyu "Brian" Park
Department of Civil Engineering
Center for Transportation Studies
University of Virginia

August 2003

Research Report No. UVACTS-15-0-44
August 2003

Stochastic Traffic Signal Timing Optimization

By:

Anil Kamarajugadda
Dr. Byungkyu “Brian” Park

A Research Project Report
For the Center for ITS Implementation Research (ITS)
A U.S. DOT University Transportation Center

Anil Kamarajugadda
Department of Civil Engineering
Email: adk3w@virginia.edu

Dr. Byungkyu "Brian" Park
Department of Civil Engineering
Email: bpark@virginia.edu

Center for Transportation Studies at the University of Virginia produces outstanding transportation professionals, innovative research results and provides important public service. The Center for Transportation Studies is committed to academic excellence, multi-disciplinary research and to developing state-of-the-art facilities. Through a partnership with the Virginia Department of Transportation's (VDOT) Research Council (VTRC), CTS faculty hold joint appointments, VTRC research scientists teach specialized courses, and graduate student work is supported through a Graduate Research Assistantship Program. CTS receives substantial financial support from two federal University Transportation Center Grants: the Mid-Atlantic Universities Transportation Center (MAUTC), and through the National ITS Implementation Research Center (ITS Center). Other related research activities of the faculty include funding through FHWA, NSF, US Department of Transportation, VDOT, other governmental agencies and private companies.

CTS Website
http://cts.virginia.edu

Center for Transportation Studies
University of Virginia
351 McCormick Road, P.O. Box 400742
Charlottesville, VA 22904-4742
434.924.6362

1. Report No. UVACTS-15-0-44	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Stochastic Traffic Signal Timing Optimization		5. Report Date August 2003
		6. Performing Organization Code
7. Author(s) Anil Kamarajugadda and Byungkyu "Brian" Park		8. Performing Organization Report No.
9. Performing Organization and Address Center for Transportation Studies University of Virginia PO Box 400742 Charlottesville, VA 22904-7472		10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agencies' Name and Address Office of University Programs, Research and Special Programs Administration US Department of Transportation 400 Seventh Street, SW Washington DC 20590-0001		13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

Signalized intersections are a critical element of an urban road transportation system and maintaining these control systems at their optimal performance for different demand conditions has been the primary concern of the traffic engineers. Currently, average control delay is used as a performance measure of a signalized intersection. The control delay is estimated using the delay equation provided by the Highway Capacity Manual (HCM). The HCM delay equation is a function of multiple input parameters arising from geometry, traffic and signal conditions. Variables like volume, green time and saturation flow rate that influence delay computations are stochastic variables, which follow their characteristic distribution. This implies that delay has to be estimated as a distribution as against the point estimate, the average delay.

Various simulation programs and optimization techniques have evolved that aid the traffic engineer in the optimization process. None of the optimization programs consider the day-to-day stochastic variability in the delay during the optimization process. The purpose of this research is to estimate variability in delay at signalized intersections and incorporate the variability in the optimization process.

An analytical methodology to compute the variance of delay for an isolated intersection and arterial intersections is developed. First, delay variance is computed for an isolated intersection using expectation function method for undersaturated conditions and integration method for oversaturated conditions. The variance computation for an isolated intersection is expanded to arterial intersections using the integration method and the analytically approximated platoon dispersion model.

The delay variance estimates are then utilized in the optimization of intersections. A genetic algorithm approach is used in the optimization process using either average delay or the 95th percentile delay as an objective function. The results of the optimization, especially for isolated intersections, have shown considerable improvement over SYNCHRO, a signal optimization program, when evaluated using microscopic simulation programs SIMTRAFFIC and CORSIM. However, the results of arterial optimization did not show any significant improvement over the SYNCHRO.

17 Key Words	18. Distribution Statement
Stochastic Optimization, Traffic Signal, Simulation, Variability, HCM, Genetic Algorithms, Level of Service.	No restrictions. This document is available to the public.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	130	

ABSTRACT

Signalized intersections are a critical element of an urban road transportation system and maintaining these control systems at their optimal performance for different demand conditions has been the primary concern of the traffic engineers. Currently, average control delay is used as a performance measure of a signalized intersection. The control delay is estimated using the delay equation provided by the Highway Capacity Manual (HCM). The HCM delay equation is a function of multiple input parameters arising from geometry, traffic and signal conditions. Variables like volume, green time and saturation flow rate that influence delay computations are stochastic variables, which follow their characteristic distribution. This implies that delay has to be estimated as a distribution as against the point estimate, the average delay.

Various simulation programs and optimization techniques have evolved that aid the traffic engineer in the optimization process. None of the optimization programs consider the day-to-day stochastic variability in the delay during the optimization process. The purpose of this research is to estimate variability in delay at signalized intersections and incorporate the variability in the optimization process.

An analytical methodology to compute the variance of delay for an isolated intersection and arterial intersections is developed. First, delay variance is computed for an isolated intersection using expectation function method for undersaturated conditions and integration method for oversaturated conditions. The variance computation for an isolated intersection is expanded to arterial intersections using the integration method and the analytically approximated platoon dispersion model.

The delay variance estimates are then utilized in the optimization of intersections. A genetic algorithm approach is used in the optimization process using either average delay or the 95th percentile delay as an objective function. The results of the optimization, especially for isolated intersections, have shown considerable improvement over SYNCHRO, a signal optimization program, when evaluated using microscopic simulation programs SIMTRAFFIC and CORSIM. However, the results of arterial optimization did not show any significant improvement over the SYNCHRO.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

1.1 Background

Transportation systems are an integral part of a modern day society designed to provide efficient and economical movement between the component parts of the system and offer maximum possible mobility to all elements of our society. A competitive, growing economy requires a transportation system that can move people, goods, and services quickly and effectively. Road transportation is a critical link between all the other modes of transportation and proper functioning of road transportation, both by itself and as a part of a larger interconnected system, ensures a better performance of the transportation system as a whole.

Signalized intersections, as a critical element of an urban road transportation system, regulate the flow of vehicles through urban areas. Traffic flows through signalized intersections are filtered by the signal system (stopping of vehicles during red time) causing vehicular delays. Vehicular delay at signalized intersections increases the total travel time through an urban road network, resulting in a reduction in the speed, reliability, and cost-effectiveness of the transportation system. Increase in delay results in the degradation of the environment through increases in air and sound pollution. Thus, delay can be perceived as an obstacle that has a detrimental effect on the economy. It has been the traffic engineers' endeavor to quantify delay and optimize the signal system to perform at a minimum delay.

1.1.1 Delay Estimation

Delay estimate measures reflect the driver discomfort, frustration, fuel consumption and lost travel time. Numerous equations have been developed for the estimation of delay. In the U.S., delay is estimated using the Highway Capacity Manual (HCM). The HCM equation is a function of multiple input parameters arising from geometry, traffic and signal conditions. The HCM procedure for a signalized intersection uses average demand flow rate and saturation flow rate in order to estimate volume to capacity ratio and the corresponding performance measure, delay. The Level of Service (LOS) is then determined from a predefined range of average control delay values.

In practice, for gathering inputs for the evaluation of delay in the field, most efforts are given to the estimation of capacity, while traffic volumes are collected just for a day or two, with the exception of locations with existing surveillance systems. The delays are then calculated based on the average demand observed from the data collection process. Conversely, the demand volumes are subject to stochastic variability and usually follow a Normal distribution for daily variability and Poisson distribution for cyclic variability. Further, variables like green time and saturation flow rate (in addition to the volume) that influence delay computations are stochastic variables which follow their characteristic distribution. In addition, the average control delay at signalized intersection in the real world might vary depending on traffic conditions including different arrival distributions, percentage of trucks, and driver characteristics. Considering that delay is governed by a number of stochastic variables, it is imperative that delay also be considered a stochastic variable. This implies that delay has to be defined through a distribution as against the point estimate, the average delay.

This implies that an average delay value obtained from the data collection may not reflect the actual performance of the intersection. In other words, an average delay value of say 35 seconds per vehicle (LOS D) obtained on the data collection day has no significance if the 95^{th} percentile confidence interval of delay varies from 20 to 50 seconds per vehicle (LOS B to LOS D). Further, LOS is not a good performance measure when the delay value lies borderline of two LOS categories. For example, an average delay of 34.9 seconds is considered as LOS C, while 35.0 is LOS D.

Computation of the variability of delay often requires information on the demand variability like variance. With the use of advanced vehicle detection and communication technology, traffic count data from signalized intersections are extensively archived in places such as the Smart Travel Laboratory (STL) at the University of Virginia. The data in the STL is provided from Northern Virginia Smart Travel Signal System. The Management Information System for Transportation (MIST) at Northern Virginia controls over 1,000 signalized intersections, and system detectors report vehicle counts, speed, and occupancy every 15 minutes. Thus, vehicle demand variations can be easily captured and analyzed. Then, the delay variance could be estimated through sampling processes like Monte Carlo Simulations and Latin Hypercube Design. However, sampling techniques provide inconsistent results from different runs and requires a fairly large sample size to get close to the analytical values. Thus an analytical methodology that overcomes the drawbacks of the sampling procedures is desirable.

1.1.2 Optimization

Various simulation programs and optimization techniques have evolved that aid the traffic engineer in the optimization process. Delay and its derivative are used as the objective function in most optimization software. For example, SYNCHRO optimizes based on the percentile delay while TRANSYT-7F optimizes based on factor that involves the average delay called the disutility index. However, delay is a stochastic variable and optimizing a signalized intersection for an average value fails the system performance for extreme demand conditions. Thus, a methodology that optimizes the intersection considering stochastic variability would be very useful.

1.2 Objectives

The objectives of this analytical research are to

a) Develop an analytical methodology that estimates the variability of HCM delay equation for both undersaturated and oversaturated conditions, and
b) Optimize the signalized intersections considering stochastic variability, and evaluate its performance using microscopic simulation programs like SIMTRAFFIC and CORSIM.

1.3 Scope and Structure

An analytical methodology to compute the variance of delay for isolated and arterial intersections is developed. First, delay variance is computed for an isolated intersection using expectation functions for undersaturated conditions and numerical integration for oversaturated conditions. The variance computation for an isolated intersection is

expanded to an arterial intersection using the platoon dispersion model. The HCM delay equation is utilized in the computations and is assumed to be valid. However, the same methodology could be used with any analytical delay computation equation. The stochastic variability in delay is studied for the variability in the demand volumes only and the effect of variable saturation flow rates and green times is not studied.

The delay variance estimates are then utilized in the optimization of intersections. A genetic algorithm approach is used in the optimization process. Average delay, 95th percentile delay, and the total delay are some of the delay derivatives used in the optimization process. In addition to optimizing an isolated intersection, arterial network optimization is also presented. In the optimization of arterials, only two intersections are considered in the present report. However, the methodology could be expanded to a full fledged arterial.

This report commences with a literature review. Chapter 2 reviews the delay computations, uncertainty quantification, the platoon dispersion model and the genetic algorithm technique. Chapter 3 outlines the methodology in detail with sufficient examples. Chapters 4 - 7 present the results and comparisons with simulation models. Finally, the report concludes with conclusions and recommendations.

CHAPTER 2 LITERATURE REVIEW

This chapter presents the relevant literature that has been reviewed as a part of this research. This chapter commences with a discussion on the delay equations that have been developed beginning with the early 20th century. The platoon dispersion model, which is the link between isolated and arterial intersections, is reviewed next. This is followed by an overview of the uncertainty analysis where the pros and cons of the sampling procedures are presented. Finally, the optimization procedure utilized by various simulation programs is discussed followed by the Genetic Algorithm procedure.

2.1 Delay Background

Signalized intersections were developed in England in the early 20th century. With the introduction of these controls to maneuver conflicting streams of vehicular and passenger traffic, researchers have concentrated on estimating delays due to these controls and in developing the optimum signal timings to minimize delay especially for pre-timed signals. Webster's equation is one of the fore most delay equations developed in 1958 assuming practical distributions like Poisson (random) arrivals with uniform discharge headways [1]. Webster introduced three terms to the delay equation as shown below.

$$d = \frac{c(1-\lambda)^2}{2(1-\lambda x)} + \frac{x^2}{2q(1-x)} - 0.65\left(\frac{c}{q^2}\right)^{1/2} x^{(2+5\lambda)} \qquad (1)$$

where,

d is the average delay per vehicle,

c is the cycle time,

λ is the ratio of the effective green to the cycle length,

q is the flow rate,

s is the saturation flow rate, and

x is the degree of saturation.

The first term in Equation (1) represents the delay when traffic is considered to arrive at a uniform rate. The second term is a correction to consider the random nature of the arrivals. The third term is the empirical correction term introduced to give a closer fit to the simulated delay values. Furthermore, Webster used differential calculus techniques on the developed delay estimate to compute the cycle length for the minimum average delay.

Akcelik further developed the delay equation by utilizing the coordinate transformation technique to obtain a time-dependent equation that is applicable to signalized intersections [2]. A generalized delay equation of the form shown in Equation (2) was developed that embraces the Australian and Canadian delay formulas as well.

$$d = \frac{0.5c(1-u)^2}{1-ux} + 900Tx^n\left[(x-1) + \sqrt{(x-1)^2 + m(x-x_0)/QT}\right] \qquad (2)$$

where,

d = average overall delay

c = signal cycle time in seconds

u = g/C (ratio of effective green to the cycle length)

x = degree of saturation

T = flow period in hrs

m, n = calibration parameters

x_0 = degree of saturation below which the second term of the delay formula is zero

In the U. S., the Highway Capacity Manual (HCM) delay equation is utilized in delay computations. The HCM 2000 propounds that delay be computed using the following equation [3]:

$$D = \frac{0.5 \times C \times (1 - g/C)^2}{1 - \left[\min(1, X)\frac{g}{C}\right]} + 900 \times T \times \left[(X - 1) + \sqrt{(X - 1)^2 + \frac{8kIX}{cT}}\right] \quad (3)$$

Where,

C = cycle length in seconds,

g = effective green time in seconds,

X = degree of saturation (v/c),

v = demand volume in vehicles/hour

T = duration of analysis period hours,

k = incremental delay factor, 0.5 for pre-timed signals,

i = upstream filtering/metering adjustment factor, 1 for isolated intersection, and

c = capacity in vehicles per hour.

Engelbrecht et al. have validated the HCM delay equation for oversaturated conditions and for different period of analysis [4]. Delays estimated by the HCM 2000

delay model were observed to be in close agreement with the delay estimates from TRAF-NETSIM simulations.

In addition to computing point estimates like the average delay, attempts have been made to quantify the variability in delay due to random nature of the arrivals. Olsweski developed numerical methods to calculate average delay and estimate the distribution of the average cyclic delay. The methodology is based on sequential calculation of queue length probabilities with different arrival processes and was not applicable to practical situations [5]. Fu et al. modeled analytical equations to compute the variance of delay based on a simulation study [6]. The model for variance had to be calibrated extensively and the calibration depended on the delay definition used in the simulation program. Further, a signalized intersection optimization based on variance minimization was conducted and the results were similar to that obtained from average delay minimization.

The arrival process observed downstream of a traffic signal is expected to differ from that observed upstream of the signal due to platooning of the vehicles. Thus, the delay equation developed for an isolated intersection has been modified to be applicable to arterial intersections. A progression factor was introduced for estimating delays for intersections in an arterial to account for this process. Rouphail developed a set of progression factors that adjusts delays at coordinated intersections using time-space diagrams and flow counts [7]. Levinson**,** also attempted to compute the signal delay for platooned arrivals for two extreme conditions [8]:

a) When the first vehicle in the platoon arrives during a green interval and is unimpeded and

b) When the first vehicle in the platoon arrives during the red period and is impeded by queued vehicles.

Olszewski computed delay for a pretimed signal when the arrival rate is non-uniform by utilizing the step arrival rate model [9]. A significant finding of his research was that the progression effects the uniform delay term and not the overflow delay term in the HCM delay equation. Teply presented a practical system to evaluate the signal coordination at a series of intersections and studied the quality of signal progression based on the time-space charts developed from surveys and simulation [10].

2.2 Platoon Dispersion

"The on-off nature of traffic signal tends to create bunches or "platoons" of vehicles. The platoons of vehicles disperse as they travel away from the lights due to the different speeds of the individual vehicles" [11]. Thus the arrival pattern at an intersection downstream from another signal is different from an isolated intersection. Robertson developed the platoon dispersion model for the Road Research Laboratory in United Kingdom in 1969. The dispersion model was developed based on the observations made at four sites in West London at approximately 300, 600 and 1000 ft downstream of the stop bar. The predicated flow rate at any time step is expressed as a linear combination of the original platoon flow rate in the corresponding time step (with a lag of t) and the flow rate of the predicted platoon in the step immediately preceding it. Equation (24) presents the recursive model developed by Robertson. A smoothing factor 'F' is used in the model to best fit the actual and calculated platoon shapes and is inversely proportional to the travel time on the link. The arrivals at the downstream intersection are estimated

depending on the discharge patterns from upstream intersection. The smoothing factor is found to be site specific and depends on the road width, gradient, parking, opposing flow level, etc [11].

Rouphail developed a closed form solution for the recursive model developed by Robertson, and studied the effect of platoon dispersion on signal coordination and delay estimation. Flow rates in the predicted platoon measured at the k^{th} interval of the j^{th} simulated cycle are expressed in terms of the demand and capacity rates at the source intersection in addition to signal-control and travel-time parameters [12].

Successful implementation of the dispersion model needs parameters like the dispersion factor α and travel-time factor β (Refer to Equation (24)) to be calibrated according to the conditions of the arterial. McCoy et al. attempted to calibrate the platoon dispersion model for passenger cars under low friction traffic flow conditions and suggested appropriate values for α and β to 0.21 and 0.97 for two-way-two-lane street and 0.15 and 0.91 on a four lane divided highway [13]. McTrans suggests that the degree of platoon dispersion on internal links can be calibrated for local conditions by using the platoon dispersion factor (PDF). High platoon dispersion factors indicate heavy friction, (i.e., urban central business districts (CBD) areas having significant amounts of parking, turning, pedestrians, and narrow lane widths), which conspire to reduce platoon intensities. Low platoon dispersion factors indicate low friction, (i.e., ideal suburban high- type arterial street conditions) that allows increased platoon intensities [14]. A value of 0.35 for α was found to be suitable for U.S. conditions.

2.3 Uncertainty Analysis

Uncertainty is inherent in any system or model and it has been the engineers' endeavor to reduce uncertainty to the minimum wherever possible and to quantify the uncertainty in the system. Sampling techniques have been widely used in transportation engineering to quantify variability [15]. Sampling techniques involve the running of the model for a selected set of inputs based on their probability distributions to generate the probability distribution of the output. These sampling procedures lack accuracy because every simulation run produces different results and a large sample size might be necessary for the convergence to the true solution. Some of the widely used sampling methods for uncertainty analysis are Monte Carlo Simulation and Latin Hypercube Sampling.

Monte Carlo simulation, a simple random sampling (SRS) procedure, is the most widely used sampling method for computer experiments because they are quick and easy to implement for high dimension problems. Many of the initial studies in computer experiments investigated the distribution of the response given "random" inputs. However, with increasing complexity of the problems, an improved design strategy called the Latin Hypercube Design is utilized..

A Latin Hypercube Design improves the distribution of input variables in the design of a sample and the design with the best distribution of points is selected. Ideally, the Latin Hypercube design generates a minimal number of input combinations that are spread as evenly as possible in the experimental space. A Latin Hypercube design with the design points more uniformly spaced can be chosen by measuring the variability of the number of design points in a randomly located sub region of the experimental design space. To ensure that each of the input variables X_k has all portions of its distribution

represented by the input values, the range of X_k is divided into N strata of equal marginal probability 1/N and one sample is picked from each stratum [16].

Other probabilistic methods like expectation functions have evolved. Expectation functions can be used to quantify uncertainty contributed by uncertain input parameters and expectation functions overcome the drawbacks of the sampling procedures. The moments of the output variable about the mean are estimated based on the distribution of the input parameters. Exact knowledge of these moments is used in identifying the distribution of the output random variable. Tyagi et al. discusses various distributions for these input parameters and developed generic expectation functions [17].

Very little literature was found on the uncertainty analysis of delay under variable demand conditions. Olszewski attempted to develop the probability distribution of delay while Fu et al. calibrated a model that computes the variance of delay assuming that the arrivals follow a Poisson distribution [6, 18]. The developed calibrated model was compared to the theoretically developed Markov Chain model and was found to comply well within a range of degree of saturation values.

2.4 Optimization

One of the primary objectives of this research is to optimize signalized intersections using stochastic variability in delay. Delay is a stochastic variable and its variability has to be accounted for in the optimization process. Simulation and optimization programs like SYNCHRO and TRANSYT-7F do not consider the variability in delay for their optimization.

2.4.1 TRANSYT-7F

TRANSYT is macroscopic optimization and simulation tool originally developed in the United Kingdom by the Transport and Road Research Laboratory (TRRL) [11]. It is a model that considers platoon dispersion for its computations. TRANSYT-7F is a U.S. version of TRANSYT developed by the University of Florida. TRANSYT-7F uses a delay derivative (Disutility Index) as the objective function during the optimization process [14]. The delay definition used is the entire amount of time spent while not traveling at the prevailing cruise speed. TRANSYT-7F measures this by periodically counting the number of vehicles queued at a signal and integrates this series of counts over time. Uniform and residual delays are computed based on the area under the uniform queue profile (queue.out from Spyglass). Incremental delay is computed by using the Highway Capacity Manual equation, where certain input parameters (e.g., capacity) are obtained directly from TRANSYT-7F simulation, and other input parameters (e.g. duration of the analysis) are obtained directly from the input data file. TRANSYT-7F uses genetic algorithms or the hill climbing method for optimizing cycle length, splits, phase sequences and offsets.

2.4.2 SYNCHRO

SYNCHRO, developed by Trafficware Inc., is a software package that can model and optimize traffic signal timings. SYNCHRO minimizes a parameter called percentile delay in its optimization. The Percentile Delay is the weighted average of a delay corresponding to the 10th, 30th, 50th, 70th and 90th percentile volumes. SYNCHRO accommodates for progression by calculating the progression factor (PF) used in the

delay equation using the ratio of uniform delay calculated by SYNCHRO with coordination and uniform delay calculated by SYNCHRO assuming random arrivals. Furthermore, SYNCHRO uses quasi-exhaustive search in offset optimization [19].

2.4.3 Genetic Algorithms

Genetic Algorithms are search algorithms based on the mechanics of natural selection and evolution. John Holland, his colleagues and his two students at the University of Michigan developed these algorithms [20].

A genetic algorithm process starts with a random set of individuals called the population. The individuals in a population are represented in the form of binary strings. These strings are then acted upon by operators, which produce a different population every generation, and then this cycle is repeated until certain termination criteria are met. A simple genetic algorithm is composed of three operators:

- Reproduction
- Crossover
- Mutation

The reproduction is a process in which individuals are selected based upon their fitness value or the objective function. This operator is an artificial version of natural selection, the survival of the fittest. The reproduction operator is implemented in algorithmic form in a number of ways. Roulette wheel selection and tournament selection are some of them. After reproduction, a crossover operator involving two steps is operated. Firstly, members of the newly reproduced strings in the mating pool are mated at random. Secondly, each pair of strings undergoes crossover as follows: an integer

position k along the string is selected uniformly at random between 1 and the string length minus one. Two new strings are created by swapping all characters, between positions k+1 and the string length. A mutation is an operator to move the function from local maxima and minima. A simple mutation involves generating a random number for every digit in the binary string and if the number is less than a predefined mutation probability (usually 0.05) the digit in the binary string is flipped.

These genetic algorithms have been gaining significance in its applications for transportation signal system optimization. Foy et al. have used GA to develop a demand responsive (adaptive) optimization technique to control traffic signals (Traffic GA) [21]. A street network with four intersections shaped in a square configuration and each intersection being connected to two other intersections by perpendicular roadways was simulated. A two phase signal control was assumed for all the four intersections. Nine decision variables were involved in the optimization process, the total green for all the phases and two variables (for all the four intersections) one for the phase sequence and the other for the proportion of the green time allocated the phases. These nine variables were coded into a 24 bit string. In Traffic GA, the inverse of the total average wait time was used as an objective function. A simulation model was developed by them to be used as an evaluator. Each GA run consisted of 50 individuals and the program was run for 60 times for every set of conditions. After 60 generations, the member with a minimum fitness value was chosen as the best solution for the given set of conditions. Traffic GA was run repeatedly while using the newest traffic data and new signal timing plans better suited to the present conditions are displayed. Balanced conditions of green phase times and a reasonable cycle lengths were obtained as a function of the traffic demand. The

traffic GA results and the theory of convergence indicate that GAs may be able to solve more difficult problems than traditional control strategies and search methods.

Park et al. [22] have developed a procedure that optimizes all the traffic control parameters (i.e., cycle length, green split, offset, and phase sequence) for oversaturated and undersaturated conditions. The procedure utilizes genetic algorithm-based program to optimize the four parameters simultaneously as well as model queue blocking effects. Delay multiplied by -1 is utilized as an evaluation function for the optimization process. The genetic algorithm-based signal optimization program was implemented at two closely-spaced signalized intersections within 100 meters of each other. The GA optimizer utilized up to 250 generations with a population size of 10 per generation, a crossover probability of 0.4, and a mutation probability of 0.03. An elitist method was used for the GA selection process. The results of genetic algorithm optimizer indicate that the GA optimizer searches more frequently for a good cycle length range. It was found that the proposed GA-based program provides acceptable solutions within reasonable amount of time [22].

