

FINAL
CONTRACT REPORT

**EVALUATION OF TRAFFIC SIGNAL
TIMING OPTIMIZATION METHODS
USING A STOCHASTIC AND MICROSCOPIC
SIMULATION PROGRAM**

B. BRIAN PARK, Ph.D.
Research Assistant Professor
Department of Civil Engineering
University of Virginia

J. D. SCHNEEBERGER
Graduate Research Assistant



Standard Title Page - Report on Federally Funded Project

1. Report No. FHWA/VTRC		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle Evaluation of Traffic Signal Timing Optimization Methods Using A Stochastic and Microscopic Simulation Program				5. Report Date November 2002	
				6. Performing Organization Code	
7. Author(s) Brian Park and J.D. Schneeberger				8. Performing Organization Report No.	
9. Performing Organization and Address Virginia Transportation Research Council 530 Edgemont Road Charlottesville, VA 22903				10. Work Unit No. (TRAIS)	
				11. Contract or Grant No. 57935	
12. Sponsoring Agencies' Name and Address Virginia Department of Transportation FHWA 1401 E. Broad Street P.O. Box 10249 Richmond, VA 23219 Richmond, VA 23240				13. Type of Report and Period Covered Final: February 2001 – September 2002	
				14. Sponsoring Agency Code	
15. Supplementary Notes					
16. Abstract This study evaluated existing traffic signal optimization programs including Synchro, TRANSYT-7F, and genetic algorithm optimization using real world data collected in Virginia. As a first step, a microscopic simulation model, VISSIM, was extensively calibrated and validated using field data. Multiple simulation runs were then made for signal timing plans such that drivers' behavior, day-to-day traffic variation, etc were considered in the evaluation. Finally, long-term demand growth or changes were statistically modeled and evaluated, again using multiple simulation runs. Five timing plans were evaluated using the simulation test bed: 1) VDOT's former timing plan, 2) VDOT's current timing plan, 3) the genetic algorithm optimized timing plan, 4) the Synchro optimized timing plan, and 5) the TRANSYT-7F optimized timing plan. The simulation study results indicated that the current practice of VDOT signal optimization procedure significantly improves upon its former one by reducing travel times by 17% and total system delay by 36%. The three "optimized" timing plans did not provide significant improvements. Evaluation of the Lee Jackson Memorial Highway network showed that the current VDOT signal optimization procedure significantly improved the performance of network operations. Thus, the study recommended that VDOT continue using its procedure for developing new timing plans, but that it evaluate its signal timing plan regularly so that it does not become outdated.					
17 Key Words Traffic Signal Optimization Stochastic Simulation Model Simulation Model Calibration and Validation				18. Distribution Statement No restrictions. This document is available to the public through NTIS, Springfield, VA 22161.	
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of Pages 33	22. Price

FINAL CONTRACT REPORT

**EVALUATION OF TRAFFIC SIGNAL TIMING OPTIMIZATION METHODS USING
A STOCHASTIC AND MICROSCOPIC SIMULATION PROGRAM**

B. Brian Park, Ph.D.
Research Assistant Professor
Department of Civil Engineering
University of Virginia

J. D. Schneeberger
Graduate Research Assistant

Project Managers

Michael A. Perfater, Virginia Transportation Research Council
Catherine C. McGhee, Virginia Transportation Research Council

Contract Research Sponsored by
the Virginia Transportation Research Council

Virginia Transportation Research Council
(A Cooperative Organization Sponsored Jointly by the
Virginia Department of Transportation and
the University of Virginia)

Charlottesville, Virginia

January 2003
VTRC 03-CR12

NOTICE

The project that is the subject of this report was done under contract for the Virginia Department of Transportation, Virginia Transportation Research Council. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Virginia Department of Transportation, the Commonwealth Transportation Board, or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

Each contract report is peer reviewed and accepted for publication by Research Council staff with expertise in related technical areas. Final editing and proofreading of the report are performed by the contractor.

Copyright 2003 by the Commonwealth of Virginia.

ABSTRACT

Traffic signal timing optimization has been recognized as one of the most cost-effective methods for improving mobility within the urban transportation system. Inappropriate signal timing plans can cause not only discomfort (extra delay) to drivers but also increased emissions and fuel consumption. Thus, it is important to investigate the practice of signal optimization methodology to ensure that newly developed timing plans will improve the system performance. The investigation can be conducted via either field testing or the use of a reliable simulation tool. Due to the risk and time requirements of field testing, simulation tools are being widely used. However, simulation models must be properly calibrated and validated so that the model output can be trusted.

This study evaluated existing traffic signal optimization programs including Synchro, TRANSYT-7F, and genetic algorithm optimization using real-world data collected in Virginia. As a first step, a microscopic simulation model, VISSIM, was extensively calibrated and validated using field data. Multiple simulation runs were then made for signal timing plans such that drivers' behavior, day-to-day traffic variation, etc. were considered in the evaluation. Finally, long-term demand growth or changes were statistically modeled and evaluated, again using multiple simulation runs.

Five timing plans were evaluated using the simulation test bed. The timing plans under evaluation included (1) the former timing plan of the Virginia Department of Transportation (VDOT), (2) VDOT's current timing plan, (3) the genetic algorithm optimized timing plan, (4) the Synchro optimized timing plan, and (5) the TRANSYT-7F optimized timing plan. The simulation study results indicated that the current practice of VDOT's current signal optimization procedure significantly improved upon its former one by reducing travel times by 17% and total system delay by 36%. The three "optimized" timing plans did not provide significant improvements.

Evaluation of the Lee-Jackson Memorial Highway network showed that the current VDOT signal optimization procedure significantly improved the performance of network operations. Thus, the study recommended that VDOT continue using its procedure for developing new timing plans but that it evaluate its signal timing plan regularly so that it does not become outdated.

FINAL CONTRACT REPORT

EVALUATION OF TRAFFIC SIGNAL TIMING OPTIMIZATION METHODS USING A STOCHASTIC AND MICROSCOPIC SIMULATION PROGRAM

B. Brian Park, Ph.D.
Research Assistant Professor
Department of Civil Engineering
University of Virginia

J. D. Schneeberger
Graduate Research Assistant

INTRODUCTION

A great deal of effort has been dedicated to the design of better traffic signal timing plans for the last four decades, but very little research has been conducted in the evaluation of signal timing plans, especially in the context of stochastic and microscopic simulation environments. Furthermore, stochastic and microscopic simulation-based signal optimization has been very limited mainly due to heavy computational burden. Thus, the state of the practice has been to use programs including TRANSYT-7F, Synchro, and PASSER-II to optimize traffic signal timing plans based on the embedded macroscopic simulation models. With advances in computation technology, the use of microscopic simulation becomes more feasible. Thus, an optimization and evaluation of traffic signal timing plans based on stochastic and microscopic simulation seems very natural since the urban street network itself contains a variety of stochastic aspects including different drivers' behavior, vehicle mix, and day-to-day demand variations. The use of deterministic and macroscopic simulation-based signal optimization methods could easily result in a local optimum or even a poor solution. A recent study also indicates that a signal timing plan based on a direct signal optimization using a stochastic and microscopic simulation model produces better performance than that of a macroscopic simulation-based method (Rouphail et al., 2000). The similar study also indicated that the use of a well-calibrated simulation program is crucial in the evaluation of signal timing plans (Park et al., 2001).

It is often practical to use a computer simulation model in evaluating a newly developed traffic signal timing plan before its actual field implementation. The value of a stochastic simulation model is its ability to account for system variability through repeated model runs. If a sufficient number of runs are conducted, the daily variations in traffic flow can be factored into the results. Even though a few researchers have used stochastic and microscopic simulation programs for such purposes, the issue of variability was not well addressed in their efforts. In some cases, the number of replications is relatively small such that the distribution of system performance was not investigated. At times, with a highly variable system, the results (usually the mean of system performance) could be misleading if only a small number of replications were used. This project develops and uses the distribution of system performance in evaluating the performance of traffic signal timing plans. It is important to investigate system performance

with respect to not only the mean value but also its variability. It might be advantageous to implement a timing plan that produces a slightly higher average system delay with tighter delay variations than a plan that yields a lower average system delay with higher system delay variations. For example, a timing plan that produces the average system delay of 32 seconds per vehicle with a standard deviation of 5 should be implemented rather than a timing plan that produces the average system delay of 28 seconds per vehicle with a standard deviation of 10.

When a new timing plan is evaluated via a stochastic and microscopic computer simulation program, demand fluctuations are usually discarded. Even though multiple runs of stochastic simulation produce random variability, they do not account for significant demand fluctuations. When the evaluation of a signal timing plan is conducted in an area such as Northern Virginia, demand fluctuations ought to be carefully considered. A simple Monte Carlo simulation, one of the most common simulation methods, could be used to account for such demand fluctuations. However, this would require a significant amount of resources and might not be practical. This project uses an advanced statistical method so that demand fluctuations are systematically considered in an efficient manner.

PURPOSE AND SCOPE

The purpose of this study was to evaluate current signal optimization methods. The investigated methods included Synchro, TRANSYT-7F, and genetic algorithm optimization. It was anticipated that the evaluation would yield the following:

- calibrated input parameters based on field data collected in Virginia for a stochastic and microscopic simulation program tested in this study
- a well-calibrated traffic signal evaluation test bed
- a methodology that can evaluate the distribution of system performance of a newly developed traffic signal timing plan for various demand and input conditions before its actual field implementation.

Case studies used in the evaluation of simulation optimization methods were confined to the test-bed network consisting of 12 signalized intersections along Route 50 in Northern Virginia.

METHODS

The research approach in this study involved (1) a literature review, (2) development of a test bed, (3) calibration and validation of the simulation model, and (4) signal timing plan optimization and evaluation.

Literature Review

Literature was reviewed the current practices in the simulation model calibration and validation, traffic signal optimization programs (e.g., Synchro, TRANSYT-7F, and Genetic Algorithm approach), and current optimizations practices.

Test Bed Development

Site Selection

The test site was chosen with the help of Northern Virginia Smart Traffic Signal System personnel. The research team made site visits before making final decisions.

Data Collection

Data collection was required to provide simulation program input parameters and output measures of performance for the calibration and validation of the microscopic simulation model. Some data were provided from VDOT plans; other data were collected directly from the field.

The geometric characteristics of the test network were obtained from a Synchro file used by VDOT. Distances between intersections, lengths of left- and right-turn lanes, intersection grades, detector locations, and speed limits were collected from this file. The timing plan from this file was also taken. This timing plan was labeled as the “base” timing plan, the timing plan used by VDOT before it implemented its new timing plan, which was determined using Synchro. The signal timing plan currently used by VDOT in the field (referred to as the “field” timing plan) for the 12 intersections were extracted from the Management Information System for Transportation (MIST) terminal located in the Smart Travel Laboratory at the University of Virginia. This system is directly linked to the timing plans used in the field at the test site and therefore gives real-time data. Phases, splits, minimum green times, offsets, and gap out times were all taken from the system and used modeling in the microscopic simulation model, called VISSIM (2001).

Although the MIST system terminal provided detector data, turning movement counts, especially for shared lane approaches, had to be collected in the field. The MIST system terminal provides 15-minute local and system detector data on most approaches in the test network, but not all of them. Additionally, the unreliability of traffic volume counts from loop detectors made field collections of volumes a necessity. Both manual and video counts were conducted in the field along the test network. Counts were collected on a normal weekday, Wednesday, July 11, 2001, between 4:45 p.m. and 6:15 p.m.

A group of 16 people performed simultaneous manual counts along the test network. Manual counts were conducted at locations along the test site where detector data from MIST were not suitable or obtainable. In order to use the data collected from the manual counts simultaneously with the MIST data, synchronization between clocks used by the manual counters and the clock at the MIST system was performed before data collection. Shared lane approaches were counted manually since loop detectors cannot distinguish between turning and

through movements. Manual counts were also conducted where there were no loop detectors or where the loop detectors were not working properly.

Four PATH dual cameras were used to conduct the video counts. These cameras, provided by VDOT, are a pole-mounted video system. The cameras were positioned on Lee-Jackson Memorial Highway at the intersections of Highland Oaks Drive, Intel Country Club Road, Stringfellow Road, and Lees Corner Road, where they recorded traffic conditions at the intersections.

The videotapes of these intersections were used to obtain accurate volume counts and turning percentages along the arterial and local streets. These video counts are more accurate than manual or loop detector-based MIST counts for several reasons. Manual counts involve some human error, especially when a counter is required to observe more than one movement at high volumes. Since counters are observing at real time, it is possible that they may miss some vehicles when conducting counts. For example, it was often the case that manual counters missed right-on-red counts because they were busy watching other movements. Loop detector data on congested arterials do not always prove to be accurate either. Thresholds for monitoring a vehicle's presence are not always accurately set. As a result, volumes can be either over- or underestimated depending on how the threshold for determining a vehicle's presence is set. Accuracy in video counts is possible mainly because the viewer may view the videotape more than once. Therefore, the viewer can concentrate on a single movement and then, when finished, can review the tape and observe a different movement.

Eastbound leftmost lane travel times on Lee-Jackson Memorial Highway were used as a measure of performance for the calibration process. This measure of performance was selected based on the ease of data collection from the field and output results from microscopic simulation program. A direct comparison could easily be made between the field travel times and simulation travel times.

Travel times were determined from two cameras with synchronized clocks that were positioned on the bridges of Sully Road and Fairfax County Parkway overlooking the test site. The cameras were focused on the license plates of vehicles traveling in the leftmost lane along Lee-Jackson Memorial Highway and recorded a vehicle's license plate number when it entered and left the test network. License plate numbers and times were recorded for every vehicle at the beginning and end of the network and later matched. Subtracting the time the vehicle left the network from the time the vehicle entered gave the vehicle's eastbound travel time in the leftmost lane.