The above study was extended to deal with three different optimization strategies and was tested under different intersection spacing: 100, 200, and 300 meters [23]. Three types of objective functions are considered namely, throughput maximization, average delay minimization, and modified average delay minimization with a penalty function. An arbitrary arterial system consisting of four intersections was selected in order to test the GA-based program. Of the three objective functions, the delay minimization strategy is observed to be applicable to both undersaturated and oversaturated conditions. The GA-based program and TRANSYT-7F timing plans were compared. The GA-based

program yielded less queue time than that of TRANSYT-7F on the basis of multiple CORSIM runs [23].

CHAPTER 3 METHODOLOGY

As mentioned earlier, the primary objectives of this research are to develop an analytical methodology that estimates the variability of HCM delay equation for both undersaturated and oversaturated conditions, and further, to optimize the signalized intersections considering stochastic variability in the demand volumes. The methodology for the realization of these objectives will be presented in this chapter. Furthermore, the results of the optimization are compared with that of SYNCHRO and the timing plans from both the optimization processes are evaluated using microscopic simulation programs SIMTRAFFIC and CORSIM. The results will be presented in the chapter 7.

The following is an overview of the methodology chapter. Firstly, the HCM delay equation for a signalized intersection is presented and its stochastic variables are identified. Then, the methodology developed for the delay variability estimation of an isolated intersection is outlined. Undersaturated and oversaturated intersections are dealt differently and the methodology for each condition is presented separately. The delay variability estimation for an arterial intersection is presented next. Finally, the genetic algorithm based optimization procedure is presented along with a brief overview of the C++ program developed for the optimization. The results of the methodology will be presented in the later chapters.

3.1 Highway Capacity Manual Delay Equation for Isolated intersection

The Highway Capacity Manual (HCM) delay equation has been presented in the literature review. The HCM delay equation (Equation (3)) has a number of input variables some of which are subject to stochastic variability. The following section presents in detail the stochastic input variables to the HCM delay equation.

3.1.1 Stochastic Variables Identified

Of the multitude of input variables presented in Equation (3), effective green time, saturation flow rate, and degree of saturation are identified to be stochastic variables. The reasons for identifying only three variables are presented here.

The effective green time is computed using the following equation

$$g = G + Y + R_c - L \quad (4)$$

Where,

g = effective green time,

G = displayed green time,

Y = yellow time,

R_c = red clearance time, and

L = lost time.

The lost time in Equation (4) changes from driver to driver and hence induces stochastic variability into the effective green time. However, the magnitude of the variability in lost times is negligible compared to the variability of the volumes and it is difficult to quantify the variability in the lost times that driver's experience, as it would involve collecting data on human perception. Likewise, saturation flow rate also has stochastic variability arising from the presence of heavy vehicles and the temporal variations in saturation flow estimation. Nevertheless, as the variability of effective green time and saturation flow rate are relatively small, for this exploratory research the variations in traffic demand are only considered.

The variability in the degree of saturation arises from the random nature of the traffic arrivals at an intersection. The degree of saturation is nothing but the ratio of the volume to the capacity. The assumptions of constant green time and saturation flow rate fixes the capacity which is a function of green time and saturation flow rate. This implies that the degree of saturation (X), which is governed by the volume (V) and capacity, will vary according to the volume only. The variability in the traffic demand (V) is inherent to any traffic system due to the random nature of the arrivals at an intersection and this induces stochastic variability into the degree of saturation. The variability of the arrivals could be identified from one cycle to the other and/or from one day to another. Since the degree of saturation (X) is equal to demand volume (V) divided by a constant capacity (c), the mean of X will be the average demand divided by a constant capacity and the standard deviation of X will be the standard deviation of demand divided by the capacity. Further, the distribution of X will be the same as that of V with the mean and standard deviation scaled down by a factor (capacity) as mentioned above. Hence, given the mean, variance and distribution of the demand volumes, the stochastic properties of X can be computed and these values of mean and variance are further used for delay variance computations.

3.2 Inputs

The proposed methodology computes the delay variance at a signalized intersection. The variables required for delay variance computations are mentioned here. Firstly, HCM delay for an intersection is computed by taking the weighted average of the delays from all the lane groups with their corresponding average volumes. Similarly, the delay

variance is also computed for all the lane groups separately and then aggregated. Therefore, the data required for the delay variance computations are required for each lane group separately.

As mentioned earlier, computation of the variability of delay requires the mean, variance and distribution of all the stochastic variables. As volume is the only variable that is being considered for stochastic variability, information on the average volume, variance and distribution is needed for every lane group. In addition to stochastic details, other information like the signal timing plan, saturation flow rates and other variables that are essential to HCM delay computation are necessary. The following is a detailed list of all the variables required for an isolated intersection computation

- Average volume, variance and distribution for every lane group
- Green times for every lane group
- Cycle length for the intersection
- Saturation flow rate by lane groups (left turn and through volumes may have different saturation flow rates)
- Duration of analysis period (T)
- Incremental delay factor, upstream filtering/metering adjustment factor.

For an arterial intersection delay variance computation, the following additional inputs are required:

- Cruise time on the arterial link
- Platoon dispersion factor (α) and the empirical factor (β)

These additional factors are used to calibrate the platoon dispersion model and are explained in section 3.6.2.

The mean and variance of volumes could be obtained from data collection. A signal control system like the Management Information System for Transportation (MIST) of Northern Virginia can provide volume information. The distributions of the volumes are usually assumed to be Poisson for cycle-by-cycle variability and Normal, Uniform or any other feasible distribution for day-to-day variability. However, the distributions mentioned are valid only for an isolated intersection. For an intersection in an arterial, the arrivals are screened through an upstream intersection and therefore, the distribution of arrivals at a downstream intersection is different from the standard distributions and has to be estimated. The platoon dispersion model is used in the estimation process and the parameters required to calibrate the model are presented in section 2.2.

3.3 Overview of the Methodology

The following flow chart (Figure 1) is an overview of the methodology involved in the optimization procedure. Two modules are marked as the delay variance computation module and the optimization module. The delay variance computation module involves inputting the data and computing the delay variance for different demand conditions. The optimization module involves the signal timing development process using the genetic algorithms. The different aspects involved in the flowchart will be presented in the later sections of the methodology chapter.

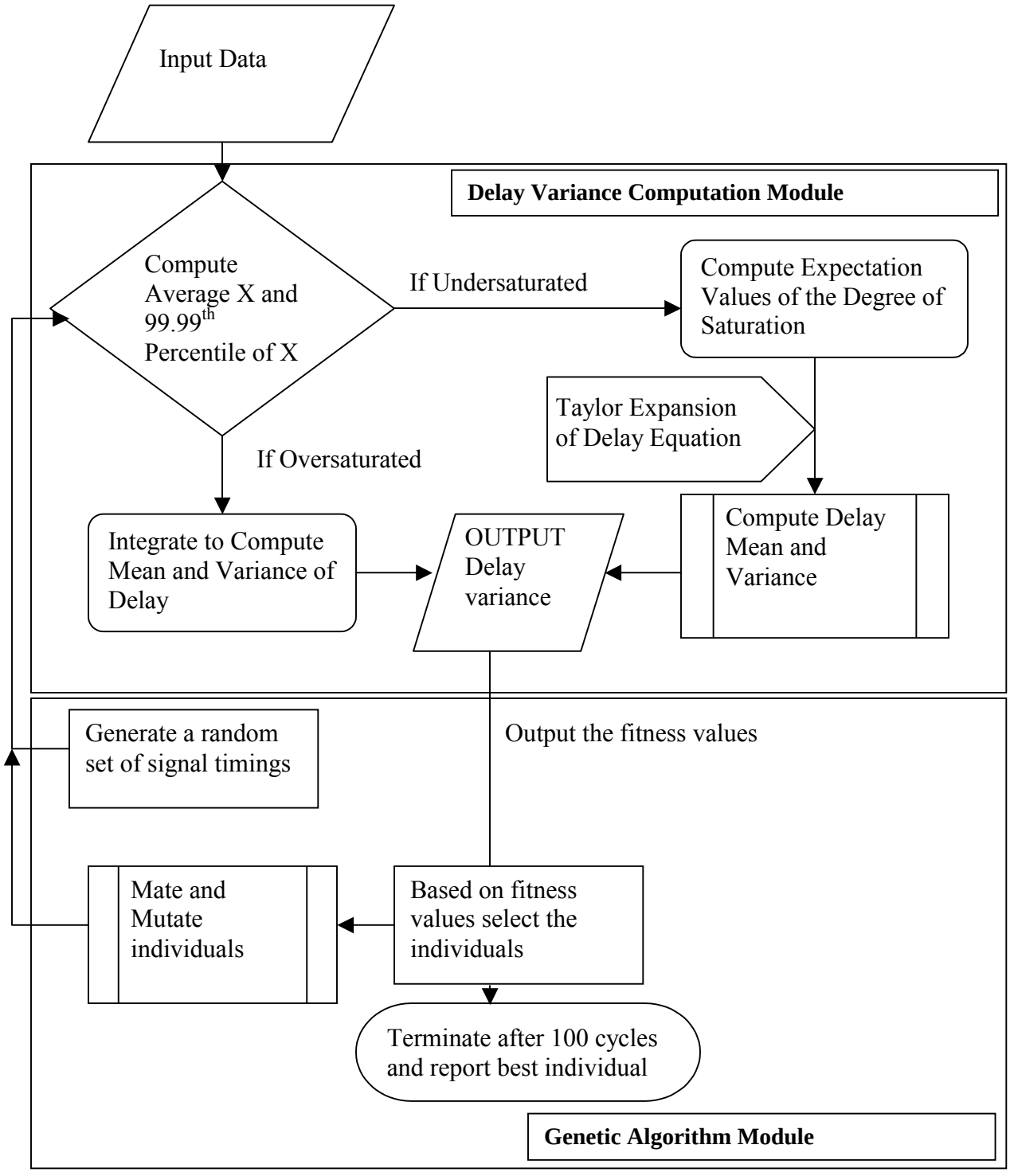

Figure 1. Flowchart of intersection signal timing optimization

3.4 Undersaturated isolated intersection

The following section presents the methodology involved in the delay variance computations for an undersaturated intersection. The methodology for an isolated undersaturated intersection involves the expectation function method. First, the assumptions are highlighted followed by the expectation function method. Then, the Taylor series expansion used to simplify the HCM delay equation is explained. Finally, the procedure for HCM delay variance computation is presented.

3.4.1 Assumptions

The following assumptions are made in developing the methodology.

- HCM delay equation is valid.
- Intersection is operating under undersaturated conditions.
- Saturation flow rate and effective green times are constant.

The first assumption is valid as has been observed in the literature review. The second assumption implies that after considering the variability in the volumes, the 99.99^{th} percentile upper confidence limit on the degree of saturation has to be less than one. Otherwise, the methodology for an oversaturated intersection has to be applied. The validity of the third assumption is discussed earlier in this chapter under the inputs section 3.2.

3.4.2 Expectation Function Method

The expectation method is an analytic procedure that overcomes the shortcomings of sampling procedures. This method involves expectation functions. In this methodology, each stochastic input variable is considered a random variable following a distribution with known mean and variance. Since the output (from the function) is dependent on the input variables, the output is also a random variable whose higher order moments are to be calculated based on the variation of the input variables. The expectation method can be used to calculate the first and higher order moments of an output variable that is a function of several independent random variables in multiplicative, additive and combined forms. The following section explains the expectation functions utilized in the methodology.

3.4.2.1 Expectation Functions

Expectation functions provide the expectation values of the powers of the variable given the mean, variance, and distribution of the variable. That is, given the mean (μ), variance (σ^2), and the distribution of a random variable X, expectation functions provide the expectation of X^n as a function of $\overline{X}$ and σ^2 (i.e., $E(X^n) = f(\overline{X}, \sigma^2)$). The nature of the function depends on the distribution of X.

Tyagi et al. developed generic expectation function equations, which are functions of the mean and coefficient of variance (COV) of an input random variable [17]. The input random variable can have a Uniform, Triangular, Lognormal, Gamma, Exponential, or Normal distribution. With a prior knowledge of the mean, variance and the distribution of the input random variable the expectation values of the higher powers of the variables

can be computed. For example, consider an input variable (X) that follows a distribution with a mean value of μ_X and a coefficient of variance of CV_X (i.e., the ratio of Standard deviation to mean). The expectation values of powers of X (depending on the distribution) are presented below for Normal, Uniform, Log-normal and Gamma distributions [17].

For Normal distribution,

$$E\left[X^r\right] = \mu_X^r \sum_{n=0}^{r/2} \binom{r}{2n} CV_X^{2n} E\left[z^{2n}\right] \qquad \text{When r is even} \tag{5a}$$

$$E\left[X^r\right] = \mu_X^r \sum_{n=0}^{(r-1)/2} \binom{r}{2n} CV_X^{2n} E\left[z^{2n}\right] \qquad \text{When r is odd} \tag{5b}$$

Where,

r is the power of the variable for which the expectation value is being computed, and

z is the unit normal variable.

For Uniform Distribution,

$$E(X^r) = \frac{\mu_X^r}{2\sqrt{3}(r+1)CV_X} \cdot \left[\left(1 + CV_X\sqrt{3}\right)^{r+1} - \left(1 - CV_X\sqrt{3}\right)^{r+1}\right] \tag{6}$$

For Lognormal distribution,

$$E\left(X^{r}\right)=\mu_{X}^{r}\left(1+CV_{X}^{2}\right)^{r(r-1)/2} \tag{7}$$

For Gamma distribution,

$$E(X^{r})=\mu_{X}^{r}CV_{X}^{2r}\exp\left\{\ln\left[\Gamma\left(CV_{X}^{-2}+r\right)\right]-\ln\left[\Gamma\left(CV_{X}^{-2}\right)\right]\right\} \tag{8}$$

Where,

Γ is a gamma function.

As pointed-out previously, the variability in the volumes could be identified from day-to-day and/or from cycle-to-cycle. The above distributions could be used for day-to-day variability condition in the demand volumes. However, the demand conditions are expected to follow a Poisson distribution for a cycle-to-cycle variability. Hence to compute the cyclic variability in delay, expectation values for Poisson distribution are required. These expectation values are generated from the basic definition of a Poisson distribution.

Poisson is a count distribution. A cycle-by-cycle distribution for volumes implies that the volume counts from different cycles follow a Poisson distribution. Therefore, the expectation values are generated for these cyclic volume counts and then translated to hourly volumes.

The generalized expression for Poisson distribution is as shown in equation below

$$P(X=x)=\frac{e^{-\lambda}\lambda^{x}}{x!}$$

Where,

λ is the mean value of the variable x.

Poisson distribution is a discrete distribution. Therefore, the generalized expression for computing the expectation values could be written as

$$E(X^n) = \sum x^n \cdot \frac{e^{-\lambda}\lambda^x}{x!}.$$

For n=1,

$$E(X) = \sum x \cdot \frac{e^{-\lambda}\lambda^x}{x!} = \sum \frac{e^{-\lambda}\lambda^x}{(x-1)!} = e^{-\lambda} \cdot \lambda \cdot \sum \frac{\lambda^{x-1}}{(x-1)!} = e^{-\lambda} \cdot \lambda \cdot e^{\lambda} = \lambda.$$

This equation implies that the mean of a Poisson distribution is λ! which is true.

For n=2,

$E(X^2) = \sum x^2 \cdot \frac{e^{-\lambda}\lambda^x}{x!}$. x^2 in the Right Hand Side (RHS) of this equation can be expressed as $x \cdot (x-1) + x$ and simplified as follows:

$$E(X^2) = \sum x \cdot (x-1) \cdot \frac{e^{-\lambda}\lambda^x}{x!} + \sum x \cdot \frac{e^{-\lambda}\lambda^x}{x!} = \sum \frac{e^{-\lambda}\lambda^x}{(x-2)!} + \sum \frac{e^{-\lambda}\lambda^x}{(x-1)!}$$

$$\Rightarrow E(X^2) = e^{-\lambda} \cdot \lambda^2 \cdot \sum \frac{\lambda^{(x-2)}}{(x-2)!} + e^{-\lambda} \cdot \lambda \cdot \sum \frac{\lambda^{(x-1)}}{(x-1)!} = \lambda^2 + \lambda$$

The above equation implies that the variance of Poisson distribution is $\lambda^2+\lambda-\lambda^2$ (i.e., Var(x) = $E(x^2)$-$E(x)^2$) = λ which is the basic property of a Poisson distribution. Similarly, computing the expectation values for up to n=6 the following equations are obtained. The expectation values for Poisson distribution are as shown below

$$
\begin{aligned}
E(U) &= \lambda \\
E(U^2) &= \lambda + \lambda^2 \\
E(U^3) &= \lambda + 3\lambda^2 + \lambda^3 \\
E(U^4) &= -\lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4 \\
E(U^5) &= \lambda + 15\lambda^2 + 25\lambda^3 + 10\lambda^4 + \lambda^5 \\
E(U^6) &= -9\lambda + \lambda^2 + 80\lambda^3 + 65\lambda^4 + 15\lambda^5 + \lambda^6
\end{aligned}
\qquad (9)
$$

Where,

λ is the mean of the random variable, and

U is a random variable representing the cyclic counts.

The expectation values in Equation (9) are applied onto the counts and, these count expectation values have to be translated to that of the hourly volumes and then the degree of saturation (X). The count (U), vehicles per cycle, could be related to the degree of saturation X using the following process.

The U in the above equation represents the volume counts for a time period equal to a cycle length (i.e., C secs). Volume is the vehicle count per hour, therefore the counts per cycle have to be converted to the counts per hour. This implies that if the vehicles are counted for every C seconds (where C is the cycle length), the volume will turn out to be (3600/C) times the count.

The following equations are developed:

$$
\begin{aligned}
U &= V \big/ (3600/C) \\
\Rightarrow V &= U \times \frac{3600}{C} \\
\Rightarrow X &= \frac{U \times 3600}{C \times c}
\end{aligned}
\qquad (10)
$$

Where,

U is the count for a period of C secs,

V is the arrival volume, and

c is the capacity.

Using Equation (10), the following expression for the degree of saturation (X) is obtained. The expectation values of U and X can also be related as shown.

$$X = \frac{3600 \times U}{c * C}$$
$$E(X) = \frac{E(U) \times 3600}{C * c} \tag{11}$$
$$E(X^n) = E(U^n) \times \left(\frac{3600}{C * c} \right)^n$$

Using Equation (11) and the expectation values for counts Equation (9), the expectation values for the degree of saturation (X) are obtained. The expectation values of X obtained using the above equations are as presented in Equation (12):

$$E(X) = \lambda \times \left(3600 / C * c\right)$$
$$E(X^2) = \left(\lambda + \lambda^2\right) \cdot \left(3600 / C * c\right)^2$$
$$E(X^3) = \left(\lambda + 3\lambda^2 + \lambda^3\right) \cdot \left(3600 / C * c\right)^3$$
$$E(X^4) = \left(-\lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4\right) \cdot \left(3600 / C * c\right)^4 \tag{12}$$
$$E(X^5) = \left(\lambda + 15\lambda^2 + 25\lambda^3 + 10\lambda^4 + \lambda^5\right) \cdot \left(3600 / C * c\right)^5$$
$$E(X^6) = \left(-9\lambda + \lambda^2 + 80\lambda^3 + 65\lambda^4 + 15\lambda^5 + \lambda^6\right) \cdot \left(3600 / C * c\right)^6$$

Using these expectation values in a simplified delay equation, an approximated polynomial obtained from the Taylor series expansion, the expectation values for delay and its higher powers are obtained. The variance of delay from day-to-day and cycle-to-

cycle is computed. The reasons for simplifying the delay equation and the process involved are discussed in the following section.

3.4.3 Taylor series expansion

The HCM delay equation could be conceptualized as a function of demand volume if the other stochastic variables are considered to be constant. However, it is noted that since expectation values have been developed for power functions or additive and multiplicative terms of these power functions, this method cannot be applied directly on the HCM delay equation. The HCM delay equation has to be transformed to an equation involving only additive and multiplicative terms of the powers of X. This can be realized by approximating the delay equation to a polynomial as a function of X. Taylor series expansion on the delay equation is used for the transformation.

Given that the delay equation is a function of a single stochastic variable X, the delay equation is approximated as a univariate polynomial of X using Taylor series expansion.

The generalized Taylor Series expansion of any function F(x) is:

$$F(X) = F(X_0) + \sum_{n=1}^{j} \frac{1}{n!} \frac{d^n F(X_0)}{dX^n} (X - X_0)^n \qquad (13)$$

Where,

F(X) = the function being approximated, and

X_0 = the point about which the equation is expanded

The X_o in the above equation is the point where the function is expanded. The approximated function F(X) yields values that are very close to the true values around the point X_o and is exact at the point X_o. Since traffic demand volume is a random variable

with a particular mean and variance, traffic demand volumes vary around the mean volume in a pattern depending on the distribution of volume. When Taylor Series expansion (Equation (13)) is used to approximate the HCM delay equation, it is logical to utilize a value of X_0 equal to the mean volume. This will result in the best approximations. It is noted that different mean values yield different approximation equations.

In Equation (13), the F(X) is replaced with D(X) implying that delay is a function of X only. This would yield an approximate equation of the generalized form

$$D(X) = \sum_{j=1}^{n} \alpha_j X^j \tag{14}$$

Where,

D(X) is the delay function,

X is the degree of saturation,

α_j are the constants, and

n is the number of terms or the order of the polynomial.

The n value is determined on the basis of how well the approximation replicates the HCM delay curve within reasonable percentile confidence intervals for X. Usually, for low degrees of saturation, a value of 3 or 4 should suffice. An example of the comparison between the delay equation and the approximation is shown in chapter 4.

3.4.4 Calculation of HCM delay variability

This section presents half of the variance computations module depicted in Figure 1. The methodology involved in the delay variance computations for an undersaturated

intersection, which is the right portion of the variance computation module, is presented here. Figure 2 depicts the delay variance computation in detail through a flow chart.

The steps I and II in Figure 2have already been explained in the sections 3.4.2 and 3.4.3. Step III involves computing the expectation values for the powers of X which has been explained in section 3.4.2. The expectation values for X^r are computed depending upon the probabilistic distribution, mean and variance of the arrival flow as elicited earlier from Equations (5)-(8), (9) and (12). Step IV involves the computation of the expectation values for delay. The expectation values for the delay (D) and for D^2 are calculated from the expectation values of X^r as shown below. Using Equation (14) and the expectation values generated from the Equations (5)-(8), (9) and (12), the expectation of delay is calculated as in Equation (15).

$$E(D) = \sum_{j=1}^{n} \alpha_j E(X^j) \tag{15}$$

Similarly, the delay equation is squared, the expectation of D^2 is computed, and from these values, the variance of delay is calculated as

$$\sigma_D^2 = E(D^2) - [E(D)]^2 \tag{16}$$

Equation (16) is the step V portion of the flow chart where the delay variance is obtained. Furthermore, if the distribution of delay and its percentile values are known, the confidence intervals of delay can be computed. Using the standard deviation of delay (σ) from Equation (16) and mean value of delay (μ) from Equation (15), the confidence interval is computed as follows.

$$\text{C.I.} = \mu \pm \sigma \times (\text{percentile value}) \tag{17}$$

The percentile values are calculated from statistical tables depending upon the distribution and the percentile. For example, in case of Normal distribution, the 95th percentile value is 1.96, while 99.99 percentile uses the value of 3. The day-to-day variability computations are the simplest as the expectation values can be computed directly from the Equations (5)-(8). The cycle-to-cycle variability computations involve the expectation values from the Poisson distribution Equation (12).