The Smart Travel Laboratory Van, which uses an AUTOSCOPE video detection system, was used to measure and collect queue lengths, the validation measure of performance. Queue lengths were collected on a weekday in August 2001 so they could be used as an untried data set for the validation process. The van was parked along the shoulder of Lee-Jackson Memorial Highway between the intersections of Muirfield Lane and Intel Country Club Road. The van provided video of a small segment of Lee-Jackson Memorial Highway between the two intersections. Queue lengths in the eastbound direction on the highway were collected from the

van's videotape by counting the number of vehicles in a queue at the end of red time for each cycle during the data collection period.

Simulation Program and Network Coding

Microscopic Simulation Model: VISSIM

A stochastic and microscopic program, VISSIM (VISSIM, 2001), was chosen for this study because it provides application programming interface (API) and powerful 3D animations. VISSIM version 3.50 is a microscopic simulation model with similar structures and capabilities as CORSIM (CORSIM, 1997). VISSIM is a microscopic, time step, and behavior-based simulation model. The model was developed at the University of Karlsruhe, Germany, during the early 1970s. Commercial distribution of VISSIM began in 1993 by PTV Transworld AG, who continues to distribute and maintain VISSIM today.

Essential to the accuracy of a traffic simulation model is the quality of the actual modeling of vehicles or the methodology of moving vehicles through the network. In contrast to less complex models using constant speeds and deterministic car following logic, VISSIM uses the psychophysical driver behavior model developed by Wiedemann in 1974 (VISSIM, 2001). The basic concept of this model is that the driver of a faster moving vehicle starts to decelerate as he or she reaches his or her individual perception threshold to a slower moving vehicle. Since the driver cannot exactly determine the speed of that vehicle, his or her speed will fall below that vehicle's speed until he or she starts to slightly accelerate again after reaching another perception threshold. This results in an iterative process of acceleration and deceleration. Stochastic distributions of speed and spacing thresholds replicate individual driver behavior characteristics. The model has been calibrated through multiple field measurements at the Technical University of Karlsruhe, Germany.

The basic idea of the Wiedemann model is the assumption that a driver can be in one of four driving modes:

1. *Free Driving*: No influences of preceding vehicles are observable. In this mode the driver seeks to reach and maintain a certain speed, his or her individually desired speed. In reality, the speed in free driving cannot be kept constant but oscillates around the desired speed due to imperfect throttle control.
2. *Approaching*: The process of adapting the driver's own speed to the lower speed of a preceding vehicle. While approaching, a driver applies a deceleration so that the speed difference of the two vehicles is zero in the moment he or she reaches his or her desired safety distance.
3. *Following*: The driver follows the preceding car without any conscious acceleration or deceleration. He or she keeps the safety distance more or less constant, but again due to imperfect throttle control and imperfect estimation, the speed difference oscillates around zero.

4. *Braking*: The application of medium-to-high deceleration rates if the distance falls below the desired safety distance. This can happen if the preceding car changes speed abruptly or if a third car changes lanes in front of the observed driver.

For each mode, the acceleration is described as a result of speed, speed difference between the follower and leader, distance between the follower and leader, and the individual characteristics of driver and vehicle. The driver switches from one mode to another as soon as he or she reaches a certain threshold that can be expressed as a combination of speed difference and distance. For example, a small speed difference can be realized in only small distances, whereas large speed differences are apparent and require drivers to react much earlier. The ability to perceive speed differences and to estimate distances, as well as the desired speeds and safety distances, varies among the driver population. Because of the combination of psychological aspects and physiological restrictions of the driver's perception, the model is called a psychophysical car-following model.

Network Coding

The test site was coded in VISSIM Version 3.50 using the data collected from the field and VDOT information. The VISSIM model consists of two components: a simulator and signal state generator. The simulator is responsible for generating traffic and is where the network is graphically built. An aerial photograph of the study area was imported into the simulator. The network was then "digitized" over the aerial photograph, and attributes from the data collection were applied (e.g., lane widths, speed zones, detector locations). Although links are used in the simulator, VISSIM does not have a traditional node structure. This link-based structure allows flexibility to control traffic operation (e.g., yield conditions) and vehicle paths within an intersection.

The signal state generator is separate from the simulator and is where signal logic presides. Signal control logic is input here for each intersection as VAP (vehicle actuated programming) files. In these VAP files, signal characteristics are entered for the actuated signals including phase sequences, minimum green times, force-offs, and gap out times. The signal state generator reads detector information from the simulator every time step. The signal state generator decides the signal display during a time step based on the detector information.

Model Calibration and Validation

Individual parameters in VISSIM were adjusted or tuned so that the model accurately represented field measured or observed traffic conditions. An eight-step calibration and validation procedure was decided upon:

Determination of Measures of Effectiveness

The first step was to determine measures of effectiveness appropriate for calibration and validation. Performance measures, uncontrollable input parameters, and controllable input parameters were identified. It was important to identify all measures of effectiveness clearly before proceeding forward in the calibration and validation process.

Identification of Calibration Parameters

All calibration parameters within the microscopic simulation model were identified, and acceptable ranges for each of the calibration parameter were determined.

Experimental Design for Calibration

The number of combinations among feasible controllable parameters is too large such that possible scenarios cannot be evaluated in a reasonable time. A Latin hypercube design algorithm was used to reduce the number of combinations of calibration parameter scenarios.

Multiple Runs

In order to reduce stochastic variability, multiple runs were conducted for each scenario from the experimental design. The average performance measure and standard deviation were recorded for each of the runs.

Development of a Surface Function

A surface function, using the calibration parameters and measure of performance, was created from the results of the multiple runs.

Determination of Parameter Sets Based on Surface Function

The purpose of this step was to find an optimal parameter set that provides a close match with the field performance measure. Due to the fact that there could exist several parameter sets providing output close to the target (i.e., field performance measure), several parameter sets were considered.

Evaluation of Parameter Sets

Multiple runs were conducted to verify whether the parameter sets identified in the previous step generate statistically significant results. For each parameter set, a distribution of performance measure was developed and compared with the field measure. Visualization was also used to evaluate the models.

Collection of New Data Set for Validation

In order to validate the microscopic simulation model a new set of field data under untried conditions was collected. The “calibrated” model was then evaluated with the new data set.

Signal Timing Plan Optimization and Evaluation

Three traffic signal optimization tools were used to optimize the offsets on Lee-Jackson Memorial Highway: (1) linking a genetic algorithm (GA) to the calibrated VISSIM model, (2) coding and optimizing the network in Synchro, and (3) coding and optimizing the test site in TRANSYT-7F. The Synchro and TRANSYT-7F optimizations were straightforward. The test network was coded into the programs and optimized. The experimental setup for the GA optimization worked as follows:

- The GA code developed in Fortran 90 language is linked to VISSIM using Rexx (Cowlshaw, 1990), a script language-based computer program.
- Signal timing plans (offsets) in binary representation form are randomly produced in the GA. The Rexx code converts the timing plans into integer values and inserts them directly into a VISSIM input file.
- VISSIM makes multiple runs (four or eight per timing plan) and outputs travel times for each run as *.rsr files.
- A second Rexx code extracts the performance measure (eastbound and westbound travel times) for each signal plan from the corresponding VISSIM output file (*.rsr). The Rexx file takes a weighted average of travel times on Lee-Jackson Memorial Highway.
- The performance measures from the Rexx program are fed back to the GA. The GA evaluates the performance measures (attempts to minimize travel time) and then generates a new set of signal timing plans (offsets).
- The new timing plans are sent back to VISSIM where the process continues until all signal timing plans proposed by the GA optimizer are run.