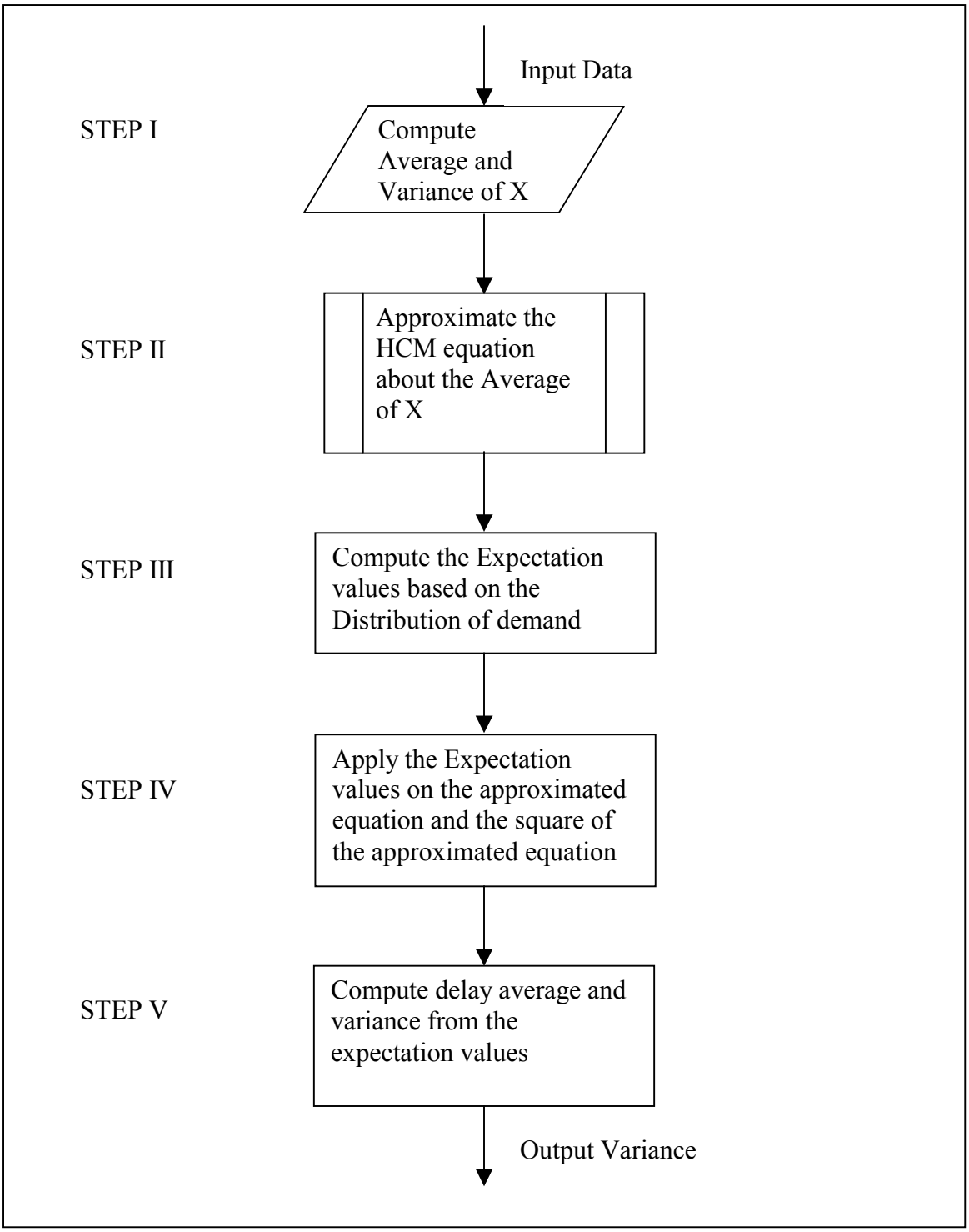

Figure 2. Flow Chart of the HCM delay variance computation for undersaturated intersection

3.5 Oversaturated Intersections

The methodology for undersaturated intersections involved expectation functions. Expectation functions could be used as long as the function is continuous i.e. retains its analytical equation throughout its domain. The methodology for oversaturated intersections differ from that of the undersaturated intersection case because of the discontinuity of the delay equation at X=1. This discontinuity arises from the definition of delay equation provided in the HCM. The HCM delay equation is different for oversaturated and undersaturated conditions as shown in Equation (18). The expectation values cannot be used here as the function changes its equation.

$$D = \begin{cases} \dfrac{0.5 \times C \times (1 - g/C)^2}{1 - X * \dfrac{g}{C}} + 900 \times T \times \left[(X - 1) + \sqrt{(X - 1)^2 + \dfrac{8kIX}{cT}} \right] & \text{When } X<1 \\ 0.5 \times C \times (1 - g/C) + 900 \times T \times \left[(X - 1) + \sqrt{(X - 1)^2 + \dfrac{8kIX}{cT}} \right] & \text{When } X \geq 1 \end{cases} \quad \textbf{(18)}$$

For an undersaturated intersection, by definition the volumes are less than the capacity implying that the degree of saturation (X) is less than one. Considering the stochastic variability in the volumes, if the 99.99^{th} percentile confidence intervals of X are found to be less then 1, the expectation methodology could be used.

Conversely, by the definition of an oversaturated intersection, the volumes are very high at an intersection such that they do exceed the capacity. Furthermore, with the inherent variability in the volumes, the degree of saturation encompasses the regions of $X \leq 1$ and $X > 1$ and delay changes its equation at the value of X=1 and hence expectation methodology provides results that might not be accurate.

Therefore, to compute the variance of delay for an oversaturated intersection, an integration technique is utilized. The flow chart presented in Figure 3 depicts the process of delay variability estimation for oversaturated conditions. A detailed methodology is also explained below.

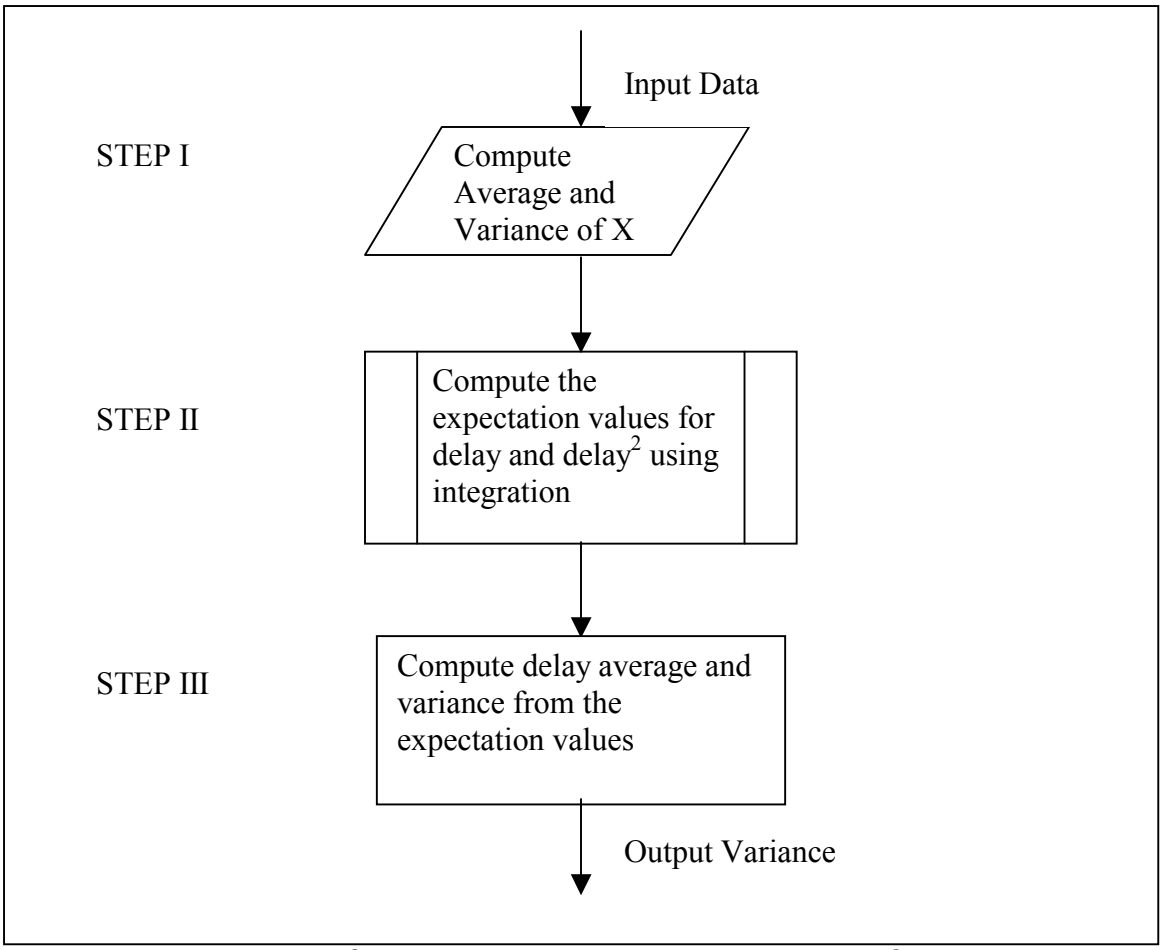

Figure 3. Flow Chart of the HCM delay variance computation for oversaturated intersections

The first step in delay variance computations involves gathering the relevant data and computing the average and variance of the degree of saturation which will be used for further calculations. The next step involves computing the expectation values for delay and delay squared using integration as is explained below.

The expectation value for any function f(x) can be computed from the following integral:

$$E(f(x)) = \int_{Over\ the\ Limits\ of\ X} f(x)P(x)dx \ . \qquad (19)$$

Where,

f(x) = the function for which expectation values are needed, and

P(x) = probability distribution of x.

The above integration is computed within the limits over which x varies and in the present case; x is the degree of saturation (X) and is assumed to be in the range [0, 3] as the degree of saturation is not usually expected to exceed 3.

For the present case, the function f(x) is of the form shown in Equation (18)

$$f(x) = \begin{cases} f_1(x) & \text{When } x<1 \\ f_2(x) & \text{When } x\geq 1 \end{cases}$$

The expectation of f(x) could be computed using the following equation

$$E(f(x)) = \int_{x\leq 1} f_1(x)P(x)dx + \int_{x\geq 1} f_2(x)P(x)dx \qquad (20)$$

The $f_1(x)$ and $f_2(x)$ are replaced by the two equations shown in Equation (18). The p(x) used is the probability distribution of the variable x that is known prior to the

computation. P(x), as is mentioned earlier, is Poisson for cyclic variability and Normal for daily variations.

$$E(f^2(x)) = \int_{x \le 1} f_1^2(x)P(x)dx + \int_{x \ge 1} f_2^2(x)P(x)dx \quad (21)$$

Similarly, the delay equation is squared and the integrals are computed again but this time with the squared values of the functions $f_1(x)$ and $f_2(x)$ and with the same probability distributions as shown in Equation (21). This would give the expectation value of D^2 and the average delay value is obtained from Equation (20).

The next step (step III) involves using the expectation values of delay and delay squared to compute the variance of delay and the variance of delay is computed as in Equations (16). Finally, if the distribution of delay is known with percentile values then the confidence intervals on delay could be computed using Equation (17).

3.6 Arterial Intersections

The delay estimation procedure for an intersection at an arterial is the same as that of an isolated intersection except for the through movements on the arterial link. A different delay equation is applied as is shown in Equation (22). The delay estimation process for through movements at an arterial is different from that of other movements in two aspects. First, arrivals at the downstream intersection depend upon the turning vehicles that join the main stream in addition to the through vehicles at the upstream intersection. Thus the mean and variance of the arrivals downstream have to be estimated based on the component movements mean and variance. Furthermore, the arrival distribution for

through movements at an arterial intersection may or may not be Poisson (for cyclic variations) or Normal (for daily variations) as was assumed for an isolated intersection.

Second, the arrival pattern within a cycle at the downstream intersection is not uniform due to filtering of vehicles at the upstream signal and consequent platoon dispersion along the arterial in addition to the turning vehicles. The arrival pattern at the downstream intersection has to be estimated using the platoon dispersion model (section 3.6.2). The progression factor used in the HCM delay computations for an arterial intersection reflects the arrival pattern onto the delay estimate.

Thus the delay variance estimation for an arterial intersection requires information on the arrival mean, variance and distribution from cycle-to-cycle or day-to-day and the arrival pattern within a cycle at the downstream intersection. The arrival mean, variance and distribution from cycle-to-cycle or day-to-day are assumed to be input by the user. Although the distribution, mean and variance of the arrivals at the downstream intersection can be computed, based on the mean, variance and distribution of the upstream arrivals, the existence of sinks and sources along the arterial change the mean and variance of the arrivals at downstream and complicate the situation. Hence, the arrivals at the downstream intersection along with the mean, variance and distribution are assumed to be input by the user. The arrival pattern within a cycle is estimated using the platoon dispersion model presented in section 3.6.2.

The following sections illustrate the delay estimation procedures in detail and are organized as follows. The HCM equation for the delay estimation for an arterial intersection is presented along with the progression factor estimation procedure. The platoon dispersion model used in the estimation procedure is presented next followed by

the simplifications made to the model. Finally, the procedure for delay variance computation is summarized at the end of the chapter.

3.6.1 HCM Delay Equation for an Arterial Intersection

As is mentioned earlier, HCM delay equation is assumed to be valid. The HCM delay equation for an intersection at an arterial is different from that of an isolated intersection and is shown below.

$$D = \begin{cases} PF \times \left(\dfrac{0.5 \times C \times (1 - g/C)^2}{1 - X \times \dfrac{g}{C}} \right) + 900 \times T \times \left[(X-1) + \sqrt{(X-1)^2 + \dfrac{8kIX}{cT}} \right] & \text{When X<1} \\ PF \times \left(0.5 \times C \times (1 - g/C)\right) + 900 \times T \times \left[(X-1) + \sqrt{(X-1)^2 + \dfrac{8kIX}{cT}} \right] & \text{When X}\geq 1 \end{cases} \quad (22)$$

Where,

PF is the additional term called the progression factor

The progression factor is an indication of the arrival pattern at an intersection within a cycle. Progression factor accounts for proper/ improper progression arrangements in the arterial. This factor is estimated based on the arrivals during green as follows:

$$PF = \frac{(1 - R_p u)\left[1 - u + y_L (1 - R_p)\right](1 - y_L)}{(1-u)^2 (1 - R_p y_L)} \quad (23)$$

Where,

u = Green time ratio (g/C),

y_L = Lane group flow ratio defined as v_L/s_L,

R_p = Platoon ratio defined as v_{Lg} / v_L.

v_{Lg} is the arrival flow rate during the green period, v_L is the average arrival flow rate during the signal cycle, s_L is the lane group saturation flow rate per lane. The factors involved in the progression factor estimation procedure except for the arrival flow rate during the green period are available and this term is computed using the platoon dispersion model. The procedure is presented below.

3.6.2 Platoon dispersion model

The platoon dispersion model used in this methodology is the following [11]:

$$q'(t+\beta T) = F \times q(t) + [(1-F) \times q`(t+\beta T-1)] \quad (24)$$

Where,

q'(t + βT) = Predicted flow rate at the time interval t + βT,

q(t) = Flow rate of the initial platoon during the step t,

T = Cruise time on the link in steps,

β = Empirical factor ~0.8,

F = $(1 + \alpha\beta T)^{-1}$ is the smoothing factor, and

α = Platoon dispersion factor ~0.35.

This model predicts the flow rate at the downstream stop bar (ignoring the presence of queue) for a particular time interval based on the discharges from the upstream signal. The predicted flow rate depends upon the cruise time on the arterial link,

the flow rate in the previous interval and the upstream discharge rate. As is mentioned in section 3.2, all the parameters required for this model are input by the user. The factors involved in the platoon dispersion model are the empirical factor (β) and platoon dispersion factor (α). A value of 0.35 for α best represents the measured dispersion on typical urban streets in the U.S. [11]. The platoon dispersion factor (α) is a site specific factor and a value of 0.35 corresponds to light turning traffic, light pedestrian traffic and 11 to 12 foot lanes. Using the two parameters (α and β) and the travel time on the link, a smoothing factor (F) is calculated. Based on the discharge pattern at the upstream intersection and the smoothing factor, the arrivals at the downstream intersection are calculated.

3.6.2.1 Upstream discharge pattern

The platoon dispersion model requires the upstream discharge pattern as an input to predict the arrival pattern at the downstream intersection. Figure 4 shows the discharge pattern for a simple through flow at the upstream intersection. The vehicles are discharged at a saturation flow rate up to g_o (i.e., until the queue is dissipated). Once the queue is dissipated, the vehicles pass through an intersection as they arrive and hence the discharge rate is v, the rate at which the vehicles arrive. As is mentioned earlier, the saturation flow rate is assumed to be a constant and the arrivals at the upstream intersection are assumed to follow a Poisson distribution.

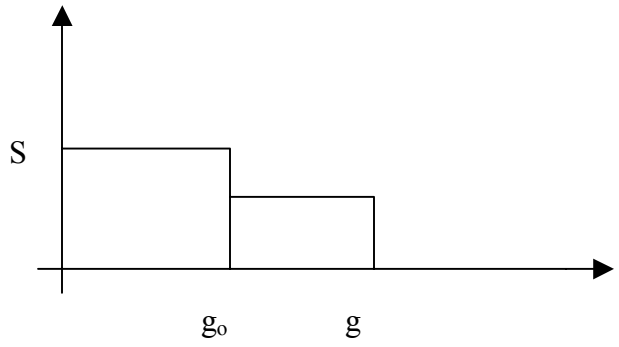

Figure 4. Discharge pattern at an upstream intersection considering only through volumes

However, the left turn volumes that join the arterial add a new component to the discharge pattern. The discharge pattern with the left volumes is as shown in Figure. At the upstream intersection, the left volumes are assumed to receive the green after the

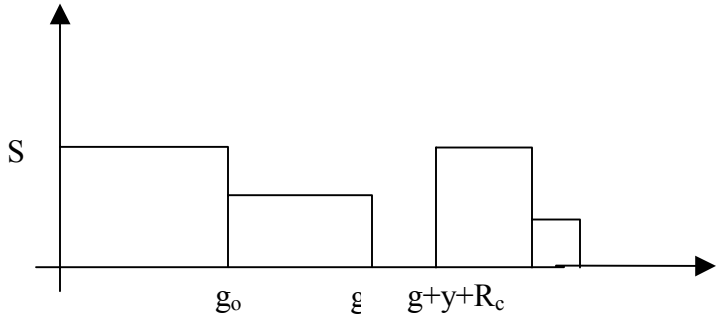

Figure 5. Discharge pattern at the upstream intersection considering through and left vehicles

through volumes. Thus the discharge pattern of the through volumes is repeated again with the left turning volumes after a time period equal to the yellow + red clearance time. This discharge pattern is based on the assumption that the signal timing plan follows a

NEMA phasing where the left turn movement leads the through volumes as is shown in Figure. The movements in dark represent the through and left turning movements at the upstream intersection. The two movement volumes arrive downstream as is shown in the layout in Figure. Furthermore, the left and through arrivals are independent of each other. Hence, the platoon dispersion model can be applied to each of the component discharge patterns separately and later combined.

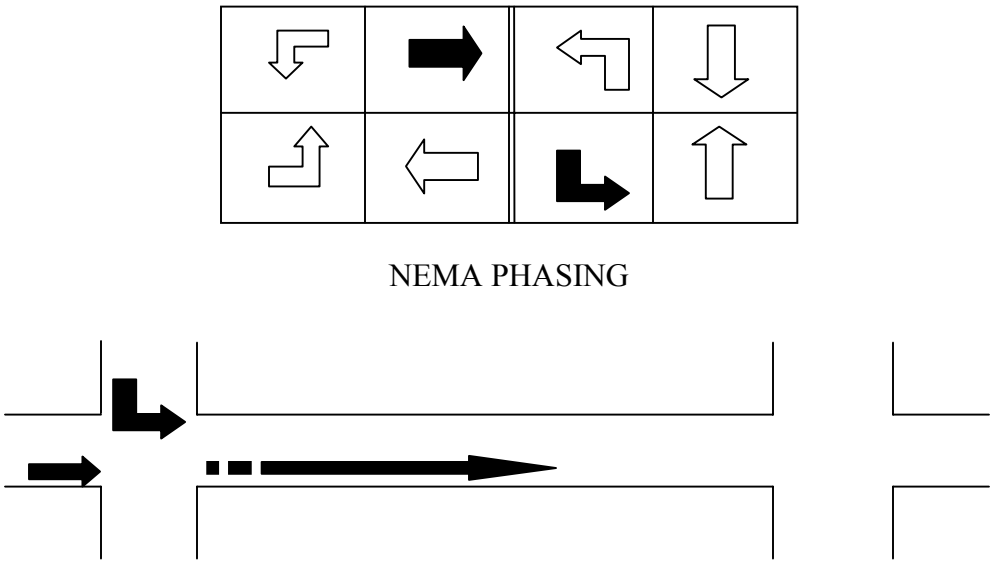

Figure 6. Figure depicting the Platooned arrivals at downstream

Since the movements are independent of each other and follow a similar discharge pattern, a generic platoon dispersion model based on the discharge pattern in Figure 4 is developed and then the volumes at the downstream intersection are cumulated.

3.6.2.2 Simplification of the platoon dispersion model

The platoon dispersion model from Equation (24) can be further simplified by using the equation sequentially to the n^{th} term. That is, the Left Hand Side (LHS) of Equation (24) is q`(t+βT) which can be substituted in the RHS of equation as q'(t+βT-1) for the next time interval. Suppose, the flow is predicted for a time interval 1, using Equation (24), it is evident that q'(1) is depended on q'(0) and a known upstream discharge rate. Furthermore, q'(0) is assumed to be equal to zero as there are no discharges prior to the time t = 0. Thus q'(1) is calculated from q'(0). Similarly, q'(2) could be calculated from q'(1) and so on . Thus by repeating the above process q'(n) is computed. This q'(n), the n^{th} term (i.e., the flow rate for the n^{th} interval) turns out to be a summation of a geometric series. Representing the above process in equations yields the following:

For the first interval at downstream,

$q'(t+\beta T) = Fq(t)$ because q'(t + βT-1) = 0

For the next interval,

$$q'(t+\beta T+1) = Fq(t) + Fq(t)\times(1-F)$$
$$\Rightarrow q'(t+\beta T+1) = Fq(t)\times[1+(1-F)]$$

Similarly, for the next interval, the flow turns out to be $Fq(t)\times\left[1+(1-F)+(1-F)^2\right]$. Therefore, if the discharge flow rate at upstream remains the same as q(t), the flow rate at down stream for the n^{th} interval would be:

$$q'(t+\beta T+n) = Fq(t)\times\left[1+(1-F)+(1-F)^2+.....(1-F)^n\right]$$

The expression in the square brackets on the right hand side (RHS) is the summation of a geometric series with 1 as the first term and a common ratio of (1-F). The RHS would simplify to the following:

$$q'(t+\beta T+n)=Fq(t)\times\frac{1-(1-F)^{n+1}}{1-(1-F)}=q(t)\times\left[1-(1-F)^{n+1}\right]$$

$$q'(t+\beta T+n)=s\times\left[1-(1-F)^{n+1}\right] \quad (25a)$$

Thus, for the upstream discharge pattern assumed in Figure 4, the flow rate downstream for any interval up to g_0 can be computed using Equation (25a).

However, as shown in Figure 4, the discharge flow rate from upstream changes to v at time g_o (time when the queue is dissipated). Therefore using the equation sequentially after the change in q(t), the following equation is obtained.

$$q'(t+\beta T+n+m)=Fq'(t)\times\frac{1-(1-F)^{m}}{1-(1-F)}+q(t)\times\frac{(1-F)^{m}-(1-F)^{m+n+1}}{1-(1-F)}$$

$$=q'(t)\left[1-(1-F)^{m}\right]+q(t)\left[(1-F)^{m}-(1-F)^{m+n+1}\right]$$

$$=v\left[1-(1-F)^{m}\right]+s\left[(1-F)^{m}-(1-F)^{m+n+1}\right] \quad (26a)$$

Where,

n = the interval at downstream where the flow rate changes from s to v at upstream, equal to g_0 expressed in intervals,

m = any interval after the flow has changed from saturation flow to v, and

v = the flow rate in vehicles per hour.

Further, the discharge from the upstream intersection falls down to zero at the end of the green period at the upstream intersection. Using the methodology used above, the platoon dispersion model can be simplified to predict the flow for entire two cycle intervals assuming that q(t) the discharge from upstream is zero after a time g (green time). The equation for the flow rate would be:

$$q'(t+\beta T+n+m+p)=\left[(1-F)^{p}\right]\times\left\{v\left[1+(1-F)^{m}\right]+s\left[(1-F)^{m}-(1-F)^{m+n+1}\right]\right\} \quad (27a)$$

Summarizing the above equations,

$$q'(t+\beta T+n)=s\times\left[1-(1-F)^{n+1}\right] \quad (25a)$$

$$q'(t+\beta T+n+m)=v\left[1-(1-F)^{m}\right]+s\left[(1-F)^{m}-(1-F)^{m+n+1}\right] \quad (26a)$$

$$q'(t+\beta T+n+m+p)=\left[(1-F)^{p}\right]\times\left\{v\left[1+(1-F)^{g-g_0}\right]+s\left[(1-F)^{g-g_0}-(1-F)^{g}\right]\right\} \quad (27a)$$

Where,

n varies from (0, g_0-1),

m varies from (1, g-g_0), and

p varies from (1, 2C-g) this is because the platoon is assumed to disperse within two cycles only (as is mentioned previously).

The above flow pattern represents the flow arriving at a downstream intersection that is dispersed within two cycle lengths. The Equations (25) – (27) are transformed to involve only a single variable n. The downstream arrival flow pattern from time periods 0 to 2C can be represented as:

$$Arrival\ \ Pattern=\begin{cases} s\left[1-(1-F)^{n}\right] & n\in[1,g_0-1] \quad (25b)\\ v\left[1-(1-F)^{n-g_0}\right]+s\left[(1-F)^{n-g_0}-(1-F)^{n}\right] & n\in[g_0+1,g] \quad (26b)\\ \left\{\left[1-F\right]^{n-g}\right\}\times\left\{v\left[1-(1-F)^{g-g_0}\right]+s\left[(1-F)^{g-g_0}-(1-F)^{g}\right]\right\} & n\in(g,2C] \quad (27b) \end{cases}$$

3.6.3 Estimation of the Arrival Pattern and Progression Factor

The arrival pattern at the downstream intersection is computed by first computing the total arrivals at the downstream intersection for one cycle based on the upstream discharge pattern (Figure 4). Then, the template equations are applied to the through and left turning volumes and then the total arrivals during green are computed. From the arrivals during green, the flow rate during the green period is estimated. Using the analytical expressions developed in Equations (25)-(27), cumulative arrivals are computed.