Traffic signal timing plans were evaluated using the calibrated VISSIM test network as an unbiased evaluator. Five timing plans were under investigation and include (1) VDOT's former timing plan, (2) VDOT's current timing plan, (3) the genetic algorithm optimized timing plan, (4) the Synchro optimized timing plan, and (5) the TRANSYT-7F optimized timing plan. Evaluation of each timing plan was based on 100 VISSIM simulation runs. Measures of performance used to evaluate the timing plans were travel times on Lee-Jackson Memorial Highway and total system delay.

The responses of the timing plans to changes in mean volumes were also evaluated. Wide ranges of changes in demand volumes were explored: 15% from the base demands at each entry node. Since the number of possible demand pattern combinations was too large to explore, a Latin hypercube design was applied to the problem. Fifty demand variation combinations were created, and each timing plan was evaluated via multiple VISSIM simulations.

RESULTS AND DISCUSSION

Literature Review

Relevant literature was reviewed to gain insight into concepts and issues related to the microscopic simulation model, calibration and validation, and optimization. Information was obtained through an extensive search of publications.

Current Practices in Simulation Model Calibration and Validation

Microscopic simulation models contain numerous independent parameters to describe traffic control operation, traffic flow characteristics, and driver behavior. These models contain default values for each variable, but they also allow users to input a range of values for the parameters. Changes to these parameters during calibration should be based on field measured or observed conditions and should be justified and defensible by the user.

Unfortunately, many of the parameters used in simulation models are difficult to measure in the field, yet they can have a substantial impact on the model's performance. Examples of some of these variables in microscopic simulation models could include start-up lost time, queue discharge rate, car-following sensitivity factors, time to complete a lane change, acceptable gaps, and driver's familiarity with the network. This is why skeptics often view simulation modeling as an inexact science at best and an unreliable "black-box" technology at worst (Hellinga, 2002). This skepticism usually results from unrealistic expectations of the capabilities of simulation models, use of poorly verified or validated models, and/or use of poorly calibrated models (Hellinga, 2002).

It is understood that microscopic simulation model-based analyses have been conducted often under default parameter values or best guessed values. This is mainly due to either difficulties in field data collection or the lack of readily available procedures on the simulation model calibration and validation. At times, simulation model outputs could result in unrealistic estimates of the impacts of new treatments if the simulation model is not properly calibrated and validated. Thus, the calibration and validation for simulation models are crucial steps in assessing their value in transportation policy, planning and operations. Sacks et al. (2003) indicated that simulation model calibration and validation are often discussed and informally practiced among researchers but have not been formally proposed as a procedure.

Model calibration is defined as the process by which the individual components of the simulation model are adjusted or tuned so that the model will accurately represent field-measured or observed traffic conditions (Milam, 2002). The components or parameters of a simulation model requiring calibration include traffic control operations, traffic flow characteristics, and drivers' behavior. Model calibration is not to be confused with validation. *Model validation* tests the accuracy of the model by comparing traffic flow data generated by the model with that collected from the field (Milam, 2002). Validation is directly related to the calibration process because adjustments in calibration are necessary to improve the model's ability to replicate field-measured traffic conditions.

Hellinga (2002) described a calibration process consisting of seven component steps: (1) defining study goals and objectives, (2) determining required field data, (3) choosing measures of performance, (4) establishing evaluation criteria, (5) representing the network, (6) determining driver routing behavior, and (7) evaluating model outputs. This process provides basic guidelines but does not give a direct procedure for conducting calibration and validation.

Sacks et al. (2003) recognized four key issues on model validation: (1) identifying explicit meaning of validation in particular *context*, (2) acquiring relevant *data*, (3) quantifying *uncertainties*, and (4) *predicting* performance measures under new conditions. They demonstrated an informal validation process using CORSIM simulation model and emphasized the importance of data quality and visualization. The authors have not established any formal procedure for simulation model calibration and validation.

Cheu et al. (1998) and Lee and Yang (2000) used a GA to optimize parameters in FRESIM and PARAMICS, respectively. The GA was used to adjust default parameter values in the simulation models. The GA was also used to minimize differences between the 30-second loop detector output (volume and speed) from the simulation model and data collected from the real world. Variability of performance measures and visualization analysis was not emphasized in their calibration procedure.

Optimization Programs

Synchro

Synchro is a macroscopic and deterministic model for optimizing traffic signal timing plans. Synchro can optimize cycle lengths, green splits, phase sequences, and offsets. Splits are optimized by percentile, with Synchro attempting to provide enough green time to serve 90% of the flow from a lane group. If there is not enough cycle time to serve the 90% flow, 70%, 50%, etc., flow is then tried (Synchro, 1999). Any extra green time goes to the main street. Synchro attempts to determine the shortest cycle length that clears the critical percentile traffic when optimizing cycle lengths. Offset optimization is conducted through a semi-exhaustive search. It is not possible to perform an exhaustive search for every second. Instead, Synchro uses three steps to eliminate “bad” offset areas. The first step looks at every 8 seconds for offset values. The bad areas are eliminated. Second, it looks at every 4 seconds, eliminating the bad areas. Third, it looks at every second.

TRANSYT-7F

TRANSYT-7F is a macroscopic, deterministic optimization and simulation model originally developed in the United Kingdom by the Transport and Road Research Laboratory. TRANSYT-7F is a macroscopic model that considers platoons of vehicles instead of individual vehicles. The model simulates traffic flow in small time increments, so its representation of traffic is more detailed than other macroscopic models that assume uniform distributions within traffic platoons. A platoon dispersion algorithm that simulates the spreading out of platoons as they travel downstream is also used in TRANSYT-7F.

TRANSYT-7F optimizes signal timing by performing a macroscopic simulation of traffic flow within small increments while signal timing parameters are varied. Optimization can be performed two ways in TRANSYT-7F. The first approach uses GA, while the other uses the hill climbing method. GA optimization is a theoretical improvement over the traditional hill-climb optimization technique that has been employed by TRANSYT-7F for many years. The GA has the ability to avoid becoming trapped in a "local optimum" solution and is mathematically best qualified to locate the "global optimum" solution (TRANSYT-7F, 1998). Synchro enables its files to be converted to TRANSYT-7F files. The updated Synchro file with the timing plan currently in use was used to create a TRANSYT-7F file.

Genetic Algorithm

A GA is a search algorithm based on the mechanics of natural selection and evolution (Goldberg, 1989). It works with a population of individuals, each representing a possible solution to a given problem. Each individual is assigned a fitness value according to how good a solution to the problem it is. The highly fit individuals are given opportunities to reproduce by cross breeding with other individuals in the population. Selecting the best individuals from the current generation and mating them to produce a new set of individuals produce a new population of possible solutions.