In every cycle, the platooned arrivals at downstream are assumed to disperse within two cycles as shown in Figure 7. That is, vehicles that are discharged from upstream in the present cycle would reach at the downstream intersection in the next cycle. This assumption is based on the premise that if the platoon disperses over more than two cycles, the arrivals would be close to random. Therefore, for any intermediate cycle, the volume at the downstream intersection should consist of two parts one that is the portion of platoon that is departing during current cycle and is dispersed for a time of C (cycle length in secs) (Region I in Figure 7Figure) and the other is the remaining portion of the platoon that has already departed in the previous cycle from upstream (Region II in Figure 7Figure).

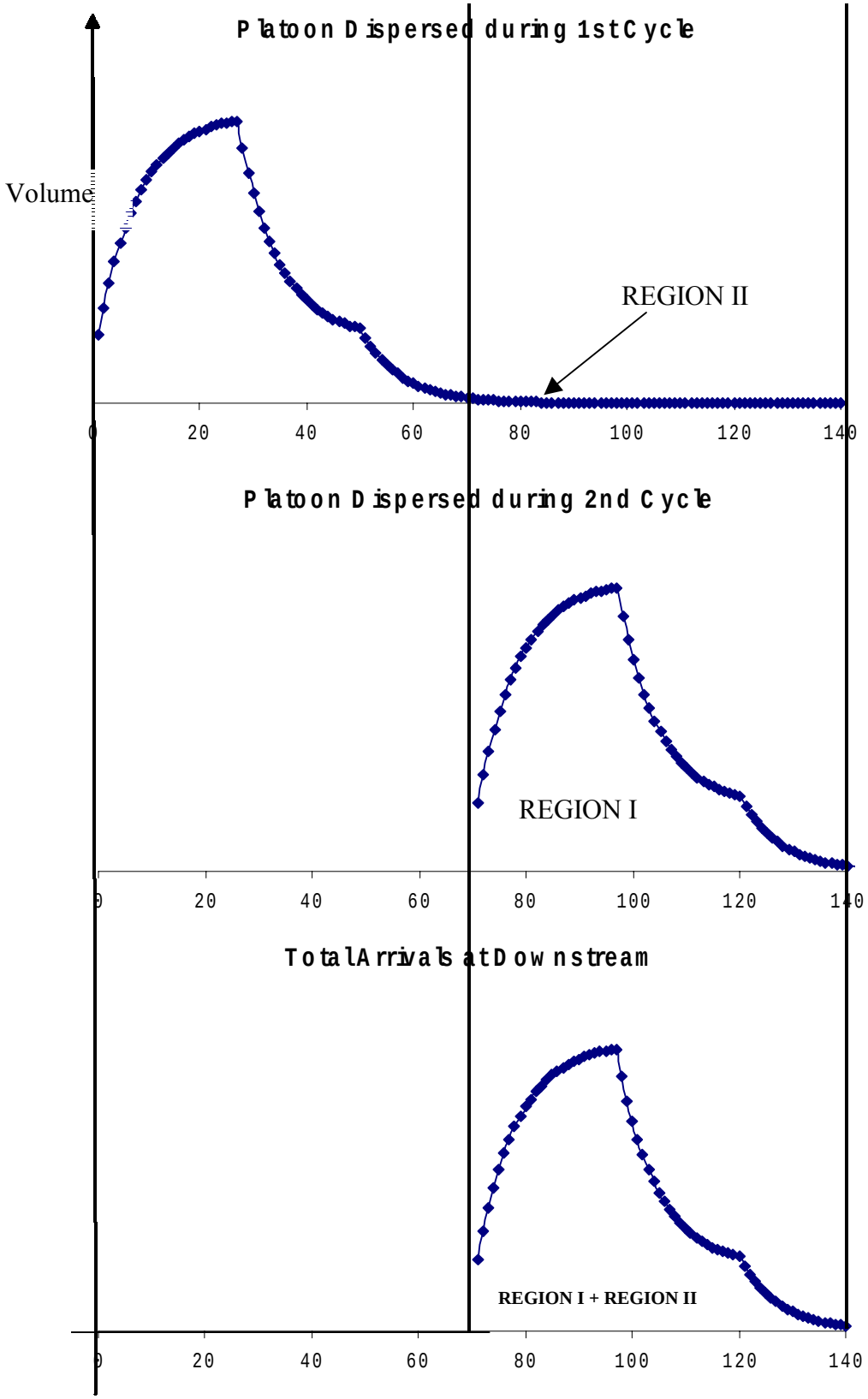

Figure 7. Depicting the platoon dispersion process

The expression for the total arrivals downstream could be presented as shown in Equation (28) – (30). The downstream arrival flow pattern from time periods 0 to C can be represented as shown in Equations (25b) – (27b):

$$Arrival\ \ Pattern = \begin{cases} s\left[1-(1-F)^{n+1}\right] & n \in [0, g_0 - 1] \\ v\left[1-(1-F)^{n-g_0}\right] + s\left[(1-F)^{n-g_0} - (1-F)^{n}\right] & n \in [g_0, g] \\ \left\{\left[1-F\right]^{n-g}\right\} \times \left\{v\left[1-(1-F)^{g-g_0}\right] + s\left[(1-F)^{g-g_0} - (1-F)^{g}\right]\right\} & n \in (g, C] \end{cases}$$

Similarly, the arrival pattern for the vehicle that would arrive downstream in the next cycle could be represented as

$$Arrival\ \ Pattern = \left\{\left[1-F\right]^{n+C-g}\right\} \times \left\{v\left[1-(1-F)^{g-g_0}\right] + s\left[(1-F)^{g-g_0} - (1-F)^{g}\right]\right\} \quad n \in [0, C]$$

As explained earlier, the arrivals during the two cycles are added to compute the effective arrivals at the downstream intersection as

$$Arrival\ \ Pattern = \begin{cases} s\left[1-(1-F)^{n+1}\right] \times \left\{\left[1-F\right]^{n+C-g}\right\} \times \begin{Bmatrix} v\left[1-(1-F)^{g-g_0}\right] + \\ s\left[(1-F)^{g-g_0} - (1-F)^{g}\right] \end{Bmatrix} & n \in [0, g_0 - 1] \quad [28] \\ v\left[1-(1-F)^{n-g_0}\right] + s\left[(1-F)^{n-g_0} - (1-F)^{n}\right] \times \left\{\left[1-F\right]^{n+C-g}\right\} \times \\ \left\{v\left[1-(1-F)^{g-g_0}\right] + s\left[(1-F)^{g-g_0} - (1-F)^{g}\right]\right\} & n \in [g_0, g] \quad [29] \\ \left\{\left[1-F\right]^{n-g}\right\} \times \left\{v\left[1-(1-F)^{g-g_0}\right] + s\left[(1-F)^{g-g_0} - (1-F)^{g}\right]\right\} \times \\ \left\{\left[1-F\right]^{n+C-g}\right\} \times \left\{v\left[1-(1-F)^{g-g_0}\right] + s\left[(1-F)^{g-g_0} - (1-F)^{g}\right]\right\} & n \in (g, C] \quad [30] \end{cases}$$

Equations (28)-(30) represent the arrivals at downstream for any intermediate cycle.

3.6.3.1 Combining the through and left turning volumes at downstream intersection

Equations (28) – (30) represent the arrival pattern at the downstream intersection for the upstream discharge pattern presented in Figure 4. These Equations (28) – (30) are used separately with the upstream through and the left turning volumes to compute the arrivals at the downstream intersection. The volumes are then combined according to the vehicular arrival times at the downstream intersection.

According to the platoon dispersion model, the first vehicle arrives at the downstream intersection at a time equal to βT after the start of green at the upstream intersection. Where β is the parameter used in the platoon dispersion model and T is the travel time estimated from the ratio of the distance between the intersections and the free flow speed. From the knowledge of the arrival times for both the upstream left and through volumes, and the offset between the intersections, the total arrivals during green are computed by cumulating the arrivals during the downstream green time.

Based on the above arrival pattern, with a knowledge of the downstream green time and cycle length (usually same as that of the upstream intersection when coordinated), the percentage of arrivals during green are computed as shown below. Equations (28)-(30) could be generalized as a function q(t) which represents the arrival pattern as a function of time. The proportion of arrivals during green is computed using the ratio of the arrivals during green to that of the total arrivals during the cycle. This computation is represented as follows:

$$P_g = \frac{\sum_{t=o}^{g_d} q(t)}{\sum_{t=o}^{C} q(t)} \qquad (31)$$

Where,

g_d is the green time for the major street movement at the downstream intersection,

C is the cycle length.

The numerator in Equation (31) is the arrivals during green and this is expressed in vehicles per hour to get the arrival flow rate during the green period.

3.6.4 Delay Variance Computations

The flow rate during the green period is computed using the simplified platoon dispersion model as presented so far. The progression factor is computed using the average volume over a cycle, green time, cycle length. This would yield a number that is to be multiplied by the Uniform delay (D1) term of the HCM delay equation.

With the computed progression factor, the HCM delay equation presented in Equation (22) can be used to compute the delay variance. For undersaturated conditions, the expectation methodology is utilized with the HCM equation for an arterial. Equation (22) is expanded as a polynomial as a function of X. The expectation values of X are computed from the mean, variance, and distribution of the arrivals using Equations (5)-(8) and (12). The generated expectation values are used on the approximated HCM equation and using Equations (15) – (17) the delay variance is computed.

For an arterial under oversaturated conditions, the progression factor is estimated and multiplied to the D1 term of the HCM delay equation as in Equation (22). Using this equation and the mean, variance, and distribution of the arrival volumes, delay variance is computed in a method similar to that of an isolated intersection as shown in Equations (18), (20), and (21).

3.7 LOS Computations

The level of service (LOS) at signalized intersections is determined based on the average delay ranges that are predefined in the HCM. Six ranges have been defined in the HCM that correspond to six different LOS levels (A – F). The following table presents the HCM delay ranges with the corresponding LOS level.

Table 1. HCM level of service criteria

Delay (secs/veh)	LOS
Less than 10	A
10 - 20	B
20 - 35	C
35 - 55	D
55 - 80	E
Greater than 80	F

For example control delay values between 20 and 35 seconds per vehicle are defined as LOS C.

However, delay is a stochastic variable. This implies that delay has a mean about which it varies according to its distribution and variance. Thus the intersection could perform in multiple LOS ranges depending on the delay variability. Hence, based on the mean and variance of delay, (assuming that delay follows a normal distribution) the probability of different LOS ranges could be quantified.

Consider the following figure which shows the delay as a normally distributed variable.

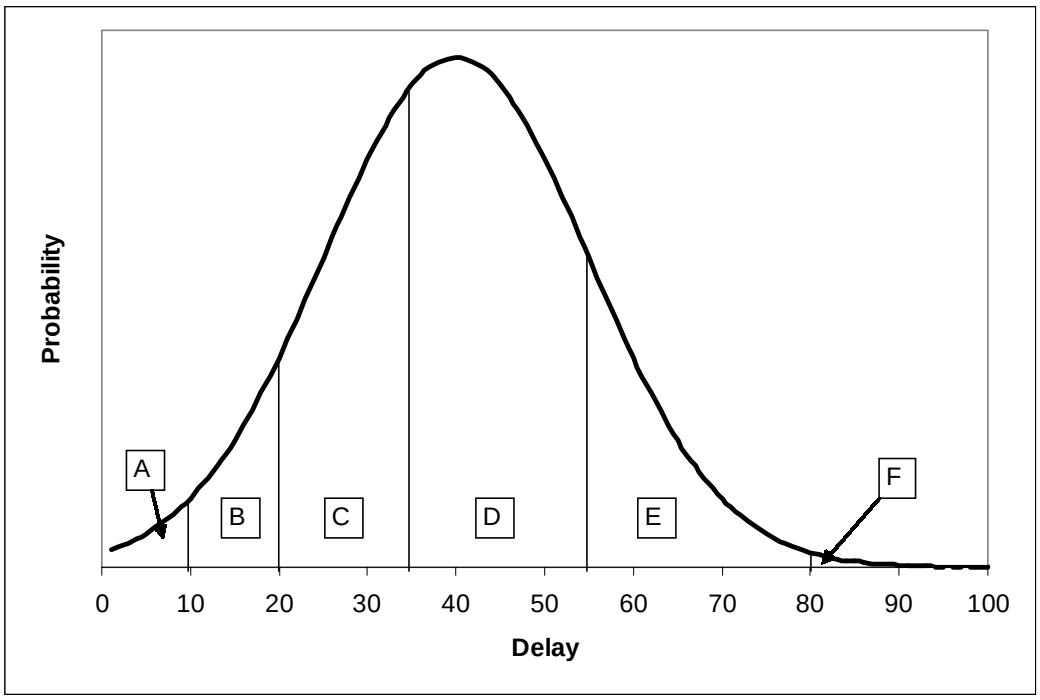

Figure 8. LOS ranges with delay distribution

The area under the probability curve for delay within different LOS limits gives the probability that the intersection performs at that LOS value. Thus, the probability of the intersection performing at LOS A is equal to the probability that the delay is less than 10 seconds per vehicle. Similarly, based on the mean and variance of delay, the probability of occurrence of different LOS levels is quantified. Thus, based on the computed probabilities, a new LOS for the intersection could be defined that would look like the following: A_{pA}-B_{pB}-C_{pC}-D_{pD}-E_{pE}-F_{pF} where the alphabets in capitals correspond to the LOS level, the subscript corresponds to the probability of occurrence at that LOS level.

Furthermore, based on the probability of occurrence of different LOS levels, a new index is developed to quantify the delay variability. This index called level of service performance index (LOSPI) is computed as shown below.

$$LOSPI = \sum_{Over\ all\ LOS\ levels} (Median\ Delay_{LOS}) * (P_{LOS})$$

where,

Median $Delay_{LOS}$ is the median of the delay interval, and

P_{LOS} is the probability of occurrence of the LOS.

3.8 Optimization

The delay variability estimation processes for isolated and arterial intersections were discussed in the previous sections. The next objective of this research is to optimize intersections with stochastic variability. Genetic algorithms are used to realize the optimization process for signalized intersections. Both average delay and the 95th percentile delay are considered as objective functions to optimize signal timing plan at an intersection with stochastic variability. Different scenarios are optimized and compared with SYNCHRO optimized timing plans by means of microscopic simulation programs CORSIM and SIMTRAFFIC. The following sections provide an introduction to the genetic algorithms and the coding procedure adopted. The results of the optimization for an isolated intersection and arterial are provided in Chapter 7.

3.8.1 Optimization procedure

The proposed methodology uses the 95th percentile delay or the average delay value as the objective function in the optimization process. This would provide a signal-timing plan that optimizes intersection performance for the 95th percentile vehicular delay. The mean and variance of delay are obtained either from the expectation functions or the

integration methodologies depending upon the lane group v/c ratios. The 95th percentile delay is computed using the Equation (17) in chapter 3 assuming that delay follows a Normal distribution.

Genetic algorithms are used to optimize an intersection signal timing plan. The general process of genetic algorithms has been outlined in the literature review section. The coding process and implementation procedure are presented here. The coding process is different for isolated and arterial intersections and hence is presented separately.

3.8.1.1 Isolated intersection

An isolated intersection with eight NEMA phases is considered as shown in Figure 9. The phase numbers are the NEMA standard phase numbers.

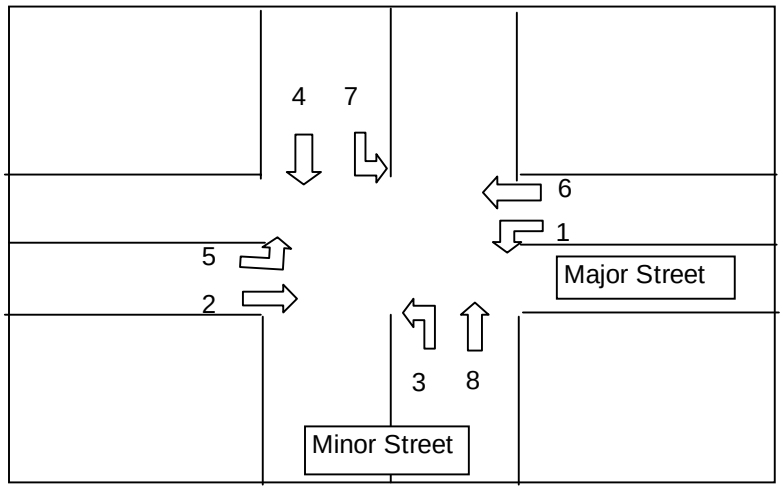

Figure 9. NEMA phase notation

Firstly, a population of n individuals (binary strings) is generated at random. The number of individuals in a population is decided depending on the type of intersection being optimized (isolated / arterial intersection). A value of n equal to 20 is chosen for an isolated intersection and a value of n ranging from 25 to 50 is chosen for an arterial

intersection depending on how well the GA converges to the optimal solution. An individual is a 36 digit binary string that is used to code one cycle length and eight green times for eight different NEMA movements for an isolated intersection. Every individual can be conceptualized as a set of signal timing conditions. The 36 digit binary string can be decoded to generate a set of feasible signal timing conditions and based on the green times generated for all the movements and the input volume conditions and average (or the 95th percentile) delay values are computed.

A coding procedure has to be adopted to generate a set of feasible green times. The coding procedure used is presented below. The 36 digit binary string is broken down into 6 strings of 6 bits each. Six fractions are generated from these strings by converting the binary string to a decimal number and dividing the number by 63 (the max possible decimal number with a 6 digit binary string). The six fractions are used to code the green times and cycle length as shown by the following equations

$$Cycle = MIN + Int((MAX - MIN) \times f(0)) \quad (32)$$

Where,

Cycle = cycle length,

MIN = minimum allowed cycle length,

MAX = maximum allowed cycle length, and

f(0) = random number generated.

Equation (32) constrains the random cycle lengths generated through the string within the maximum and minimum allowable cycle lengths. Using the cycle length

generated above, the green times are assigned to the eight movements. The left turn green times are assigned using the following equations:

$$\begin{aligned} green(1) &= 10 + Int((Cycle - 50) \times f(1) \times f(2)) - 4 \\ green(3) &= 10 + Int((Cycle - 50) \times (1 - f(1)) \times f(4)) - 4 \\ green(5) &= 10 + Int((Cycle - 50) \times f(1) \times f(3)) - 4 \\ green(7) &= 10 + Int((Cycle - 50) \times (1 - f(1)) \times f(5)) - 4 \end{aligned} \quad (33)$$

Where,

green(i) = effective green time for the NEMA phase i,

Cycle = cycle length generated in Equation (32), and

f() = random numbers/fractions generated.

The 10 seconds used in the Equation (33) is the minimum green time assigned to left turn movement. The 50 seconds is the sum of all the phases' minimum greens in a cycle. The 4 seconds is the lost time assumed to convert the display green to effective green for delay computations. Thus, these green times are generated within a feasible range.

Similarly, the through green times are also coded as shown in the Equation (34). The 15 seconds is the assumed minimum green for through movements.

$$\begin{aligned} green(2) &= 15 + Int((Cycle - 50) \times f(1) \times (1 - f(2))) - 4 \\ green(4) &= 15 + Int((Cycle - 50) \times (1 - f(1)) \times (1 - f(4))) - 4 \\ green(6) &= 15 + Int((Cycle - 50) \times f(1) \times (1 - f(3))) - 4 \\ green(8) &= 15 + Int((Cycle - 50) \times (1 - f(1)) \times (1 - f(5))) - 4 \end{aligned} \quad (34)$$

Where,

green(i) = effective green time for the NEMA phase i,

Cycle = cycle length generated in Equation (32), and

f() = random numbers/fractions generated.

These green times along with the volumes are used to compute the intersection delay and variance (as shown in the chapters 4 and 5) and the 95^{th} percentile delay value is calculated using Equation (17) to be used in the optimization procedure. The process is repeated for all the n individuals in the population. Thus n numbers of 95th percentile delay values are obtained. The offsprings are selected from the population of n using a tournament selection. In a tournament selection procedure four individuals are selected at random and the best individual (i.e. an individual with the best fitness value) is selected. The best solution from the four individuals is the one with a minimum average delay. This process is repeated n times to generate n offsprings.

These n individuals are mated at random. That is, individuals are selected at random and mated. A single point crossover procedure is utilized as mentioned in chapter 2. The individuals are then mutated to jump out of local minima. The fitness values of these individuals are then computed again and the entire process is repeated for a fixed number of generations or until a user defined criteria is met. The individual with the best fitness value in the final generation is selected as the optimum solution and results are provided after decoding the string. Thus, GA generates a solution, an individual with 36 digit binary string which results in minimum delay for the given set of conditions. The above process is realized through a C++ code. The results of the optimization process are presented in the chapter 7.

3.8.1.2 Arterial intersections

An arterial with two intersections is considered initially. As was observed with the isolated case, there are eight NEMA phase green times for every intersection. Therefore, for a network with two intersections, sixteen green times, a cycle length and an offset value have to be coded using the binary strings. The coding of the fractions used for that of the isolated intersection case is repeated for both the intersections. A binary string of length 72 is generated to produce twelve random numbers for generating the sixteen green times, cycle length and offset. The coding of the 72-digit string is as follows:

The string is broken down into 12 strings of six digits each. Each of the 12 strings are converted to a decimal number and divided by 63(i.e., $(2^6 - 1)$, the maximum possible decimal number with a six digit binary number) to generate random fractions lying between 0 and 1. The cycle length is coded using the first fraction as in Equation (32). The next fraction is used to code the offset value using the following Equation (33)

$$Offset = f(1) \times (Cycle - 1) \tag{35}$$

This would result in an offset value that lies between 0 and the cycle length minus one. The remaining 10 fractions are used to code the green times for the two intersections using the Equations (33) and (34). These green times along with the volumes are used to compute the intersection delay and variance. The process is same as that of an isolated intersection. The offsprings are generated using a tournament selection and a fitness value of inverse of delay is used. A single point crossover procedure is utilized for mating and then the individuals are mutated. This process is repeated for a fixed number

of generations or until a certain criteria is met. A C++ code is used for the optimization procedure. The results of the optimization process are presented in the chapter 7.

3.9 Evaluation of the optimization result

Signalized intersections subject to stochastic variability are optimized using the genetic algorithms in conjunction with delay variance computation methodologies. The result from GA is compared with the result of the signal optimization program SYNCHRO. Synchro is an optimization program that optimizes intersections for percentile delay. The signal timing plans from both the optimization programs are evaluated using unbiased microscopic simulators. The microscopic simulation programs chosen for the evaluation are SIMTRAFFIC and CORSIM. The following sections introduce the evaluation methodologies.

3.9.1 SIMTRAFFIC evaluation

Since signalized intersections subject to stochastic variability are being optimized, the result of the optimization should be tested under varying demand conditions. SIMTRAFFIC has the option of generating different (pre-specified) percentile volumes for different time intervals. This option in SIMTRAFFIC has been made use of and random percentile volumes are generated for a total time period of 225 mins. These volume conditions are then run using the optimized timing plans from SYNCHRO and GA and the delay results are compared. The results of the comparison are provided in the chapter 7.

3.9.2 CORSIM evaluation

CORSIM is a microscopic simulator capable of conducting multiple simulation runs of a particular demand condition. Multiple runs up to 100 could be conducted for every volume condition. This functionality of CORSIM has been made use of in the evaluation process.

Firstly, as was previously mentioned GA conducts a stochastic optimization of a signalized intersection. That means that the GA timing plan has to work well under a range of demand volume conditions. Hence different combinations of demand volume conditions have been generated using a Latin Hypercube sampling procedure. The resulting combinations of volumes are representative of the demand variability and encompass the whole domain of demand variability. The examples involving Latin Hypercube design were carried out using ACED (Algorithms for the Construction of Experimental Designs; Welch 1996). ACED is UNIX based software where it is fairly easy to add new criteria to this software which can interface with all implemented design optimization algorithms. A hundred combinations of demand conditions are generated based on the mean and variance of the lane group volumes. Every combination is run in CORSIM using the signal timing plans from SYNCHRO and GA for five times. The average delay value of these five runs is used in the evaluation procedure. A hundred different average delay values are compared by movement and by approach. Finally these results are tested statistically and are presented in chapter 7.

CHAPTER 4. DELAY VARIABILITY FOR UNDERSATURATED ISOLATED INTERSECTIONS

This chapter explains through examples the delay variance computations using the expectation methodology for an isolated intersection. Firstly, a lane group at an intersection is considered and delay variance computations are provided for the lane group subject to day-to-day variability. The result of the example is verified through Monte Carlo Simulation. Then the delay variability is studied with the degree of saturation for a lane group with daily demand variations. Finally, the mean and variance of delay are compared for different input demand distributions.

4.1 Example Delay Variance Estimation For an Undersaturated Intersection

The delay variance estimation process follows the flow chart in Figure 2 and every step will be highlighted for easy readability. A hypothetical situation is assumed with the numerical values assigned as shown below. Consider a lane group in a pretimed signalized intersection with the following conditions

Average volume	=	300 vph
COV (Coefficient of Variance) of volume	=	0.25
Cycle length	=	100 sec
Effective Green time	=	30 sec
Saturation flow rate	=	1800 vphplg

The traffic volume and v/c ratio are assumed to follow normal distribution.

The analysis period was 15 minutes, i.e., T = 0.25

For a pretimed signalized intersection K = 0.5

Also for an isolated intersection, I = 1

STEP I: Computation of average and variance of the degree of saturation (X)

The HCM delay equation is approximated using Taylor Series expansion about the value of X corresponding to the average demand. Therefore, the X_0 in Equation (13) has to be computed from the ratio of volume to capacity. The capacity can be found as follows: Capacity (c) = $s \times \frac{g}{C} = 1800 \times \frac{30}{100} = 540 \ \ vphpl$

Average X = Average volume / Capacity = 300 ÷ 540 = 0.55.

Standard deviation of Volume = 300 × 0.25 = 75

Variance X = Variance volume / Capacity^2 = 75^2 / 540^2 = 0.019

Standard deviation X = √0.019 = 0.138

STEP II: Approximate the HCM delay equation about 0.55 (Average X)

Using the Equation (13) and substituting all the above values one can obtain the approximated equation (i.e., Taylor Series equation) as follow.

Average delay = $18.4467 + 48.6496 \times X - 76.8930 \times X^2 + 68.4761 \times X^3$ (36)

An average delay of 33.48 seconds per vehicle is obtained from HCM delay Equation (3). Equation 36 provides a delay value of 33.48 at X=0.58 which is the exact value provided by the approximate delay curve.

The equation is expanded only to four terms (n = 4) as it is observed to replicate the HCM delay equation with a fair degree of accuracy, within the 95th percentile confidence interval. This is determined by comparing the delay curve with the approximation and Figure 10 shows the comparison.

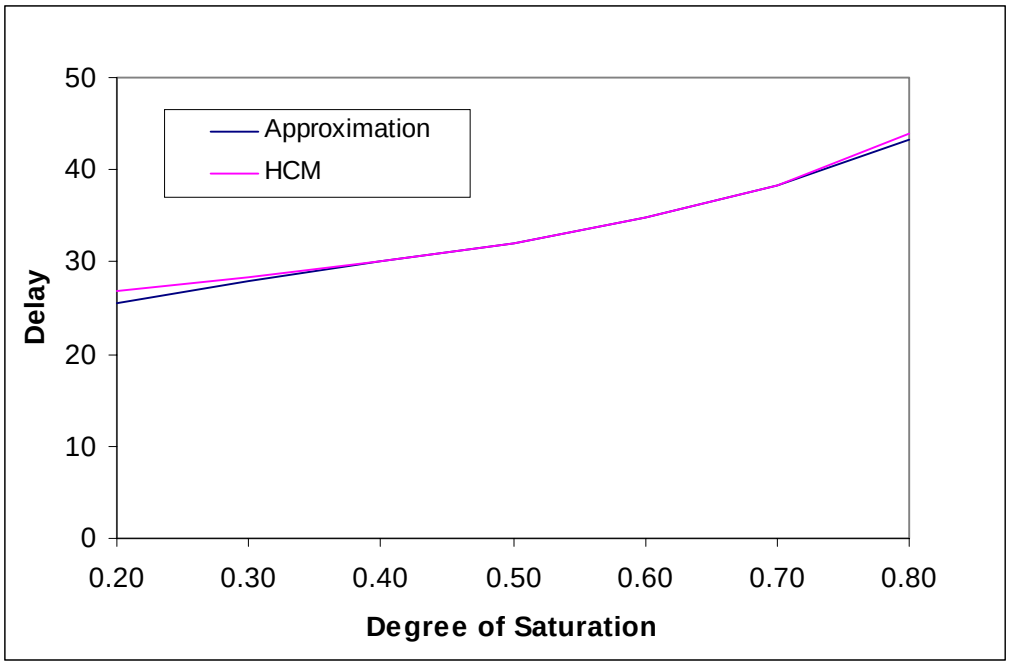

Figure 10. Comparison between approximate equation and HCM delay curve n=5

The 95th percentile confidence intervals for X turn out to be (0.3, 0.8) and within this range, the delay curve is well represented by the approximation. However, a higher value of n could be utilized to get an exact replication.

STEP III: Compute the expectation values of X based on the distribution

Given that X follows a normal distribution, and using the mean and variance (or COV) from step I and Equation (5) following values are obtained.

$$E(X) = \mu = 300/540 = 0.556$$
$$E(X^2) = \mu^2(1+CV^2) = 0.328$$
$$E(X^3) = \mu^3(1+3*CV^2) = 0.204$$
$$E(X^4) = \mu^4(1+6*CV^2+CV^4) = 0.131$$
$$E(X^5) = \mu^5(1+10*CV^2+5*CV^4) = 0.087$$
$$E(X^6) = \mu^6(1+15*CV^2+15*CV^4+CV^6) = 0.059$$

STEP IV: Apply the expectation values onto delay and delay squared

E(D) can be computed using Equation (14) as follows using the approximated Equation (36).

$E(D) = 18.4467 + 48.6496 \times E(X) - 76.8930 \times E(X^2) + 68.4761 \times E(X^3) = 34.20$

Similarly squaring the delay equation and applying the expectation values, gives the following result:

$E(D^2) = 340.28 + 1794.85 \times E(X) - 470.06 \times E(X^2) - 4955.31 \times E(X^3) + 12575.20 \times E(X^4) - 10530.67 \times E(X^5) + 4688.98 \times E(X^6) = 1171.7$

STEP V: Compute delay average and its variance from the expectation values

Using the expectation values generated in Step IV and the Equation (16), a delay variance of 15.06 is produced. The average delay is nothing but the expectation of delay and is equal to 34.20. The 95th percentile confidence interval assuming Normal distribution on delay, $34.20 \pm 1.96 \times 3.88$ or (26.59, 41.80), is obtained from Equation (17).

4.2 Evaluation of Expectation Function Method

In order to verify the proposed expectation function method for day-to-day variability, Monte Carlo simulation is utilized. For the above example, 500 data points that follow a Normal distribution with mean of 300 and coefficient of variance of 0.25 are generated. These data points can be considered as realizations of hourly traffic volumes over 500 days. Then, delays for 500 days are calculated using HCM delay Equation (3). The mean delay and its variance are calculated from these 500 delay values. This process is

repeated twenty five times and the frequency histogram is plotted for all the twenty five cases. These histograms are compared to the distribution generated from expectation function method as shown in Figure 11.

As in the example in the previous section 4.1, the expectation methodology produced a mean value of 34.20 and variance of 15.06. For presentation purpose, it is assumed that the delay follows a normal distribution. Note that delay could follow any distribution with a mean of 34.20 and a variance of 15.06.

As indicated earlier, Monte Carlo simulation generates inconsistent results. The Monte Carlo simulation could produce the mean and variance that are very close to the expectation function method if a large number of random samplings were utilized. As shown in Figure 11, the expectation function method based distribution is well represented by distributions generated from Monte Carlo simulation.

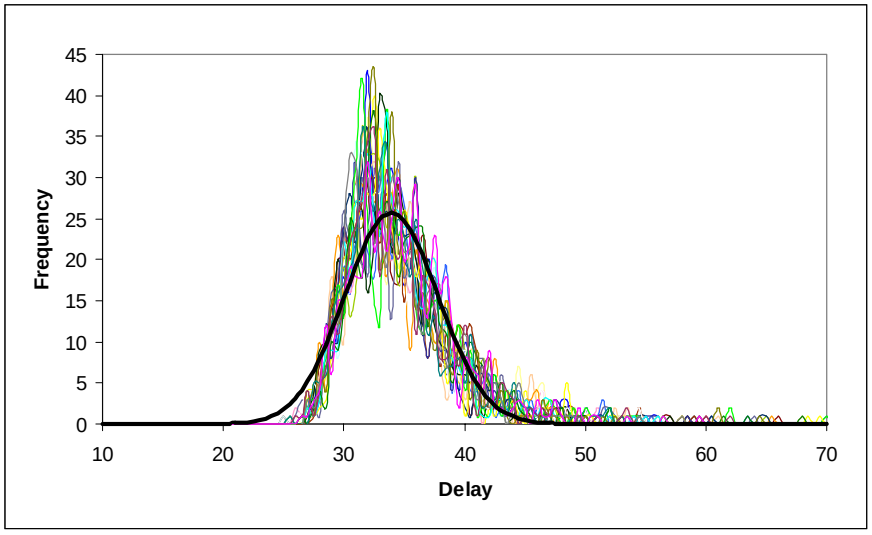

Figure 11. Expectation Function Method versus Monte Carlo Simulation

4.3 Evaluation of the Variance of Delay under Different Demand Conditions

The variance of delay is computed for demand conditions ranging from 0.2 to 0.7 for the example problem used previously (section 4.1). Inputting different X_0 values into the Taylor series expansion and applying the expectation values on these different curves provide the mean and variance (Repeating the procedure shown above using the Equations (5)-(16). Figure 12 shows the 95th percentile confidence intervals on delay for different degrees of saturation.

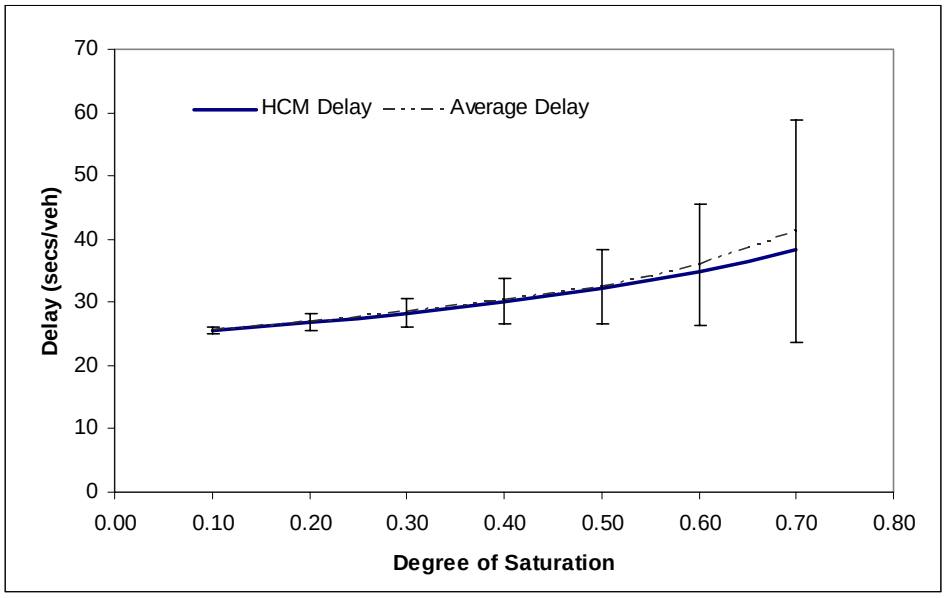

Figure 12. Confidence intervals on delay for different degrees of saturation

Some of the features of Figure 12 are that as the degree of saturation increases, the variability in the delay increases. This is shown by the increase in the confidence interval range as the degree of saturation increases.

As shown in Figure 12, the HCM delay and the average delay show discrepancies when the degree of saturation is higher than 0.6. It should be noted that the discrepancies is due to the nature of HCM equation – the delay increases exponentially when degree of saturation is higher than 0.6. The HCM delay line is obtained using average traffic volume per each degree of saturation in computing delay, while the average delay line is obtained from the average value of HCM delays calculated from varying demands per each degree of saturation. Thus, when degree of saturation value is higher than 0.6, the impact of volume that is bigger than average would result in exponentially higher delay values. Therefore, the average delay from varying demand would be higher than the delay at average volume level. However, when degree of saturation is less than 0.6 or so the HCM delay and average delay yield very similar values since the HCM delay equation is close to linear. It is also noted that Figure only shows the degree of saturation up to 0.7 since the confidence interval of the degree of saturation 0.7 or higher is almost meaningless due to very wide interval range and also, this methodology is designed for undersaturated intersections

4.4 Evaluation of the Mean and Variance of Delay For Different Distributions

The expectation function method can be used for any distribution of the input volumes as long as distribution parameters are known. These distributions include normal, gamma, uniform and lognormal distributions to name a few. In this section, we utilized normal, uniform, gamma, and lognormal distributions in order to explore HCM delay and its variability.

An example is utilized. Four different traffic volume distributions are applied under the same conditions of traffic demand and capacity. Also, the traffic demand

conditions are consistent in all the examples. The degree of saturation is varied from 0.0 to 1.0. As shown in Figure 13, regardless of distributions average delay appears to be fairly consistent. However, variability varies significantly as volume to capacity ratio closes to 1.0 (see Figure 14). As one would expect, lognormal distribution produced the highest delay variance.

The lognormal distribution is a well spread-out distribution with a big tail on both sides causing high variability, while the gamma distribution is mostly skewed positively with a very long tail on one side resulting in a lower variance than that of lognormal. The uniform distribution is uniform throughout the entire range and hence yields a lower variability than those of the above two distributions. The normal distribution has neither any skewness nor big tail and hence has the lowest variability.

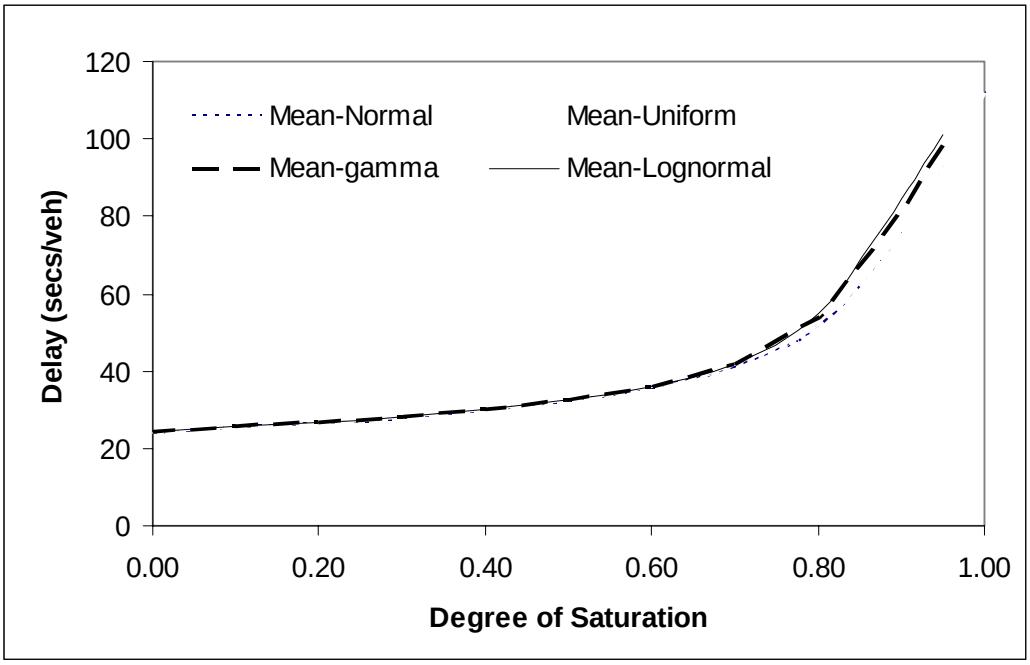

Figure 13. Expected delay under selected input distributions

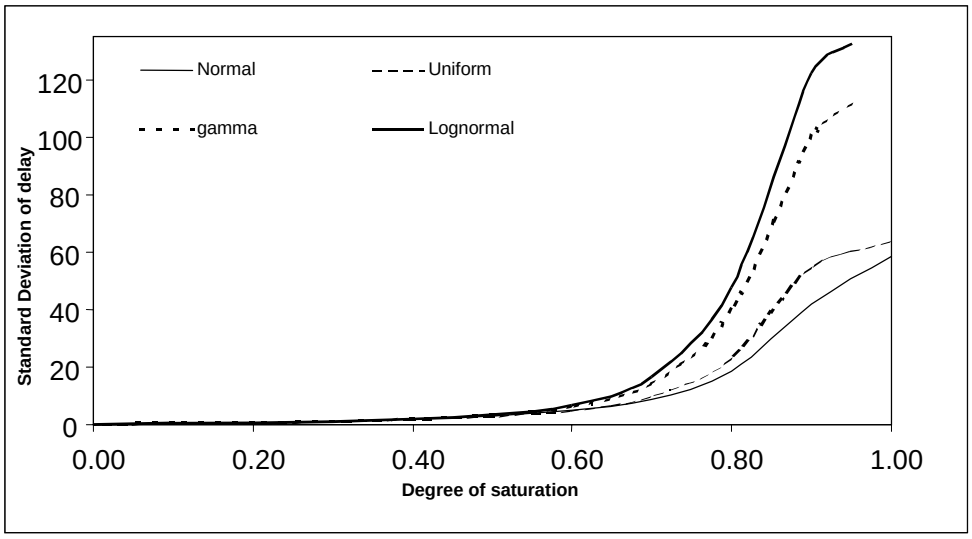

Figure 14. Standard deviation of delay for different distributions of the input volume

4.5 Evaluation of Variance of Delay under Different Variance Conditions

The example provided previously was studied for different coefficient of variance (CV) values. The delay variability was calculated for CV values ranging from 0.2 to 0.6 and the results are provided in the Table 2. The new LOS values are also defined with the Performance Index. It is evident that as the variability increases, the performance of the intersection decreases.

Table 2. Delay Confidence Interval with respect to Coefficient of Variance

COV	Delay (sec/veh)	SD	Upper C.I.	Lower C.I.	Current HCM LOS Range	Proposed LOS Range	LOSPI
0.1	33.60	1.49	36.52	30.68	C-D	$C_{0.83}$-$D_{0.17}$	30.47
0.2	33.94	3.05	39.92	27.96	C-D	$C_{0.64}$-$D_{0.36}$	33.80
0.3	34.52	4.75	43.83	25.20	C-D	$C_{0.54}$-$D_{0.46}$	35.55
0.4	35.32	6.67	48.39	22.26	C-D	$B_{0.01}$-$C_{0.47}$-$D_{0.52}$	36.47
0.5	36.36	8.86	53.72	18.99	B-C-D	$B_{0.03}$-$C_{0.41}$-$D_{0.54}$-$E_{0.02}$	37.37
0.6	37.62	11.41	59.98	15.26	B-C-D-E	$B_{0.05}$-$C_{0.35}$-$D_{0.53}$-$E_{0.07}$	38.95

CHAPTER 5. DELAY ESTIMATION FOR OVERSATURATED INTERSECTION

The methodology involved in the delay variance computations for oversaturated intersections are explained in the Chapter 3 and is highlighted through the Figure 3. The methodology proposed for oversaturated intersections uses integration techniques to quantify the HCM delay variability. A C++ program is coded to compute the integrals using approximation techniques.

This chapter provides the results of this methodology for oversaturated intersections. First, single approach delay variability is computed and the variance values are verified through Monte Carlo Simulation for cycle-to-cycle variability condition. Then, the average and variance of delay are studied for different degrees of saturation. Finally, the variance of delay for an intersection is computed through an example.

5.1 Single Approach

The delay variance estimation process follows the flow chart in Figure 3 and every step in the chart is highlighted for easy readability. A hypothetical situation is assumed with the numerical values assigned as shown below. Consider a lane group in a pretimed signalized intersection with the following conditions

Average volume = 500 vph

Cycle length = 100 sec

Effective Green time = 30 sec

Saturation flow rate = 1800 vphgpl

The traffic volume and v/c ratio are assumed to follow Poisson distribution.

The analysis period was 15 minutes, i.e., T = 0.25

For a Pretimed signalized intersection K = 0.5

Also for an isolated intersection, I = 1

STEP I: Computation of average and variance of the degree of saturation (X)

Average and variance of X calculation requires capacity to be computed. The capacity of the approach can be found as follows:

Capacity (c) = $s \times \frac{g}{C} = 1800 \times \frac{30}{100} = 540 \ \ vphpl$

Average X = Average volume /Capacity

Average X is found from Average volume to capacity ratio, as 500 ÷ 540 = 0.92. Demand is assumed to follow a Poisson distribution therefore variance computations are redundant. Furthermore, with the variability in the volumes, the degree of saturation crosses one and the intersection will be oversaturated.

STEP II and STEP III

As is mentioned earlier, the integration process is coded into a C++ program and the Steps II and III from the flowchart in Figure 3 have been integrated into the code. Hence the final outcome of the program would be the delay average and variance. The C++ program is run and the average and variance of delay obtained from the program are 59.01 seconds/vehicle and 51.58 respectively. The variance value of 51.58 corresponds to a standard deviation of 7.18. An average delay of 57.12 seconds per vehicle is obtained from HCM delay Equation (3).

Further, assuming that the percentile value for the 95th percentile confidence intervals is 1.96, the 95th percentile confidence interval (59.01 ± 1.96 × 7.18) or (44.79, 73.08) is obtained using Equation (17).

The variance of delay increases with the degree of saturation (as observed with the undersaturated condition) and will be discussed in later sections. This implies that with increasing degree of saturation, the variance of delay would no longer have any significance as the 95th percentile confidence interval range gets to gigantic proportions. The variance of delay is studied with increasing X and is presented a few sections later. The variance value obtained in the example presented in this section is validated using Monte Carlo simulation as follows.

5.2 Evaluation of Integration Method

In order to verify the proposed integration method, Monte Carlo simulation is utilized. The simulation utilized here is similar to that of the undersaturated case. For the above example, 500 data points that follow a Poisson distribution with mean of 500 are generated. These data points can be considered as realizations of traffic volumes arriving over 500 cycles. As is mentioned earlier, it is assumed that the delay equation is valid. The HCM delay considers overflow conditions and computes delay for all the vehicles that arrive during the analysis period. Thus, delays for the 500 cycles are calculated using the HCM delay equation (Equation (3)). The mean delay and its variance are calculated from these 500 delay values and the histogram is plotted as in Figure 15. This process is repeated twenty five times and the distributions of these data sets are compared to the distribution generated from integration method. As in the example in the previous section, a mean value of 59.01 and variance of 51.58 is obtained. It is evident that the mean is almost equal to the variance of the delay from the result of the computations in section 5.1. Thus the delay distribution is assumed to be close to a Poisson distribution.

For the purpose of demonstration, a theoretical distribution, which follows a Poisson distribution with a mean of 59.01, is plotted on the twenty five histograms as shown in Figure 15.

As indicated earlier, Monte Carlo simulation generates inconsistent results. The Monte Carlo simulation could produce the mean and variance that are very close to actual values if extremely large number of random samplings were utilized. As shown in Figure 15, the integration method based distribution is well represented by distributions generated from Monte Carlo simulation.

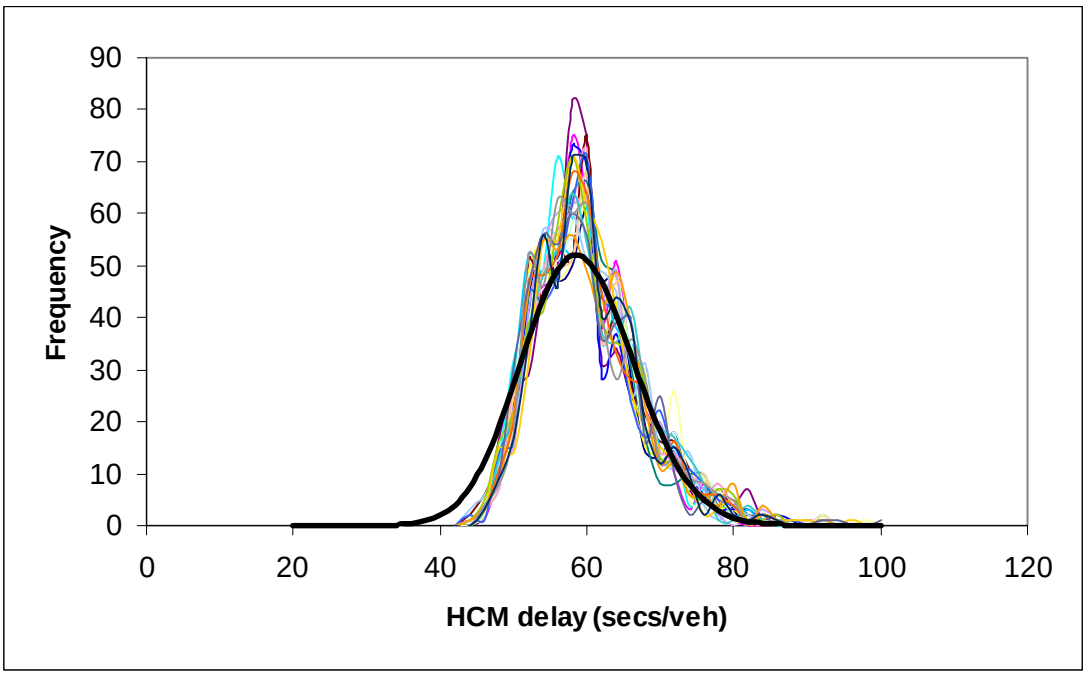

Figure 15. Expectation Function Method versus Monte Carlo Simulation

5.3 Evaluation of the Average and Variance of Delay under Different Demand Conditions

The variance of delay is computed for demand conditions ranging from 0.8 to 1.5 for the example problem used previously. The X range of 0.8 to 1.5 corresponds to a range of 432 to 810 for the volumes. Inputting different average volume values into the program provides the mean and variance (Repeating the procedure shown above using the steps I to III in Figure 3).