A GA uses three basic operators: reproduction, crossover, and mutation, although further enhanced operators have been suggested and implemented. The reproduction operator selects individuals with higher fitness, whereas the crossover operator creates the next population from the intermediate population. Finally, the mutation operator is used to explore some areas that have not been searched. More details of GA can be found in related literature (Goldberg, 1989). Schema theorem and building blocks hypothesis are rigorous explanations of how GAs work. Simply put, schema theorem and building block hypothesis state the number of good components is likely to proliferate as the number of generations evolves (Beasley et al., 1993).

The GA-based signal optimization program consists of two main components: a GA optimizer and a microscopic traffic simulator, VISSIM. Figure 1 depicts the conceptual framework of the proposed program. The GA optimizer starts by randomly producing a generation of individuals (i.e., offset values). Each individual timing plan is then evaluated through the microscopic simulator. The next generation will be evolved from the GA optimizer on the basis of those fitness values obtained from the microscopic traffic simulator (Park et al., 1999). Weighted east and westbound travel times on Lee-Jackson Memorial Highway were used as the fitness value in this study. For example, in the case of a maximization problem, individuals showing higher fitness values are selected for mating to generate offspring through GA operators. The circulation process of Figure 1 is continued until the maximum number of generations is reached.

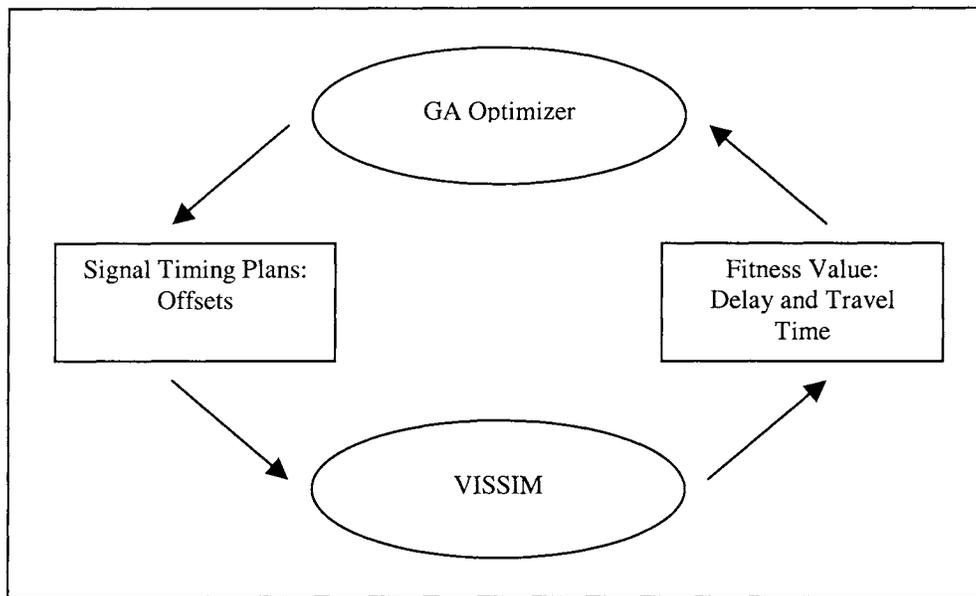


Figure 1. Conceptual Framework for GA-Based Signal Optimization Program

Traffic Signal Optimization Practices

There are a variety of computer software programs to aid transportation engineers in the analysis and optimization of signal timing plans. Intersection analysis helps to improve traffic signal operation (reduce delays, queues, and travel times) and reduce vehicle-operating costs (reduce fuel consumption). Arterial signal synchronization is one of the most cost-effective methods for reducing vehicle operating costs and improving traffic flow performance along urban arterials. Arterial signal optimization models, such as Synchro and TRANSYT-7F, have been developed to assist traffic engineers in coordinating traffic signal settings along urban arterials and around networks. Additionally, limited efforts have been made to use GAs for signal optimization.

Paracha (1999) conducted a study optimizing five intersections using Synchro and TRANSYT-7F. For each program, multiple simulation runs were made using CORSIM, a stochastic and microscopic simulation program. The simulations of the CORSIM simulation model were used to approximate how the timing plans would work in the real world. Results from the study showed that no single software package provided the best solution to all of the scenarios. The results indicated that both Synchro and TRANSYT-7F can be used effectively for optimization of signal timings at intersections with approximately equal effectiveness. This study was limited to isolated intersections, and distribution of variability was not considered.

Yang (2001) compared Synchro and TRANSYT-7F optimization programs. The goal of the study was to determine which package could best provide a timing plan to improve existing traffic performance along an arterial in Lawrence, Kansas. The test site included nine signalized intersections 16,050 feet in length. CORSIM was to evaluate the effectiveness of the signal timing plans. The study showed that Synchro coordination produced great improvement in

measures of effectiveness. TRANSYT-7F did not perform well. The study did not conduct simulation model calibration and validation. Additionally, only 12 model runs were conducted, so the distribution of variability was not considered.

Park et al. (2001) developed a test bed in Chicago consisting of nine signalized intersections using CORSIM and a GA to optimize signal timings. Taking CORSIM as the best representation of reality, the performance of the GA plan sets a ceiling on how good any (fixed) signal plan can be. An important aspect of this approach is its accommodations of variability. Also discussed was the robustness of an optimal plan under changes in demand. This benchmark was used to assess the best signal plan generated by TRANSYT-7F from among 12 reasonable strategies. The performance of the best plan fell short of the benchmark on several counts, reflecting the need to account for variability in the highly stochastic system of traffic operations, which is not possible under the deterministic conditions intrinsic to TRANSYT-7F. As a sidelight, the performance of the GA plan within TRANSYT-7F was also computed and was found to perform nearly as well as the optimum TRANSYT-7F plan.

Test Bed Development

Site Selection

An urban arterial street network in Fairfax, Virginia, was chosen for the test site. The site consists of an arterial, Lee-Jackson Memorial Highway (U.S. Route 50), and 12 coordinated actuated signals between Sully Road and the Fairfax County Parkway. The site was also chosen because of the ease at which signal timing plans and detector data for the 12 intersections could be extracted from the Management Information System for Transportation (MIST) workstation located in the Smart Travel Laboratory at the University of Virginia. This system is directly linked to the timing plans used in the field test site and therefore provides an access to real-time data.