Figure 16 shows the effect of stochastic variability on Average Delay. As explained in chapter 5, the discrepancy between HCM delay and average delay arises when the HCM curve is not linear. Thus the two curves do not concur when the degree of saturation is around one. For very high X values, the delay curve is almost linear as it is designed to go asymptotically to a straight line. Therefore, the average delay and HCM delay coincides for very high X values.

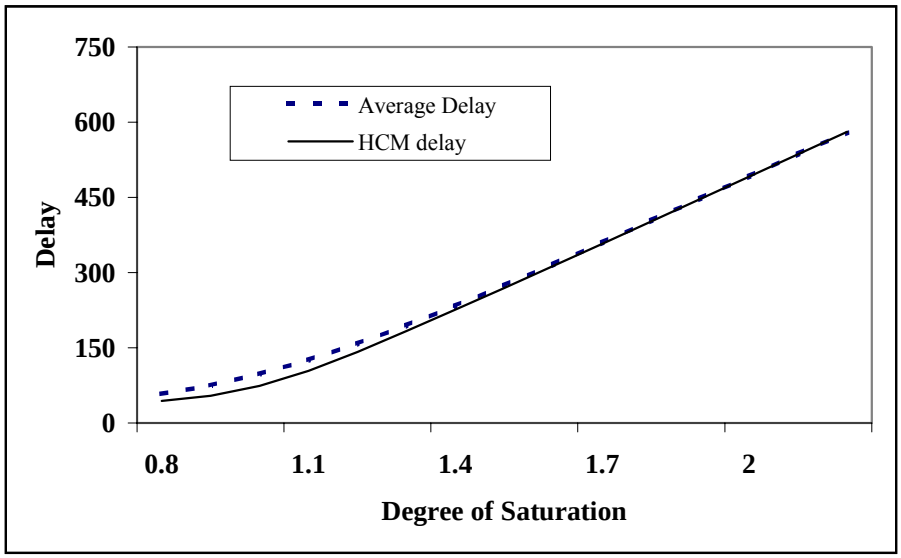

Figure 16. Average delay compared with HCM delay for over saturated conditions

As explained in the chapter 4 the HCM delay values and the average delay values diverge. The reasons for the divergence of the two curves provided in the Figure 16 are provided in the undersaturated intersection chapter.

Figure 17 shows the trend in the standard deviation of delay with increasing degree of saturation. As was observed for the undersaturated intersection, the delay standard deviation is observed to increase with increasing degree of saturation.

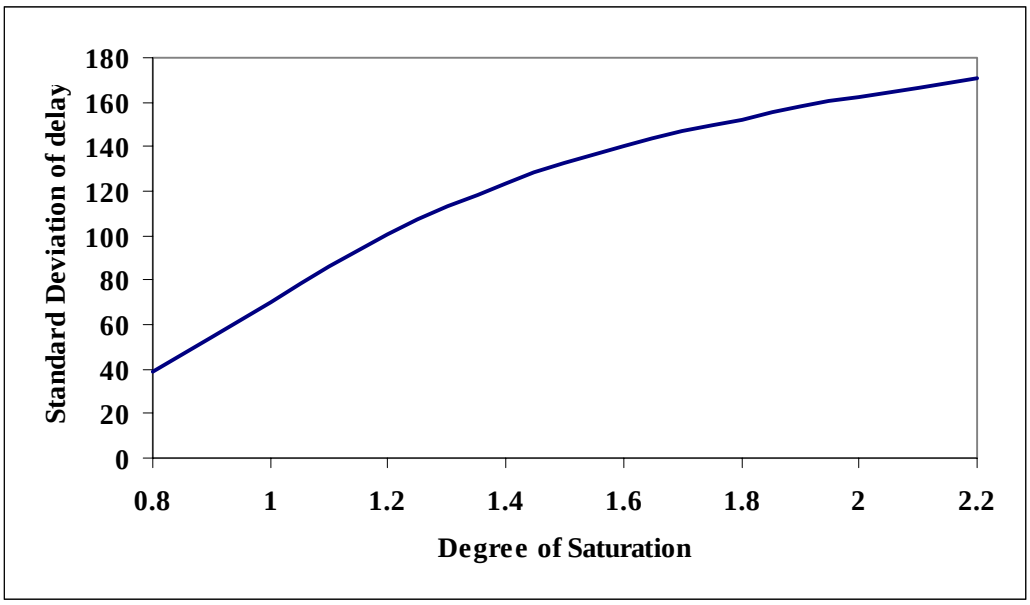

Figure 17. Standard Deviation of delay with degree of saturation

5.4 Intersection

The computations shown so far in this chapter are for a particular lane group in an approach. This methodology is expanded to compute the delay mean and variance for an intersection. The delay computation methodology shown so far is repeated for all the lane groups. Delay mean and variance are obtained separately for all the lane groups in an

intersection. The delay average and variance for an intersection are obtained by taking the weighted average of the delay average / variance values for every lane-group. The following equations further elicit the process.

Consider an intersection with n lane groups with the average delay and variance for a lane group i as μ_i and σ_i^2 respectively. The number of lane groups in an intersection depends on the lane assignment and the volumes in the lanes. Based on the lane group delay average and variance, the average delay μ_D for an intersection is calculated as follows

$$\mu_D = \frac{\sum_1^n V_i \times \mu_i}{\sum_1^n V_i} \tag{37}$$

and the variance of delay for the intersection is calculated as follows

$$\sigma_D^2 = \frac{\sum_1^n V^2{}_i \times \sigma_i^2}{\left(\sum_1^n V_i\right)^2} \tag{38}$$

Where,

V_i is the volume for the i^{th} lane group

μ_i is the average delay for the i^{th} lane group

σ_i^2 is the delay variance for the i^{th} lane group

5.4.1 Example

The equations presented above, Equations (37) – (38), are implemented using a hypothetical example. Consider an intersection with the following layout

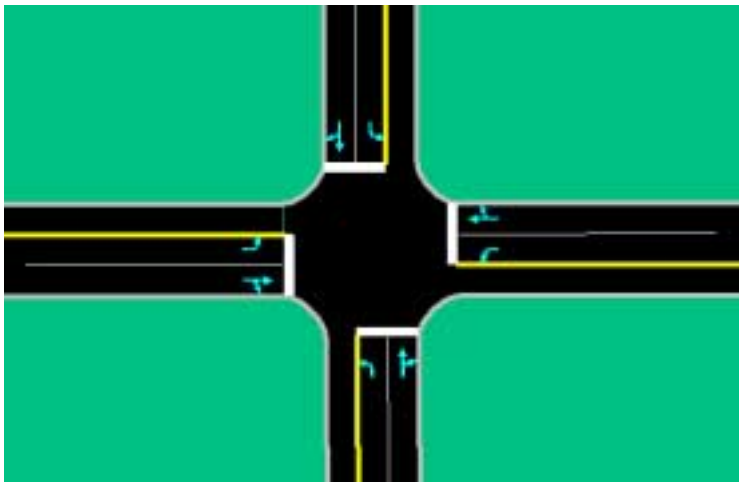

Figure 18. Example layout of an isolated intersection

Based on the guidelines provided by the HCM, eight lane groups could be identified. A left turning lane group and a through and right lane group for each approach results in a total of eight lane groups for the intersection. The volumes and green times assigned to each lane group are tabulated in Table 3.

Table 3. Example demand conditions for an isolated intersection

NEMA Phase	Average Volume (vph)	Distribution	Effective Green Time (secs)
1	100	Poisson	12
2	650	Poisson	50
3	100	Poisson	12
4	650	Poisson	50
5	100	Poisson	12
6	650	Poisson	50
7	100	Poisson	12
8	650	Poisson	50

The cycle length for the intersection is 140 seconds. The capacity for a through lane (even NEMA phase numbers) is (50/120)*1800 = 642 while the left turn lanes (odd NEMA phase numbers) have a capacity of (12/120)*1800 = 180. Therefore, the through lanes operate under oversaturated conditions while the left lanes are undersaturated. The delay average and variance is computed for all the eight lane groups using the methodologies presented previously. The delay values with their variances are highlighted in Table 4.

Table 4. Example Problem Delay Mean and Variance

NEMA Phase	Average Delay	Variance
1	81.72	27.6
2	84.05	113.01
3	81.72	27.6
4	84.05	113.01
5	81.72	27.6
6	84.05	113.01
7	81.72	27.6
8	84.05	113.01

The intersection average and variance are calculated as follows using Equations (37)-(38) as follows

$$\text{Average delay} = \left[\frac{81.7\times100+84.05\times650+81.7\times100+84.05\times650+81.7\times100+84.05\times650+81.7\times100+84.05\times650}{100+650+100+650+100+650+100+650}\right]$$

Average delay = 251218/3000 = 83.74 secs/veh

Similarly, intersection variance is computed as $192090900/3000^2 = 21.34$. Thus the intersection delay average and variance are obtained as 83.74 seconds/vehicle, 21.34 respectively.

CHAPTER 6. ARTERIAL INTERSECTIONS

The delay variance computations for an isolated intersection required inputs like the arrival mean, variance, and distribution. Depending on the intensity of arrivals and their distribution, the expectation value (for undersaturated intersection) or integration (for oversaturated intersection) methodology was utilized appropriately. This procedure could also be extended to the arterial intersection case by utilizing a factor called the progression factor for the through movements on the arterial links. This progression factor is multiplied to the uniform delay term in the HCM delay equation and it accounts for the progression of arrivals.

The arrivals (on the arterial link) at an arterial intersection do not follow any of the standard distributions and the distribution has to be estimated for the reason that the vehicular arrival at an upstream intersection is random (Poisson distribution) and these arrivals are filtered by the upstream signalized intersection (i.e., vehicles are queued up during the red period to be discharged during the green time only). Thus, the volume mean, variance, and distribution are assumed to be input by the user and these volumes are used in the delay computations along with the calculated progression factor.

The arrival pattern is accounted for by using the progression factor and is estimated using the platoon dispersion model. This procedure is presented in the next section using a hypothetical example. This example is followed by the delay mean and variance computations for the same example major street through movement.

6.1 Estimation of Arrival Pattern and Progression Factor

The following example is utilized to explain the progression factor estimation methodology. Consider an arterial with two intersections as shown in Figure 19 with the following conditions.

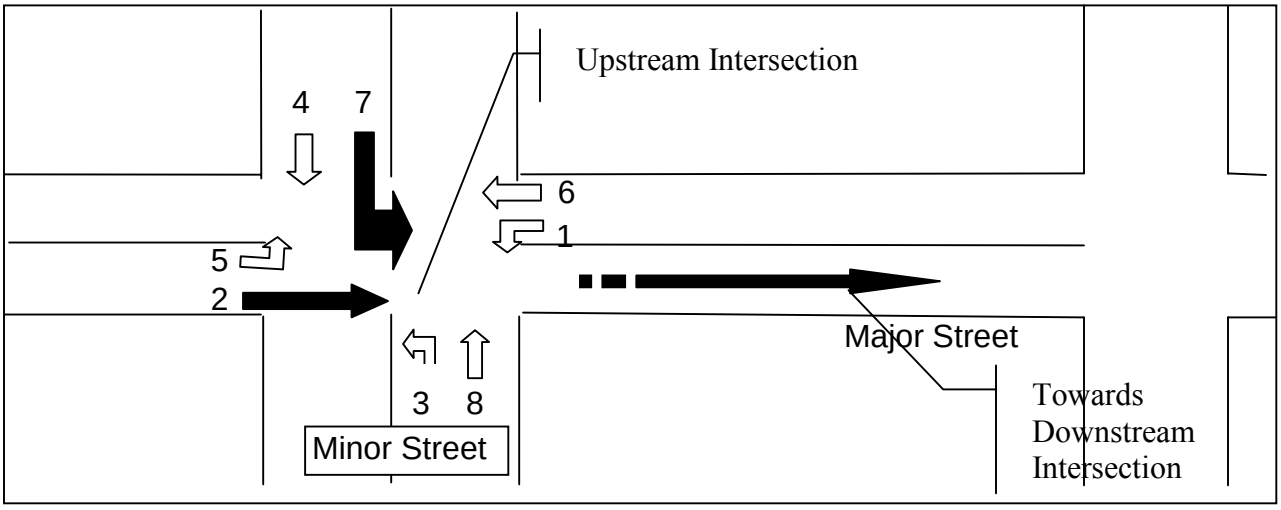

Figure 19. Example situation for arterial intersections

- The spacing between the intersections is such that the travel time is 20 secs from one intersection to the other
- There are no right turning vehicles during the red period

The example situation is explained through Figure 19. The movements at the upstream intersection are numbered according to NEMA standard phasing. The movements in dark (NEMA phases 2 and 7) contribute to the volumes onto the arterial (Movement in dark on the major street) that undergoes platoon dispersion. The progression factor is estimated for this movement based on the mean volumes, green times for movements 2 and 7.

Inputs for the platoon dispersion model

T is the cruise time on the link in steps = 20 secs

β is the empirical factor = 0.8

$F = (1 + \alpha\beta T)^{-1}$ is the smoothing factor = 0.125

α is the platoon dispersion factor = 0.35

Upstream Volume and signal timing Inputs

The arrival distribution at the upstream intersection is Poisson

Mean of the through arrivals (NEMA 2) = 400 vph

Mean of the left turning arrivals (NEMA 7) = 100 vph

Green time for the through movement (NEMA 2) = 50 secs

Green time for the left turning movement (NEMA 7)= 25 secs

Cycle length C = 140 secs

Saturation flow rate = 1800vph

Downstream signal timing Inputs

Green time downstream for major street through = 50 secs

Offset (assuming perfect) = 20 secs

Firstly, the downstream arrival pattern has to be estimated using Equations (23) – (26). The arrival pattern for the left movement (upstream NEMA 7) and the through movement (upstream NEMA 2) are computed separately and aggregated according to the time the vehicles arrive. The g_0 for each movement has to be calculated using the equation $g_0 = \frac{r \times v}{(s - v)}$

where,

r is the duration of red for the movement

v is the average flow rate over the cycle

s is the saturation flow rate

g_0 for through movement = 90×400/(1800-400) = 25.7 secs ~ 26 secs

g_0 for left movement = 115×100/(1800-100) = 6.7 secs ~ 7 secs

These values are substituted into the equations along with the cycle length, green time, smoothing factor (F), average flow rate and saturation flow rate.

The arrival pattern downstream for the through volume turns out to be

$$Q(t) = \begin{cases} 1800 - 1800 \times (0.875)^{t+1} + 462.63(0.875)^{t+90} & t \in [1,26] \\ 400 + 1400(0.875)^{t-27} - 1800(0.875)^{t} + 462.63(0.875)^{t+90} & t \in [27,49] \\ 462.63(0.875)^{t-50} + 462.63(0.875)^{t+90} & t \in [50,140] \end{cases}$$

The arrival pattern downstream for the left volumes turns out to be

$$Q(t) = \begin{cases} 1800 - 1800 \times (0.875)^{t+1} + 189.7(0.875)^{t+115} & t \in [1,7] \\ 100 + 1700(0.875)^{t-7} - 1800(0.875)^{t} + 189.7(0.875)^{t+115} & t \in [8,24] \\ 189.7(0.875)^{t-25} + 189.7(0.875)^{t+115} & t \in [25,140] \end{cases}$$

The t used in the above equations is with reference to begin of green for that particular movement. That is t = 1 is the beginning of the green for the corresponding movement. These time scales have to be transformed with reference to the downstream intersection. The offset is the time difference between the beginning of green for the upstream intersection through movement (NEMA 2) to the beginning of green for the downstream intersection through movement. In this example, it is equal to the travel time on the link, therefore, the through vehicles arrive downstream exactly at the beginning of green. Therefore, t = 1 for the through volume is the beginning of the downstream green.

Similarly, the left volume starts departing after a time equal to 50 secs, (green time for the through movement). After the beginning of green for the upstream through phase, the first vehicle (from NEMA 7) arrives downstream after a time 50secs after the begin of green at the downstream intersection. However, the downstream green time = 50 secs therefore, these vehicles arrive at the end of green. Therefore, for 90 secs (red period), the vehicles are queued up and at t = 90 the green period begins.

Thus the arrivals during green constitute the sum of volumes from t = 1 to t = 50 for the through movement and sum of t= 90 to t = 140 for the left movement (since the downstream green is 50 secs). Adding up the corresponding volumes and dividing by the green time to get the flow rate during the green time, results in an average flow rate of 1121.57 vph during the green period.

Based on the flow rate during the green period, using the Equation (23), results in a progression factor of 0.268.

6.2 Delay Variance Computations Using the Progression Factor

The same example as in the previous section is utilized here. Additional information on the arrival mean, variance, and distribution are necessary for delay variance computations.

Downstream volume inputs

Average through volume = 400 vph

Standard Deviation of through volume = 50

Further, the demand distribution is assumed to be Normal.

The major street through movement arrives downstream with a progression factor of 0.268. Equation (22) is the HCM equation to compute delay for movements in progression. Utilizing the progression factor of 0.268 and the information provided, the C++ code is run and the delay average and variance values of 14.5 and 0.534 respectively are obtained.

CHAPTER 7 OPTIMIZATION

The first objective of this research, to quantify the variability in delay due to variable demand conditions, was achieved using appropriate methodologies for different demand conditions as presented in Chapter 3. The next objective of this research, to optimize a signalized intersection subject to stochastic variability, will be realized through the examples in the present chapter. Genetic algorithms (GA) are utilized in the optimization process for the present study. The GA process commences with the generation of random input values using binary strings. These strings are acted upon by a few operators, leading to fitter individuals, every generation. The best solution in the final generation is chosen as the optimal solution. All the above steps leading to the generation of individuals has been coded into a C++ program. The program inputs the average and variance of all the demand volumes per lane group and the other necessary parameters and outputs the signal timing plan for the system. The signal timing plan should result in a minimum average delay or a minimum 95^{th} percentile delay depending on the objective function used.

Since stochastic optimization of signalized intersections is undertaken, the signal timing optimization strategy has to be tested under different average demand conditions. Scenarios of average demand conditions ranging from slightly moderate to heavy volume conditions are optimized and evaluated using microscopic simulation programs. The three scenarios namely, moderate traffic at isolated intersections, heavy traffic at isolated intersection and arterial intersections are presented separately.

The GA program is run for each scenario and the obtained signal timing plan is then compared with SYNCHRO timing plan for each scenario using neutral microscopic simulators like CORSIM or SIMTRAFFIC. The moderate volume condition is evaluated using SYNCHRO as an evaluator and then CORSIM is utilized. The heavy demand conditions are evaluated using SIMTRAFFIC as an evaluator. The arterial intersection is evaluated using CORSIM. The following sections present these evaluation procedures under various demand conditions.

7.1 Scenario I: Moderate Traffic

The GA optimization process is conducted on a hypothetical isolated intersection with moderate traffic conditions. The following are the conditions assumed to exist at the intersection

7.1.1 Setting

An isolated intersection with normal variations in the demand conditions is considered. Consider an isolated intersection with a layout as shown in the Figure 20. Based on the lane assignments, eight lane groups are identified, two for each of the four approaches. The two lane groups in an approach are left only lane group and a through and right movement lane group. As is mentioned earlier, the C++ program requires the volume mean and variance to be input by lane groups. The following Table 5 provides the average and variance of all the lane groups in the intersection. The intersection is designed such that the EB / WB approach volumes have a higher variability compared to the NB / SB approaches.

Figure 20. Layout of the hypothetical intersection

Table 5. Demand conditions for Scenario I

Nema Phase	Average Volume	Standard Deviation	Distribution
1	100	30	Normal
2	400	120	Normal
3	100	10	Normal
4	400	40	Normal
5	100	30	Normal
6	400	120	Normal
7	100	10	Normal
8	400	40	Normal

7.1.2 Timing Plan development

The GA code in C++ is run with the above input conditions. The GA optimization process has been explained earlier. Various GA optimization parameters have to be specified by the user to optimize the system more efficiently. Forty individuals are generated in every generation. A tournament selection process is utilized to select individuals to the next generation. A single point crossover is employed and a mutation probability of 0.05 is selected. The GA process is terminated after a 100 generations and the best individual in the final generation is selected as the optimal solution.

Two objective functions are considered in the optimization process namely, 95th percentile delay and the average delay the GA code is run using these objective functions separately. The results of the optimization are provided in the table below.

Table 6. Comparison of GA and SYNCHRO green times

NEMA Phases	GA (Avg Delay) optimized phase times (secs)	GA (95th %) optimized phase times (secs)	SYNCHRO optimized green times (secs)
1	10	11	10
2	28	36	20
3	10	12	10
4	27	31	20
5	10	11	10
6	28	36	20
7	10	12	10
8	27	31	20

Similarly, the average demand conditions shown in Table 5 are input into the SYNCHRO program and the intersection is optimized for Percentile delay that SYNCHRO computes. The result of the optimization is presented in the Table 66 above.

7.1.3 Evaluation

The timing plans obtained in the previous section have to be evaluated with different demand conditions as the GA is optimized for stochastic variability and considers for the demand fluctuations. To aid in generating different demand conditions, a Latin Hypercube design is utilized.

7.1.3.1 Experimental Design

As is mentioned earlier, Latin Hypercube Sampling procedure is utilized to generate different demand conditions from the combinations of volumes generated for different movements. The details of the design are presented in the methodology chapter. The varying demand volume conditions involving Latin Hypercube design were generated from ACED (Algorithms for the Construction of Experimental Designs) program. ACED is UNIX based software, which can interface with all implemented design optimization algorithms. Using the ACED program, a hundred different combinations of volume conditions are generated based on the mean and variance values specified in the Table 5. Thus 100 demand conditions for an isolated intersection are generated.

7.1.3.2 SYNCHRO Evaluation

In the present section, the delays from the GA average delay optimized timing plan are compared to the delay from the SYNCHRO percentile delay optimized timing plan using SYNCHRO as an evaluator. The optimized timing plans are presented in the section 7.1.2. As was mentioned previously, 100 different demand conditions are generated using Latin Hypercube sampling procedure. These 100 demand conditions are input into

SYNCHRO with the timing plans from GA and SYNCHRO as shown in Table 66. That is, the 100 demand conditions are input into SYNCHRO using the SYNCHRO optimized timing plan and the GA optimized timing plan and, the percentile delay and the HCM control delay from SYNCHRO are compared. The 100 demand conditions thus result in 100 percentile delay and 100 HCM delay values for both the timing plans. Scatter plots are made with the HCM delay and Percentile delay values obtained from SYNCHRO. Figure 21 and 22 show the X-Y plot of the delay results.

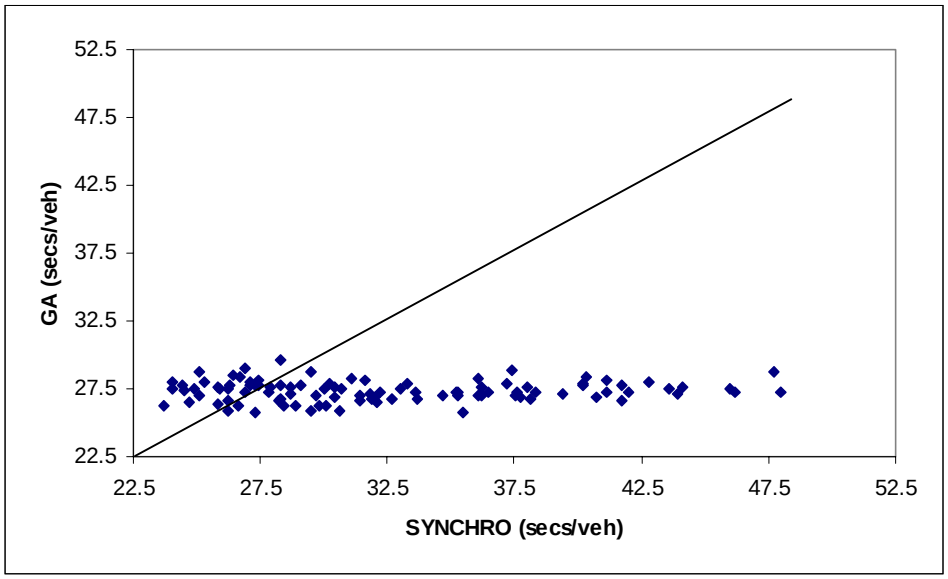

Figure 21. Comparison of SYNCHRO and GA (Average) optimized timing plan using SYNCHRO – HCM delays

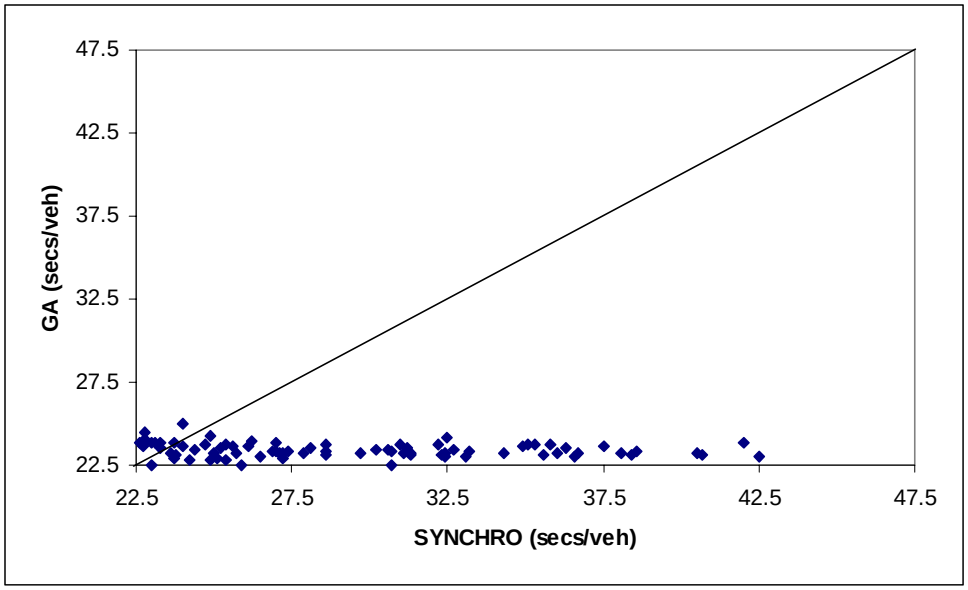

Figure 22. Comparison of SYNCHRO and GA (Average) optimized timing plan using SYNCHRO – Percentile delays

7.1.3.2.1 Discussion

The plots for both the percentile delay and the HCM control delay show that the GA optimized timing plan performs better than the SYNCHRO optimized timing plan. The 45-degree line in the Figure 22 is drawn as a visual aid in identifying the better timing plan. Based on the SYNCHRO evaluations, GA performance is better than that of SYNCHRO. The reason for this could be attributed to the stochastic optimization of delay. For extremely high volume conditions, the SYNCHRO timing plan resulted in high delays whereas the GA timing plan was not affected by huge fluctuations in demand.

7.1.3.3 CORSIM Evaluation

The different demand conditions have been generated based on the Latin Hypercube sampling procedure mentioned in the section 7.1.3.1. The Hundred different demand

conditions are simulated in CORSIM using the GA and SYNCHRO optimized timing plans. Multiple runs are made in CORSIM, for every demand condition, to account for the stochastic variations. Thus the CORSIM simulations are repeated for five times for every demand combination and the average delay from the five runs is utilized for comparison. Since 100 different combinations of volumes are generated, 100 average delay values are obtained from each of the SYNCHRO, GA average and GA 95th percentile optimized cases.

7.1.3.3.1 Evaluation criterion

The queue time from CORSIM output is used as an evaluation criterion. Queue time is used in the evaluation procedure because, the link delays in CORSIM are reported for vehicles that have been discharged through the link during the simulation period, and HCM evaluates the delay for all the vehicles present in the system during the analysis period. On the other hand, queue delay represents the cumulative measure for all the queuing in the network and includes the delay experienced by vehicles that remain in the network, or those that have yet to cross a link at the end of the simulation period. Hence, queue time is used in the evaluation procedure.

7.1.3.3.2 Discussion

The queue times obtained from the CORSIM simulation using the SYNCHRO percentile delay optimized timing plan and GA average and 95th percentile delay optimized timing plans are compared in the present section. 100 queue time values are obtained for the three timing plans separately. X-Y plots of these results are presented in the Figure 23

and Figure 24. Firstly, the queue times from the SYNCHRO optimized conditions are plotted against the queue times from the GA average optimized case. The plot in Figure 23 shows that the GA average delay optimization is doing better than the SYNCHRO optimization. Similarly, the SYNCHRO timing plan is compared with the GA 95th percentile optimized timing plan. The plot in Figure 24 shows that the GA timing plan performs better than the SYNCHRO timing plan. The reasons for this performance could be attributed to the stochastic optimization undertaken by the GA program unlike SYNCHRO, which used average delay for optimization. High volumes resulted in very high queue times for the SYNCHRO case. However, for low volume conditions, the GA timing plan is out performed by the SYNCHRO timing plan. The reason for this could be the high green time and cycle lengths resulting from the GA optimization resulting in a large red period for the movements.

However, to measure the overall performance of the timing plans, a statistical test is conducted. The paired T-test is utilized to test if the queue time values from different timing plans are statistically different or not. Table 7 presents the results of the t-test. The test is conducted with the null hypothesis that the mean queue time obtained from the two timing plans is same. However, for the SYNCHRO – GA average delay comparison, a p value of 0.00 is obtained. Thus, null-hypothesis is rejected. This indicates that the mean queue time from the GA timing plans is considerably smaller than that from the SYNCHRO timing plan and the standard deviation value for the GA case is lower indicating superior performance when compared to SYNCHRO. Similarly, the GA 95th percentile optimized timing plan is compared with the SYNCHRO timing plan. Although

the GA 95th percentile timing plan showed better performance, it performed worse than the GA average optimized timing plan.

Table 7. Comparisons of SYNCHRO and GA timing plans, T-Test result

	SYNCHRO	GA-Average
Mean	1768.2	1127.4
Standard Deviation	1762.3	742.1
Observations	100	100
P(T<=t) two tail	0	
t critical two-tail	1.984	
	SYNCHRO	GA-95th Percentile
Mean	1768.2	1384
Standard Deviation	1762.3	1113.8
Observations	100	100
P(T<=t) two tail	0	
t critical two-tail	1.984	

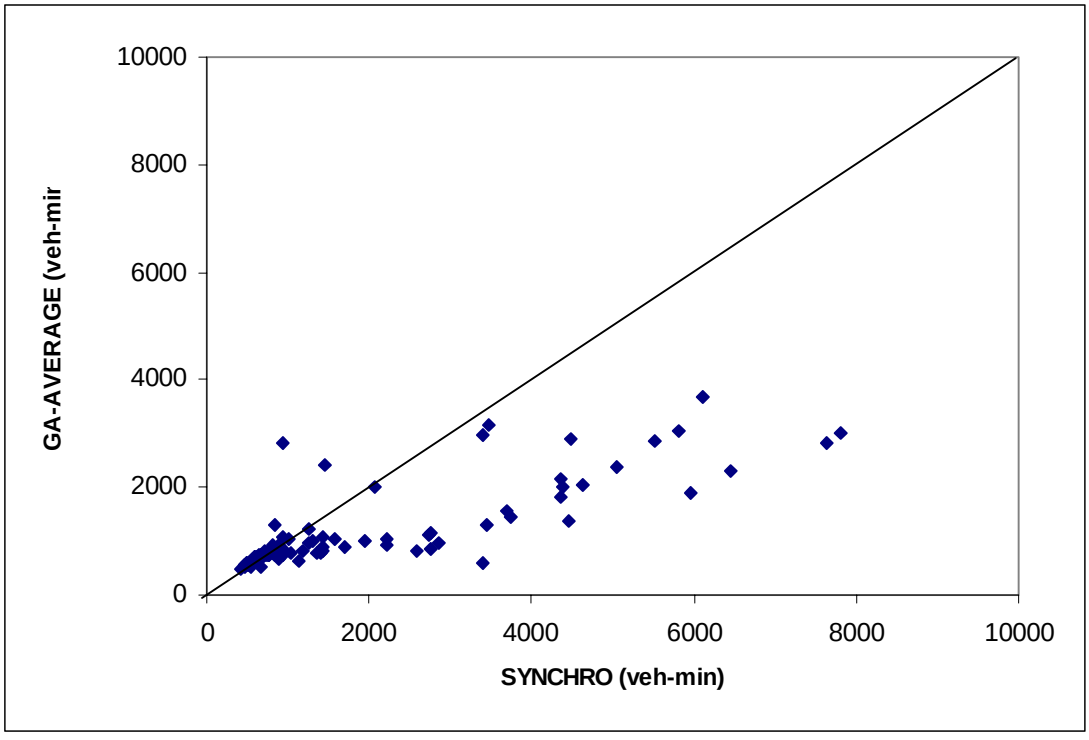

Figure 23. SYNCHRO Timing Plan compared with the GA Average optimized timing plans in CORSIM

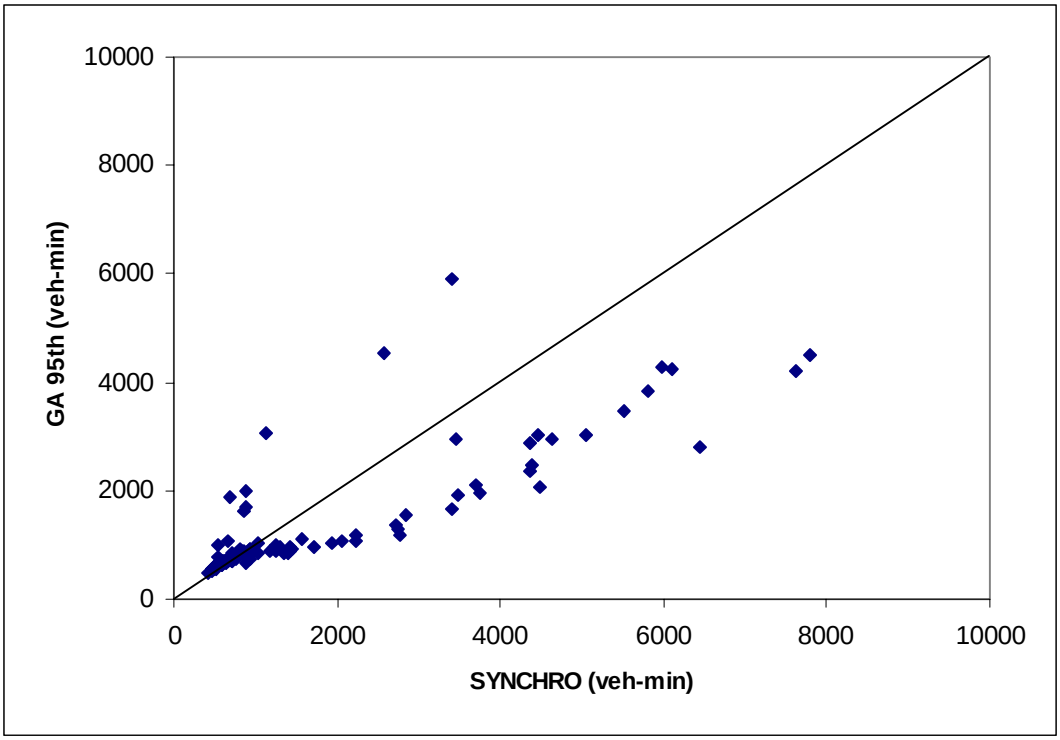

Figure 24. SYNCHRO Timing Plan Compared with the GA 95th Percentile optimized timing plans in CORSIM

7.2 Scenario II: Heavy Traffic

The GA optimization procedure is evaluated for heavy traffic demand conditions. The GA procedure is compared with SYNCHRO and evaluated using SIMTRAFFIC and CORSIM.

7.2.1 Setting

The GA optimization is conducted on a hypothetical isolated intersection. An isolated intersection with a layout as shown in the Figure 20 is considered. Based on the lane assignments, eight lane groups are identified, two for each of the four approaches. The two lane groups in an approach are a left only lane group and a through and right movement lane group. The average volumes presented in the Table 8 correspond to a degree of saturation of close to unity. Thus, the demand conditions are tagged as heavy traffic situation.

Table 8. Example inputs for Optimization of an isolated intersection

Nema Phasing	Average Volumes (vph)	Distribution
1	100	Poisson
2	650	Poisson
3	100	Poisson
4	650	Poisson
5	100	Poisson
6	650	Poisson
7	100	Poisson
8	650	Poisson

7.2.2 Timing plan development

As is mentioned earlier, the C++ program requires the volume mean and variance to be input by lane groups. Table 8 provides the average and variance of all the lane groups in the intersection. The distribution of the demand volumes are also identified as Poisson for all the lane groups hence variance values are not provided in the table. With all the input parameters at hand, the objective function has to be identified to proceed with the optimization process.

7.2.2.1 GA Average Delay

To start off, average delay is chosen as the objective function. The average delay utilized here is different from the HCM delay as it incorporates the variability of the demand volumes into delay. All the input conditions are input into the C++ program and run, a cycle length of 140 seconds is obtained from the GA process. This cycle length resulted in an average delay of 92.78 seconds/vehicle.

The GA optimization process has been explained earlier. Various GA optimization parameters have been used for this scenario and are similar to the ones used for Scenario I. For every generation, the minimum fitness value and the average fitness value of the population are plotted against the generation number. This plot gives an idea of the GA convergence to the optimal solution. Such a plot is presented in the Figure 25 for the present example.

7.2.2.2 GA 95th Percentile Delay

The hypothetical example is optimized using the 95th percentile delay as the objective function. All the other parameters are fixed including the GA operational parameters. The 95th percentile value is calculated using the Equation (17) with a percentile value of 1.96. The GA program resulted in a cycle length of 148 seconds as the optimal solution. The green times assigned to different lane groups are presented in table. The convergence of the GA to optimal solution is studied by plotting the minimum fitness value and average fitness value for every generation against the generation number. One such curve is shown for the 95th percentile optimization condition in Figure 26.

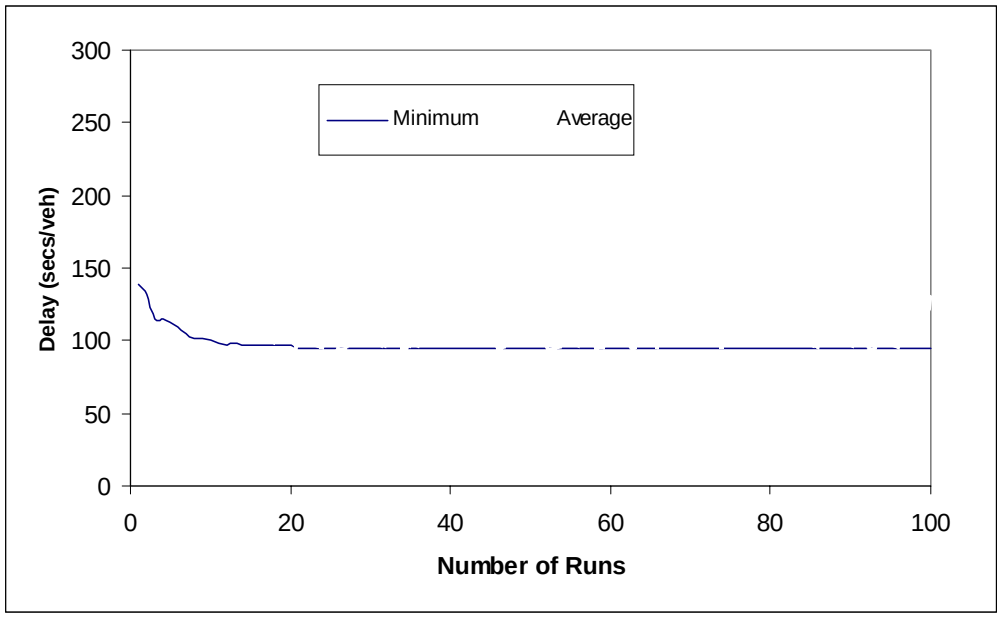

Figure 25. GA convergence using the average delay for optimization

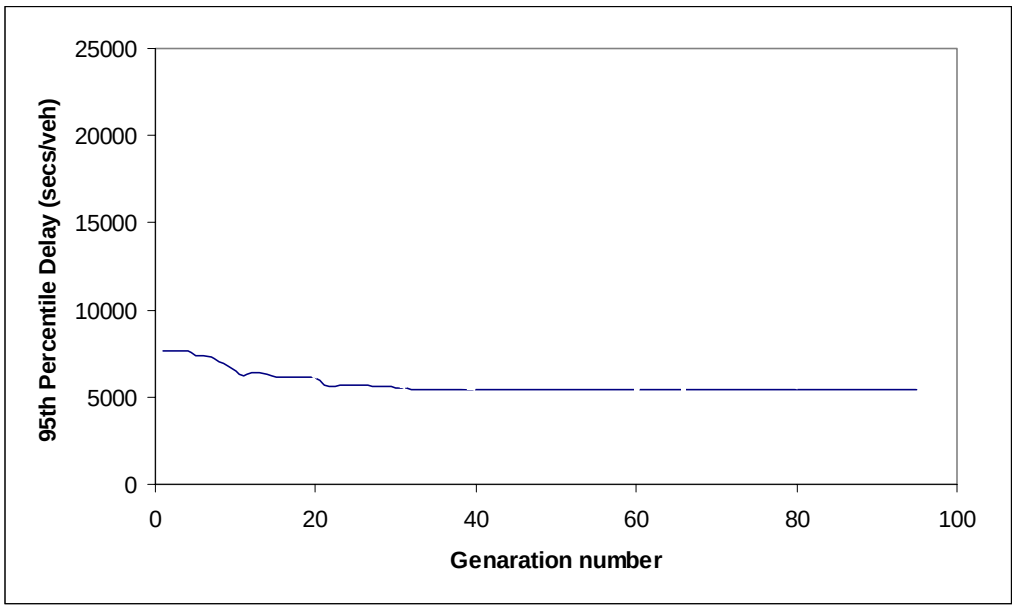

Figure 26. Convergence of GA algorithm to the optimal solution

7.2.2.3 SYNCHRO

The average volumes from the hypothetical example are input into SYNCHRO and optimized. SYNCHRO resulted in a cycle length of 100 secs. Since SYNCHRO utilizes percentile delay for its optimization process, the 100 secs cycle length resulted in an optimal percentile delay of 66.66 secs/veh. The following Table 9 shows the signal timing plan obtained from GA and SYNCHRO for the hypothetical example. The timing plans obtained from SYNCHRO and GA are then evaluated using a microscopic simulation program SIMTRAFFIC. The evaluation procedure is presented in the next section.

Table 9. Various timing plans for scenario II

Nema Phase	GA average delay optimized phase times	GA 95th % delay optimized phase times	SYNCHRO Percentile delay optimized phase times
1	12	12	10
2	59	60	40
3	15	12	10
4	57	56	40
5	16	12	10
6	55	60	40
7	15	12	10
8	57	56	40

7.2.3 SIMTRAFFIC Evaluation and Discussion

The result of the optimization from GA and SYNCHRO are evaluated using a microscopic simulator SIMTRAFFIC. Since intersections are subject to random arrivals during the course of its operation, random volumes are generated for 225 mins in SIMTRAFFIC and the signal timing plans are evaluated.

SIMTRAFFIC has the option of setting different percentile values for different time intervals during the simulation process i.e. if the simulation is being run for 30 min, the first 15 min volume could be set to a percentile of arrivals and the next 15min volume to a different percentile value. These percentile values have to be defined for the duration of the simulation prior to the evaluation. Usually, the first 15 min period is set for initializing the system while the last 15 min period is set to a zero demand condition to flush out all the vehicles from the system.

SIMTRAFFIC was run with both the timing plans three times, every time with a different random number seed. The delay values reported by SIMTRAFFIC for all the cases are plotted against the simulation time. The Figure 27 below shows the result of the SIMTRAFFIC evaluation for the present case. The vertical bars in the background of the Figure 27 indicate the random percentile value used.

From the Figure 27 it is evident that the GA optimized timing plan worked better than the Synchro optimized timing plan. Initially, when the system was subject to low percentile volumes (shown in Figure 27), GA and SYNCHRO timing plans resulted in similar delay values. However, high percentile volumes disrupted the SYNCHRO timing plans and resulted in very high delays compared to GA. It is quite clear from Figure 27 that after the second high 95th percentile volume, SYNCHRO consistently resulted in high delay values. The GA timing plan resulted in consistent performance with a minimal increase in delay (compared to SYNCHRO) even after the high percentile arrivals. The reasons for GAs' consistent performance could be attributed to the consideration given to stochastic variability in delay optimization. Since the GA program accounted for the variability, very high fluctuations in the volumes did not generate high delay values.

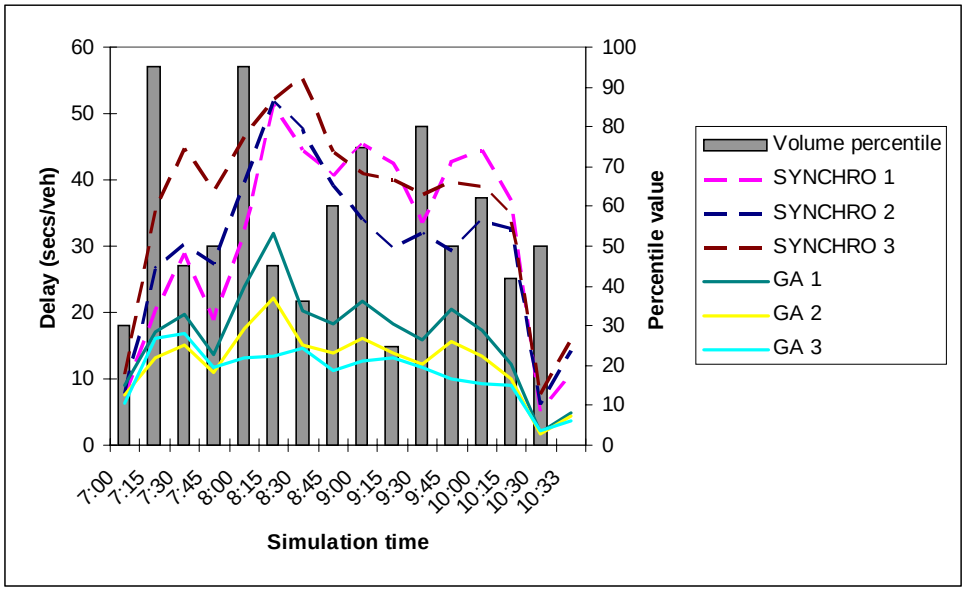

Figure 27. Comparison between SYNCHRO and GA

The result of the optimization with the 95th percentile delay and the average delay are compared using SIMTRAFFIC. Random volumes are generated for 225 min and the

simulation is run with the signal timing plans from both the cases. The simulation is run for three times for each signal timing plan using different random seeds each times. The delay values obtained are plotted against time as shown in Figure 28. It is observed that the 95th percentile optimized timing plan produced delay values slightly higher than that for the average delay optimized case. The reason for the higher delay estimate could be due to the 95th percentile value used for optimization. The optimal green times attributed to the movements are very high and hence for medium to low volumes, the delay values are higher than that for the average volume case.

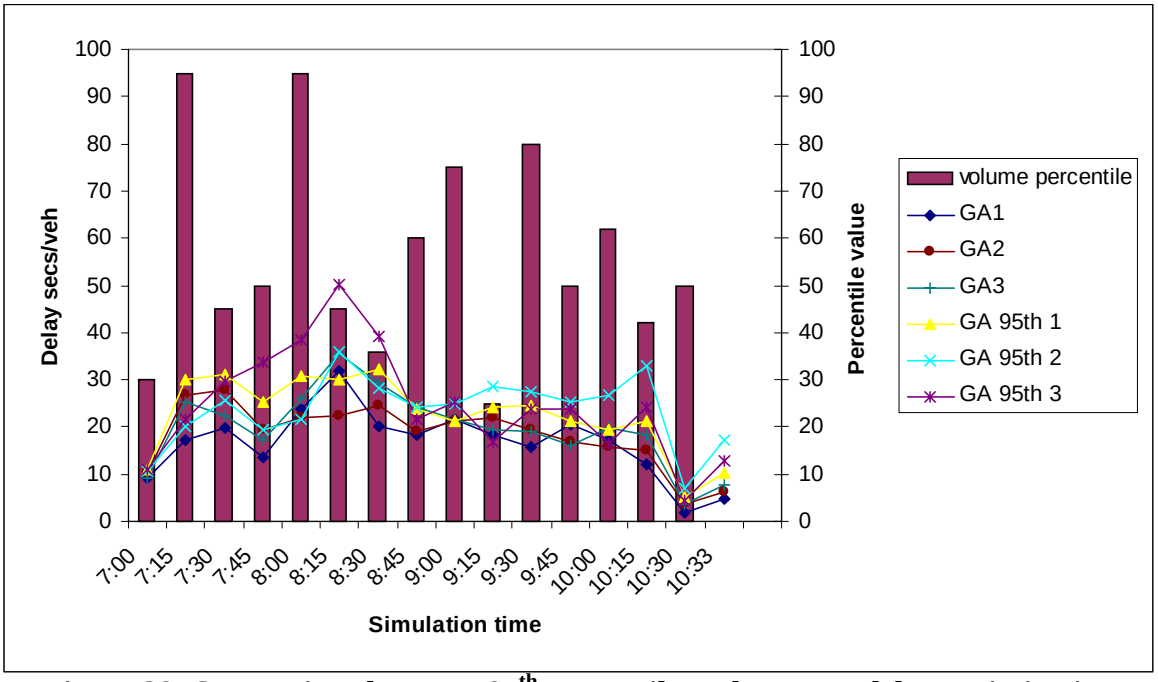

Figure 28. Comparison between 95th percentile and average delay optimization

7.2.4 CORSIM Evaluation

The different demand conditions are generated based on the Latin Hypercube sampling as shown for the Medium demand conditions. The Hundred different demand conditions are simulated in CORSIM using the GA and SYNCHRO optimized timing plans. Multiple

runs are made in CORSIM, for every demand condition, to account for the stochastic variations. Thus the CORSIM simulations are repeated for five times for every demand combination and the average delay from the five runs is utilized for comparison. Since 100 different combinations of volumes are generated, 100 average delay values are obtained from each of the SYNCHRO, GA average and GA 95th percentile optimized cases. Similar to the medium demand condition, queue time is used in the evaluation criterion.