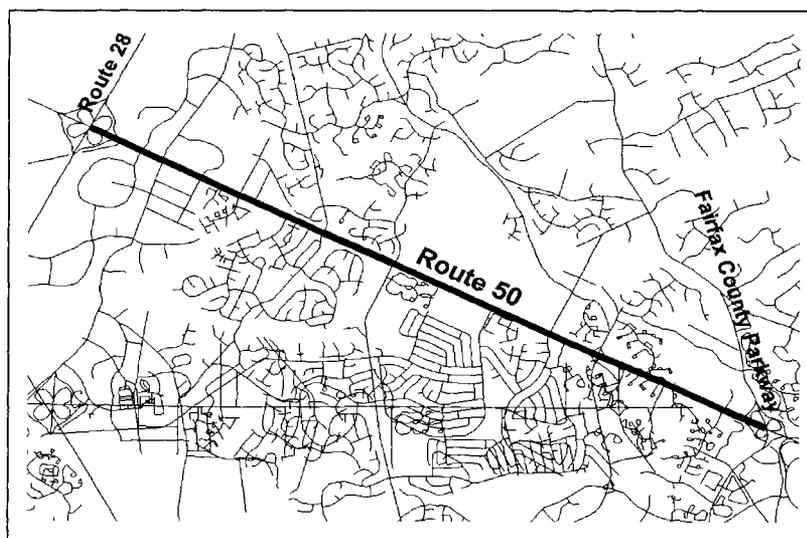


Figure 2. Test Site: Lee-Jackson Memorial Highway, Fairfax, Virginia

Data Collection

Traffic volumes, turning percentages and eastbound leftmost lane travel times on Lee-Jackson Memorial Highway and queue lengths at Muirfield Lane were collected from the field. Table 1 shows the base and field/collected traffic data aggregated by intersection on the test bed network. The base data were used in the base signal timing plan development and the field/collected traffic data were used for the development of the Synchro, TRANSYT-7F, and GA approaches and the evaluation of overall timing plans.

Table 1. Traffic Counts Data by Intersection

Intersection Name	Base ¹ (vehicles per hour)	Field/Collected ² (vehicles per hour)
Rugby Road	7328	6606
Highland Oaks Drive	6990	6025
Muirfield Lane	6688	5816
Intel Country Club	6518	5659
Stringfellow Road	7598	6707
Lees Corner Road	7062	6180
Galesbury Lane	6252	5733
Chantilly Road	6253	5767
Metrotech Drive	6808	6290
Centreville Road	8162	7010
Sullyfield Circle	7001	6058

1 Obtained from VDOT Synchro file.

2 Obtained during field data collection and from archived database.

A time series plot of left-lane eastbound travel times, collected by the license plate matching method, on Lee-Jackson Memorial Highway is shown in Figure 3. Figure 4 is a plot of the queue lengths at Muirfield Lane during the afternoon peak hour.

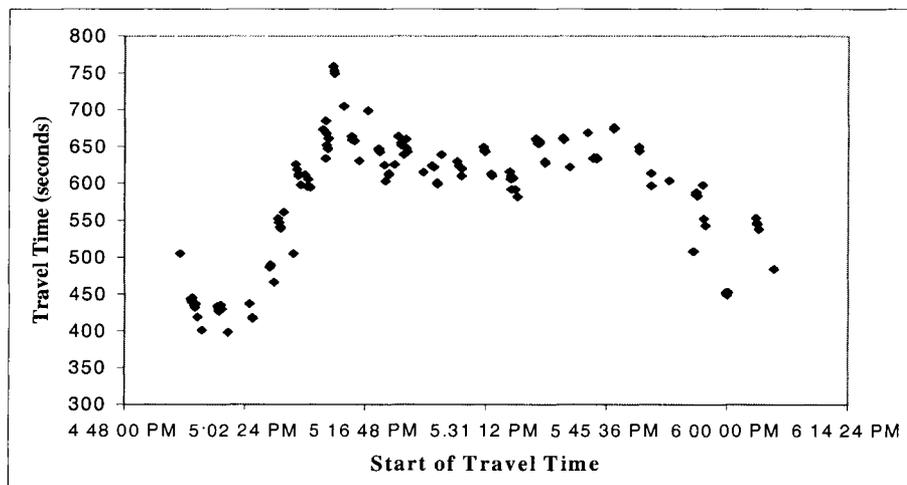


Figure 3. Eastbound Left Lane Travel Time Data

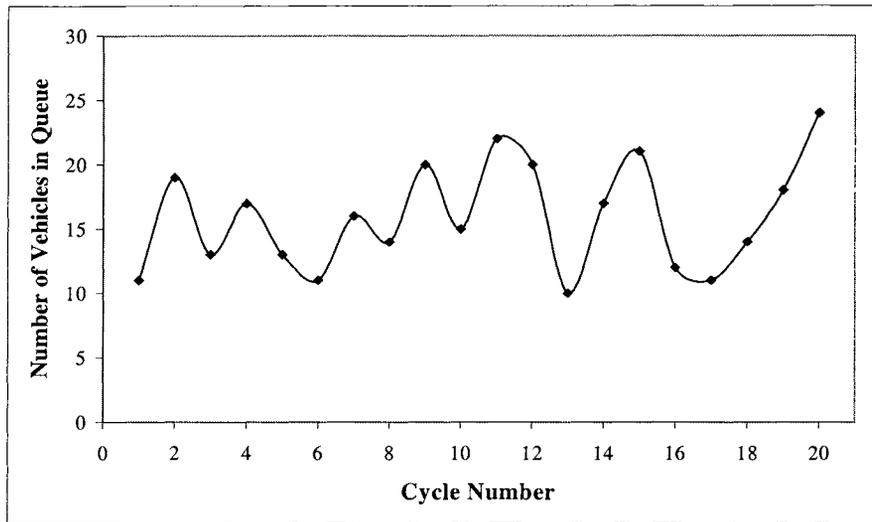


Figure 4. Queue Length Data at Muirfield Lane

Simulation Program and Network Coding

The test network was coded in VISSIM. The network coding process consisted of drawing the network in VISSIM along with inputting timing plans and traffic volumes.

An aerial photograph was used to draw the network in VISSIM. Links and connectors were traced over this image. VAP files, which contain the signal timing plans, were coded for each intersection. The offsets obtained from data collection (MIST) and optimization (Synchro) had to be converted prior to their coding into VISSIM. The offsets obtained from MIST and Synchro files reference offsets at the beginning of yellow for the coordinated phases. VISSIM required the offsets to be inputted with a reference to the start of green of the coordinated phases. The VISSIM network is depicted in Figure 5.

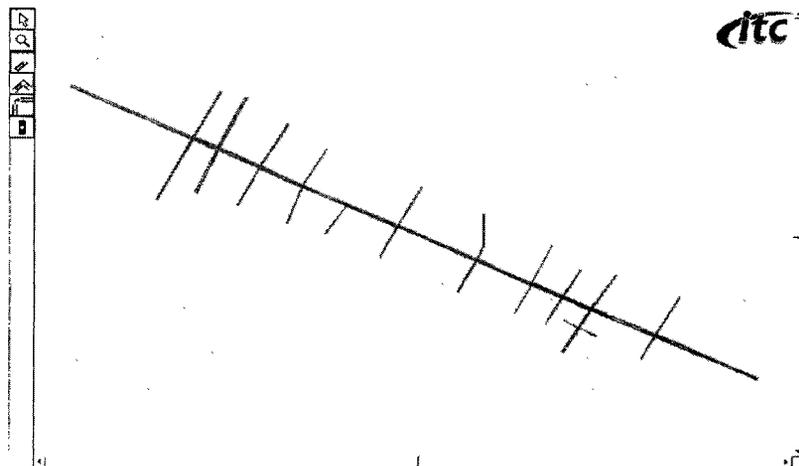


Figure 5. VISSIM Network

Simulation Model Calibration and Validation

The calibration and validation procedure applied in this study had eight steps. The following sections describe the results of the procedure.