7.2.4.1 Discussion

The queue times obtained from the CORSIM simulation using the SYNCHRO percentile delay optimized timing plan and GA average and 95th percentile delay optimized timing plans are compared in the present section. A methodology used for the medium demand condition is used here. The results of the evaluation are presented through Table 10 and Figures 29 and 30.

X-Y plots of these results are presented in the Figures 29 and 30. The plot in Figure 29 shows that the GA average delay optimization is doing better than the SYNCHRO optimization. Similarly, the SYNCHRO timing plan is compared with the GA 95th percentile optimized timing plan. The plot in Figure 30 shows that the GA timing plan performs better than the SYNCHRO timing plan.

The paired T-test is utilized to test if the queue time values from different timing plans are statistically different or not. Table 10 presents the results of the t-test. The test is conducted with the null hypothesis that the mean queue time obtained from the two timing plans is same. However, for the SYNCHRO – GA average delay comparison, a p

value of 0.02 is obtained. And hence the null hypothesis is rejected. Also, the mean queue time from the GA timing plans is considerably smaller than that from the SYNCHRO timing plan and the standard deviation value for the GA case is lower indicating superior performance when compared to SYNCHRO. Similarly, the GA 95th percentile optimized timing plan is compared with the SYNCHRO timing plan. Although the GA 95th percentile timing plan showed better performance, it performed worse than the GA average optimized timing plan.

Table 10. Comparisons of SYNCHRO and GA timing plans, T-Test result

	GA Average	*SYNCHRO*
Mean	**5045**	**5484**
Standard Deviation	**3180**	**3588**
P(T<=t) one-tail	**0.020**	
t Critical one-tail	**1.660**	
	GA 95th Percentile	*SYNCHRO*
Mean	**5215**	**5484**
Standard Deviation	**3373**	**3588**
P(T<=t) one-tail	**0.071**	
t Critical one-tail	**1.660**	

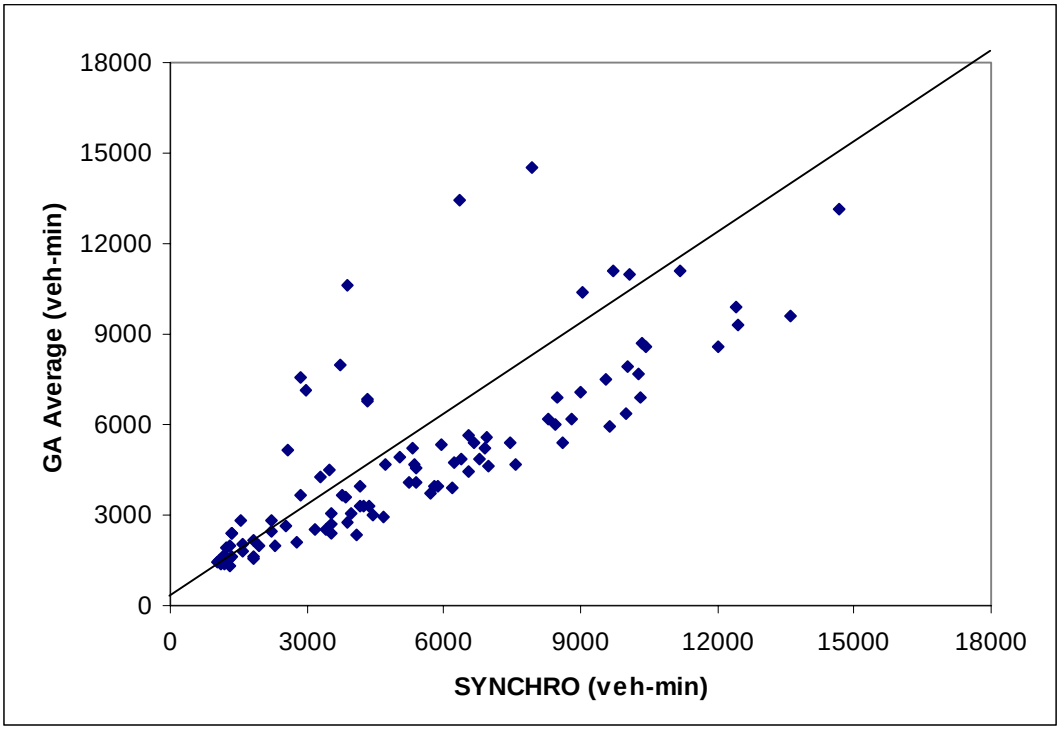

Figure 29. SYNCHRO Timing Plan compared with the GA Average optimized timing plans in CORSIM

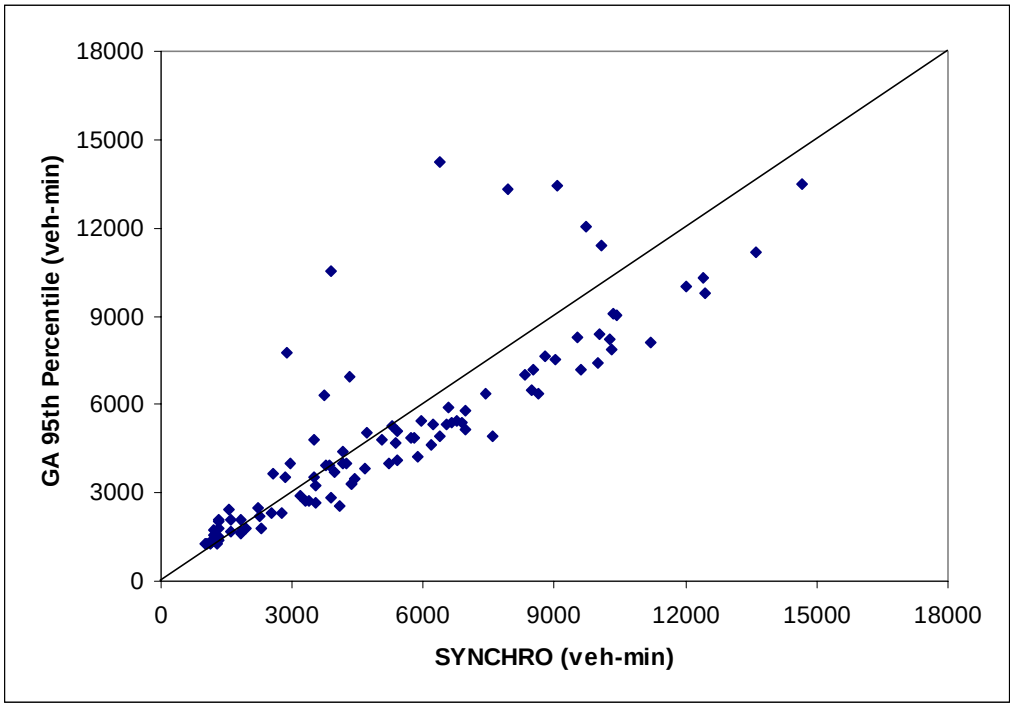

Figure 30. SYNCHRO Timing Plan Compared with the GA 95th Percentile optimized timing plans in CORSIM

7.3 Scenario III: Arterial Intersection

The signal timing plans from the optimization of an isolated intersection have been evaluated so far. In the present section, an arterial intersection is optimized using the methodology presented in the Chapter 3 and the result of the optimization is compared with SYNCHRO optimized timing plan and evaluated using CORSIM.

7.3.1 Setting

A hypothetical example is considered for the evaluation procedure. The following are the demand volume conditions existing at the intersection.

Table 11. Input demand conditions for the hypothetical intersection

Intersection	Nema Phase	Average Volume	Standard Deviation of Volume
	1	100	30
	2	400	120
	3	100	10
Downstream	4	400	40
	5	100	30
	6	400	120
	7	100	10
	8	400	40
	1	100	30
	2	400	120
	3	100	10
Upstream	4	400	40
	5	100	30
	6	400	120
	7	100	10
	8	400	40

The volumes for all the through movements are assumed to be 400 while that for the left turning movements 100 vehicles per hour. The standard deviations of the volumes are designed such that the variability on the major street or the arterial is higher than that of the other link volumes.

7.3.2 Timing Plan Development

The GA program is run with the input conditions presented in the Table 11 and the system is optimized for the average delay and the 95th percentile delay. Similarly, the average demand conditions are input into SYNCHRO and the network is optimized for the percentile delay that SYNCHRO computes. The optimized phase times for the three cases are presented in the Table 2 using the standard NEMA phase numbering. The cycle lengths and the offset vales obtained are also presented in the table. The offset value is the time difference between begin of green on the major street through movements (2/6 NEMA phase).

Table 12. Result of the optimization for the hypothetical arterial

Intersection	Nema Phase	SYNCHRO optimized phase times (secs)	GA Average delay optimized phase times(secs)	GA 95th Percentile optimized phase times(secs)
	1	10	12	12
	2	26	27	36
	3	10	10	11
Downstream	4	24	26	31
	5	10	12	12
	6	26	27	36
	7	10	10	11
	8	24	26	31
	1	10	10	12
	2	26	27	36
	3	10	11	11
Upstream	4	24	27	31
	5	10	10	12
	6	26	27	36
	7	10	11	11
	8	24	27	31
	CYCLE (secs)	70	75	90
	OFFSET (secs)	35	32	38

7.3.3 Evaluation and Discussion

The signal timing plans from GA are compared with that of SYNCHRO using the neutral microscopic simulation program CORSIM in a procedure similar to that of an isolated intersection. Firstly, demand conditions are generated using the Latin Hypercube design

and then CORSIM simulation is run with the timing plans from GA and SYNCHRO. The queue times from the CORSIM simulation output are used in the comparison. The queue time reports the delay experienced by all the vehicles in the network even if the vehicle is not discharged during the period of analysis and further, this definition is consistent with that of the HCM delay computations. Hence queue time is used as the MOE in the evaluation procedure.

A Latin hypercube sampling procedure is utilized to sample the various volumes and 100 combinations of the volumes that encompass the domain of the demand variability are generated. Thus 100 demand conditions are generated and these 100 demand conditions are run in CORSIM using the timing plans obtained from GA and SYNCHRO (shown in Table2). For every demand condition, CORSIM is run five times and the average delay from these five runs is used to compare the results from the GA and SYNCHRO optimized timing plans.

The queue time obtained from CORSIM is used in comparing the signal timing plans from GA and SYNCHRO. The 100 delay values obtained using the GA timing plan and the SYNCHRO timing plan are plotted against each other in a scatter plot. The plot shows equally good performance from the GA timing plan and the SYNCHRO optimized timing plan. These results are then tested statistically using the paired T-test. A paired T-test is conducted to verify the hypothesis that the mean of the delays from these 100 demand conditions from both the timing plans is equal. Table3 shows the comparison of delay results of the GA timing plan for the average delay optimized condition and the SYNCHRO optimized condition. The table shows that the delay values from GA and SYNCHRO optimized timing plans are not significantly different. The T-test suggests

that the average values from GA and SYNCHRO are equal and the hypothesis cannot be rejected. However, standard deviation of the delay results from the Table3 suggests that GA performed better than SYNCHRO by reducing the variance even though there is no significant improvement in the average delay estimates.

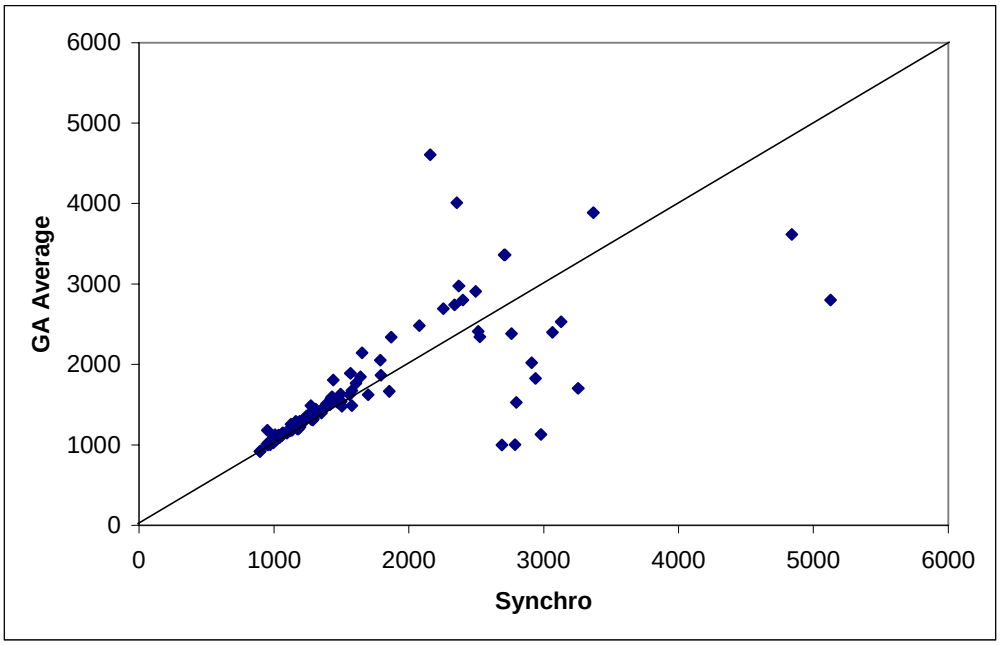

Figure 31. XY scatter plot of the delay values from GA optimized timing plan and SYNCHRO optimized timing plan

Similarly, the delay estimates from the signal timing plan obtained by optimizing the 95th percentile of delay are compared with the delay values from the SYNCHRO optimal signal timing plan. The same demand conditions generated for the previous example are utilized here. X-Y plots of SYNCHRO vs GA 95th percentile are presented in Figure 32. The plot suggests that SYNCHRO timing plan seems to be working better for most of the demand conditions when compared to the GA 95th percentile optimized condition. A T-test is also conducted on the result and the two variables are tested for the null hypothesis that the difference of the means are equal to zero. However, a t value of

considerably less than 0.05 is obtained suggesting that the means are different with a confidence level of 95%. Thus it is noted that the delays obtained from the SYNCHRO optimized timing plans works better than the GA 95th percentile optimized timing plan. Figure 32 and Table 13 present the results.

Table 13. Paired T-test on the delay estimates from GA optimized timing plan and SYNCHRO optimized timing plan.

	SYNCHRO	GA-Average
Mean	1668.7	1670.9
Standard Deviation	806.5	740.4
Observations	100	100
P(T<=t) two tail	0.972	
t critical two-tail	1.984	
	SYNCHRO	GA-95th Percentile
Mean	1668.7	1904.5
Standard Deviation	806.5	700.6
Observations	100	100
P(T<=t) two tail	0	
t critical two-tail	1.984	

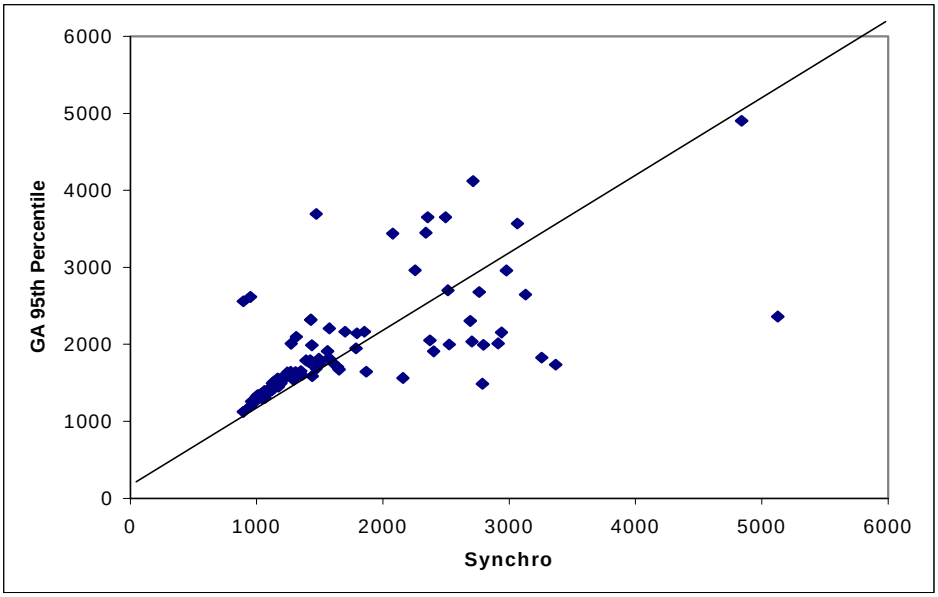

Figure 32. XY scatter plot of the delay values from GA optimized timing plan and SYNCHRO optimized timing plan

CHAPTER 8. CONCLUSIONS AND RECOMMENDATIONS

8.1 Conclusions

A series of methodologies have been developed to compute the variance of delay for undersaturated, oversaturated and arterial intersections. In order to estimate the variability in HCM delay equation, the following two methods are developed and validated through Monte Carlo simulations.

- Expectation Function Method: Works for undersaturated conditions and
- Integration Method: Works for both undersaturated and oversaturated conditions.

Cycle-by-cycle and day-to-day demand variability conditions can be modeled using the above methodologies. Distributions like Normal, Poisson, Lognormal and Exponential distributions have been utilized in delay variance computations. The average delay is consistent with different input demand distributions. However, the variance of delay is observed to depend upon the type of the demand distribution. Lognormal distribution with a big tail on both the directions resulted in a higher delay variance compared to distributions like Normal. This proves that LOS should not be determined solely based on the average control delay. In order to consider delay variability in level of service (LOS) determination, a new LOS scheme that considers variability in delay as and the level of service performance index (LOSPI) are proposed and demonstrated through an example.

Delay variability of arterial intersections is computed using the platoon dispersion model. An analytical equation of the platoon dispersion model in a closed form was developed. This equation is utilized to compute the percentage of arrivals during green at

the downstream intersection, which in turn determines the progression factor. Integration methodology was developed for the delay variance computations for an arterial.

The delay variance estimates from the methodologies were used in the optimization of signalized intersections under variable demand conditions. A genetic algorithm based approach was developed and the optimization process was coded into a C++ program. In order to evaluate the performance of the proposed program, optimization results or the GA timing plan was compared with the SYNCHRO timing plan using microscopic simulation programs CORSIM and SIMTRAFFIC. Firstly, the GA program was optimized using one of the two objective functions namely, average delay and 95th percentile delay. The timing plan from both these objective functions is compared with the SYNCHRO optimal timing plan. When evaluated in CORSIM using the queue time, both the average and 95th percentile optimized timing plans provided better results compared to SYNCHRO especially for an isolated intersection. The timing plan from the average delay optimization resulted in a better system performance than the timing plan with 95^{th} percentile optimized condition. SIMTRAFFIC evaluations were similar to that obtained from CORSIM. The GA average and 95th percentile optimized timing plans performed better than the SYNCHRO timing plan and further, the timing plan from the average delay optimization produced better results.

The proposed GA based stochastic optimization program was extended to an arterial intersection by utilizing the progression factor. The performance of the program was compared with SYNCHRO using CORSIM. For arterial intersections, GA timing plan did not show any improvement over the SYNCHRO timing plan. This could be

attributed to the progression factor definition being used in the methodology and is left to future research to study the effect of the progression factor on the optimization.

8.2 Recommendations

Based on the study conducted in this report, the following recommendations were made:

- It is stressed upon that delay has to be identified through a distribution rather than a point estimate, the average delay. Furthermore, it has also been identified that the variability in delay is sufficiently big enough for intersections to perform in more than one LOS range. Thus it is recommended that the variability in delay be considered in the determination of LOS, rather than determining the LOS at a signalized intersection based on the average delay. A newly proposed LOS scheme or LOSPI should be considered for updating existing HCM LOS criterion.
- Different demand distributions have shown consistently similar average delays and hugely different delay variance values. Well spread out distributions have resulted in high delay variance compared to compact distributions. Thus it is of primary importance to identify the distribution of demand volumes before proceeding with the determination of the LOS at a signalized intersection.
- Data on vehicular headways should be collected to determine the distribution of demand arrivals in addition to volume counts. Further, this data collection process could be extended to estimation of delay variability in the field to validate the delay variance computation methodology. So far, delay variance estimates from this research have been validated through analytical sampling procedures. Field evaluations of the results have not been considered due to fiscal and time

constraints. This element of the project is left for future research to collect the delay data in the field at different signalized intersections under different demand intensive conditions and evaluate it with the results from the corresponding methodology for a more comprehensive analysis.

- Furthermore, delay variance estimates have been used in the optimization of signalized intersections and the result of the optimization have shown considerable improvement over the SYNCHRO timing plan especially for isolated intersections. This development could be made use of in designing signalized intersections effectively.
- However, results for arterial intersections have not shown significant improvement over SYNCHRO. This has not be studied extensively and hence it is recommended that factors affecting the arterial results be studied in detail. Factors especially like progression factor are estimated differently from HCM and simulation models which result in different delay estimates and this could be one of the primary reasons for the results.
- Finally, a C++ program has been developed as a part of this research to aid in the delay variance computations. The process of data input to the program is through a notepad and has to be thoroughly understood before changing values. A program with a better user interface could be desired by researchers as it could be very useful in identifying the delay variance at signalized intersections and also in the optimization process.

REFERENCES

1. Webster, F. V. (1958). *Traffic Signal Settings*. Road Research Laboratory Technical Paper No. 39, HMSO, London, U.K.

2. Akcelik, R. The Hoghway Capacity Manual Delay Formula for Signalized Intersections, *ITE Journal*, March 1988, pp 23-27.

3. TRB, Highway Capacity Manual 2000.

4. Engelbrecht, R., D. Fambro and N. Rouphail, "Validation of Generalized Delay Model for Over saturated Conditions", *Transportation Research Record 1572*, TRB, National Research Council, Washington, D.C., pp. 122-130.

5. Olszewski P. S., "Modeling Probability-Distribution of Delay at Signalized Intersections", *Journal of Advanced Transportation*, 1994, Vol 28, No 3, pp. 253-274.

6. Fu L. P., Hellinga B., "Delay variability at signalized intersections", *Transportation Research Record 1710*, 2000, pp. 215-221

7. Rouphail, N. Progression Adjustment Factors at signalized Intersections. *Transportation Research Record 1225*, TRB, National Research Council, Washington, D.C., 1989, pp. 8-17.

8. Staniewicz, J. M., Levinson, H. S., Signal Delay with platoon Arrivals, *Transportation Research Record 1005*, TRB, National Research Council, Washington, D.C., 1985 pp. 28-37

9. Olszewski, P. Traffic Signal Delay Model for Non-Uniform Arrivals. *Transportation Research Record 1287*, TRB, National Research Council, Washington, D.C., 1990, pp. 42-53

10. Teply, S., Evans, G. D., Evaluation of the Quality of Signal Progression by Delay Distributions, *Transportation Research Record 1225*, TRB, National Research Council, Washington, D.C., 1989, pp. 01-07

11. Robertson, D. I. *TRANSYT: A Traffic Network Study Tool*. Road Research Laboratory, RRL report LR 253, Crowthorne, Berkshire 1969.

12. Rouphail N. M., Analysis of TRANSYT Platoon Dispersion Algorithm, *Transportation Research Record 905*, TRB, National Research Council, Washington, D.C., 1983, pp. 72-80.

13. Mc Coy P. T., Balderson E. A., Hsueh R. T., Mohaddes A. K., Calibration of TRANSYT Platoon Dispersion Model for Passenger Care Under Low-Friction

Traffic Flow Conditions . *Transportation Research Record 905*, TRB, National Research Council, Washington, D.C., 1983, pp. 48-52.

14. *TRANSYT-7F User's Guide*, U.S. Department of Transportation, Washington DC, 1991.

15. Anant Pradhan, Kara Maria Kockelman., 2001. Uncertainty propagation in an integrated land use transportation model. *Transport Geography* 20 August 2001

16. McKay M. D., Beckman R. J., Conover W. J., A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics,* 21: 239-245 1979.

17. Aditya Tyagi; C. T. Haan., Reliability, Risk and Uncertainty Analysis Using Generic Expectation Functions. *Journal Of Environmental Engineering*, Vol 127, No. 10, 2001, pp. 938-945.

18. Olszewski P. S., "Modeling Probability-Distribution of Delay at Signalized Intersections", *Journal of Advanced Transportation*, Vol. 28, No 3, 1994, pp. 253-274.

19. *SYNCHRO User's Guide*, Traffic ware 2000.

20. Goldberg, D. E., *Genetic Algorithms in search, Optimization and Machine learning*. Addison-Wesley Publishing Co., Massachusetts, 1989.

21. Foy, M. D., R. F. Benekohal, and D. E. Goldberg. Signal Timing Determination Using Genetic Algorithms. *Transportation Research. Record 1365*, TRB, National Research Council, Washington, D.C., 1993, pp. 108-115.

22. Park, B., C. J. Messer and T. Urbanik II, "Traffic Signal Optimization Program for Oversaturated Conditions: A Genetic Algorithm Approach." *Transportation Research Record 1683*, TRB, National Research Council, Washington, D.C., 1999, pp. 133-142.

23. Park, B., C. J. Messer and T. Urbanik II, "Enhanced Genetic Algorithm for Signal Timing Optimization of Oversaturated Intersections" *Transportation Research Record 1727*, TRB, National Research Council, Washington, D.C., 2000, pp. 32-41.

24. Akcelik R., Rouphail N. M., (1993) "Estimation of Delays at Traffic Signals for Variable Demand Conditions", *Transportation Research, Part B*, Vol 27, No 2, 1993, pp. 109-131.