Measure of Performance Selection

Two measures of performance were selected for the calibration and validation process. The first was eastbound left-lane travel times on Lee-Jackson Memorial Highway, which was used for calibration. The maximum queue length between the intersections of Muirfield Lane and Intel Country Club Road was used as the performance measure for the validation process. These performance measures were chosen because of their ease of collection from the field and from VISSIM output files. Other performance measures such as speed and delay are not easily obtainable from the field but may be obtained from simulation models.

Identification of Calibration Parameters

The following sections describe the VISSIM parameters and acceptable ranges used in the calibration process. Acceptable ranges were based on the researchers' discretion and their familiarity with VISSIM. These parameters included the emergency stopping distance, lane change distance, desired speed, number of observed preceding vehicles, average standstill distance, additive part of desired safety distance, waiting time before diffusion, and minimum headway.

Emergency Stopping Distance

The emergency stopping distance defines the last possible position for a vehicle to change lanes. For example, if a vehicle cannot change lanes due to high traffic flows but needs to change lanes in order to stay on its route, it will stop at this position to wait for an opportunity to change lanes. The emergency stopping distance is assigned for each link in the network. The default emergency stopping distance is 5.0 m. Acceptable ranges for the emergency stopping distance were determined to be between 2.0 and 7.0 m. These values were chosen because they enabled vehicles to make full use of the link for lane changes (up to 2.0, 3.0, etc., meters from the intersection). Larger values were not used because they would limit the amount of space a vehicle had on a link to attempt a lane change, thus forcing a vehicle to stop in the middle of the link waiting for an acceptable gap.

Lane Change Distance

The lane change distance parameter is used along with the emergency stopping distance parameter to model drivers' behavior in order to stay on their desired routes. The lane change distance defines the distance at which drivers will begin to attempt to change lanes. The default value for lane change distance is 200.0 m. Acceptable values for lane change distance were set between 150.0 and 300.0 m. These values were selected to ensure a driver had a reasonable distance to make a lane change before reaching the intersection. Values that are too small would force vehicles into the emergency stopped condition.

Desired Speed Distribution

The desired speed distribution is an important parameter, having a significant influence on roadway capacity and achievable travel speeds. The desired speed is the speed a vehicle “desires” to travel if it is not hindered by other vehicles. This is not necessarily the speed the vehicle travels in the simulation. If not hindered by another vehicle, a driver will travel at his or her desired speed (with small oscillations). The more vehicles differ in their desired speed, the more platoons are created. Any driver with a higher desired speed than his or her current travel speed will check for the opportunity to pass without endangering other vehicles. Minimum and maximum values can be entered in VISSIM for the desired speed distribution. The speed limit on Lee-Jackson Memorial highway was 45 mph. Acceptable ranges of speed were chosen as set distributions between 30 and 60 mph, 35 and 55 mph, and 40 and 50 mph. The desired speed distribution used in this exercise was 35 to 55 mph. This value was chosen based on prior experience with VISSIM. The 40 to 50 mph desired speed distribution was too tight. This distribution had all vehicles traveling at similar speeds, with little interaction between them. The 30 to 60 mph distribution was not chosen because it did not seem reasonable for a vehicle to have a desired speed of 30 mph.

Number of Observed Preceding Vehicles

The number of observed preceding vehicles variable affects how well drivers in the network can predict other vehicles’ movements and react accordingly. The VISSIM default value for this parameter is two vehicles. One, two, three, and four vehicles were used in this study.

Average Standstill Distance

Average standstill distance defines the average desired distance between stopped cars and also between cars and stop lines, signal heads, etc. The default value for average standstill distance is 2.0 m. Acceptable ranges of values used for this parameter were 1.0 to 3.0 m. Larger or smaller values seemed unreasonable.

Waiting Time Before Diffusion

Waiting time before diffusion defines the maximum amount of time a vehicle can wait at the emergency stop position waiting for a gap to change lanes in order to stay on its route. When this time is reached, the vehicle is deleted from the network. Sixty seconds is the default value. Other values used in the study were 20 and 40 seconds.

Minimum Headway

VISSIM defines the minimum headway distance as the minimum distance to the vehicle in front that must be available at standstill conditions for a lane change. This parameter could not be directly collected from the field. The default value is 0.5 m. The acceptable range used in the case study was between 0.5 and 7.0 m. The default value seemed too small a distance for

drivers to attempt a lane change. It did not seem realistic that a driver would attempt a lane change given headway of 0.5 m. As a result, larger values were assumed to be more reasonable.

Experimental Design for Calibration

A Latin hypercube experimental design was used for the calibration. Latin hypercube sampling provides an orthogonal array that randomly samples the entire design space broken down into equal-probability regions. This type of sampling can be looked upon as a stratified Monte Carlo sampling where the pair-wise correlations can be minimized to a small value (which is essential for uncorrelated parameter estimates) or else set to a desired value. Latin hypercube sampling is especially useful for exploring the interior of the parameter space and for limiting the experiment to a fixed (user specified) number of runs. The Latin hypercube technique ensures that the entire range of each variable is sampled. A statistical summary of the model results will produce indices of sensitivity and uncertainty that relate the effects of heterogeneity of input variables to model predictions. The Latin hypercube design consisted of 124 cases using the VISSIM parameters and three values per parameter.

Multiple Runs

Five random seeded runs were conducted in VISSIM for each of the 124 cases, for a total of 620 runs. The average eastbound left-lane travel time was recorded for each of the 620 runs. The results from the five multiple runs were then averaged to represent each of the 124 parameter sets.

Development of a Surface Function

A linear regression model was created in the S-Plus program using the calibration parameters as independent variables and the eastbound left-lane travel time from VISSIM as the dependant variable, Y. The linear regression model is:

$$Y = 400.88 - 5.10 X_1 - 0.68 X_2 + 17.80 X_3 + 28.63 X_4 + 1.77X_5 + 30.20 X_6$$

where,

- Y = eastbound left-lane travel time (sec)
- X₁ = emergency stopping distance (m), t value: -2.31, p value: < 0.0212
- X₂ = lane change distance (meters), t value: -9.15, p value: < 0.0001
- X₃ = number of observed preceding vehicles, t value: 3.69, p value: 0.0002
- X₄ = standstill distance (m), t value: 5.93, p value: < 0.0001
- X₅ = waiting time before diffusion (sec), t value: 7.32, p value: < 0.0001
- X₆ = minimum headway (m), t value: 11.37, p value: < 0.0001

Candidate Parameter Sets

Candidate parameter sets were created using the linear regression model. Microsoft Excel's Solver was used to obtain candidate parameter combination sets. The eastbound left-lane travel time, Y , was set to a target value of 613.16 seconds, the travel time observed from the field. The Excel Solver was then used to determine combinations of parameters producing travel time values close to 613.16 seconds. Eight combinations of parameters were created and are summarized in Table 2.

Table 2. Candidate Parameter Sets

Parameter	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8
Emergency Stopping Distance (m)	2.0	2.0	3.0	2.0	3.0	3.0	3.0	3.0
Lane Change Distance (m)	200	150	200	150	200	200	175	175
No. Observed Vehicles	3	2	3	3	4	4	4	4
Standstill Distance (m)	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
Waiting Time Before Diffusion (sec)	60	60	60	60	60	60	60	60
Minimum Headway (m)	2.5	2.5	3.0	2.0	3.0	3.5	3.0	2.5

Evaluation of Candidate Parameter Sets Using Multiple Runs

Fifty random seeded runs were made for each of the eight candidate parameters sets and evaluated based on two criteria: distribution of eastbound left-lane travel times produced from VISSIM and visualization. Based on these criteria, the best parameter set was selected.

Eastbound leftmost lane travel times were collected from 50 random seeded runs. Mean and median eastbound leftmost lane travel times for each of the eight cases are shown in Table 3. The results from these runs were compared to the travel times collected from the field. The average field travel time was 613.16 seconds with a standard deviation of 66.2. The field data were collected on a single day. It is not known whether the field data is an average representation of the travel times on Route 50. The data collected from the field may be average, but they may also be lower or higher than the true mean. Instead of comparing the average travel time from the 50 simulation runs and the field data, it is better to compare the field data and the distribution of the 50 runs. By comparing the field data and the distributions of travel times, variability was considered.

Table 3. Evaluation of Candidate Parameter Sets: Leftmost Lane Travel Time

Case	Mean (sec)	Median (sec)	Standard Deviation	Percentile Field Value	t test (p value)
1	449.6	439.5	57.5	100%	0.00
2	603.2	596.1	88.3	54%	0.42
3	473.7	465.3	64.3	98%	0.26
4	688.9	681.9	115.5	24%	0.40
5	467.9	455.9	59.3	98%	0.04
6	519.0	508.7	70.7	92%	0.85
7	523.2	523.2	70.6	90%	0.72
8	485.3	476.6	63.7	98%	0.00

The t test was used to determine if the travel times produced by VISSIM were statistically equal to the field travel times. In order to perform a t test, it was necessary to select the travel time results from a single VISSIM run and compare these travel times to those observed in the field. The VISSIM results that produced mean travel times closest to those observed from the field were selected for t test comparisons. Table 3 shows the results of the t tests. The percentile field values in the table show the percentage of runs, from the 50 multiple runs, which were less than the VISSIM results used for the t test.

Because of the variability in VISSIM runs and results, the t test is extremely sensitive. It is likely that in most cases, the t test would indicate that the field distribution of left-most lane travel time is statistically different from simulation-based results based on this variability. However, for a few runs close to the observed field travel times, the t test may conclude that the two are statistically the same.

The importance of visualization when using microscopic simulation models cannot be overemphasized. The purpose of the microscopic simulation model is to represent the field conditions as closely as possible. A model cannot be deemed calibrated if the animations are not realistic. For example, a parameter set may be statistically acceptable but the animations may not be realistic. Then, the model is not acceptable. Animations from the 50 multiple runs were viewed in order to identify at which percentile animations were not acceptable. An example of an unacceptable animation is depicted in Figure 6.

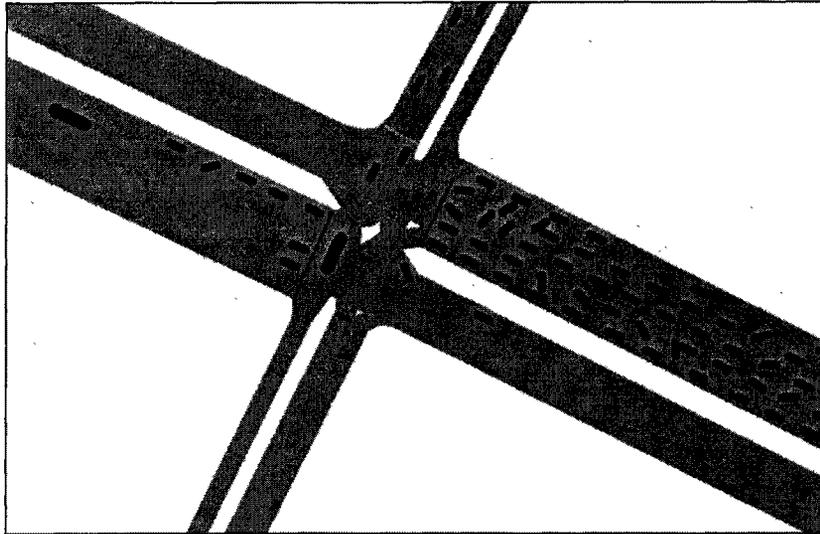


Figure 6. Example of an Unacceptable Animation in VISSIM

In this VISSIM screenshot, unrealistic animations occur in the westbound direction that were not observed in the field. Vehicles in the figure are attempting to make lane changes at the stop bar. Vehicles in the leftmost lane are trying to make right turns, and vehicles in the rightmost lane are attempting to make left turns. The vehicles were not able to make their desired lane change within the link and therefore are at an emergency stopped position. They will stay at that position, blocking other vehicles, until they are able to change lanes or they are kicked out of the system (due to waiting time before diffusion). Regardless of the travel times produced from an animation like this, this parameter set cannot be chosen because its visual does not represent the real world.

Animations of each case were viewed in order to determine whether the animations were realistic or unrealistic. Each case was viewed at several travel time percentiles in order to determine if the animations were realistic or not. It was found that cases 2 and 4 were not acceptable.

Parameter set 7 was chosen as the best parameter set based on its travel time distribution (Figure 7), statistical tests, and animations. Parameter sets 2 and 4 were eliminated because their animations were unrealistic. Parameter sets 1, 3, 5, and 8 were eliminated based on their t test results. Parameter 7 produced travel times closest to those from the field. Parameter sets 6 and 7 produced similar results, but parameter set 7 was chosen because its animations were more realistic.

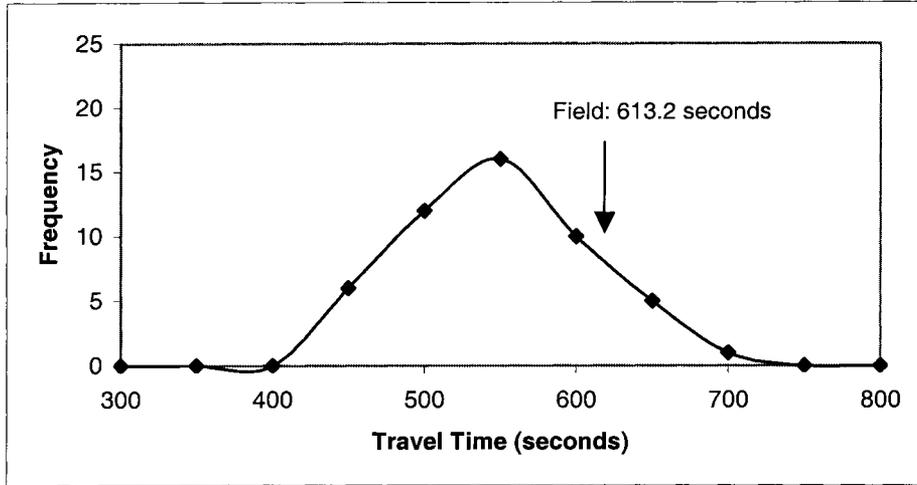


Figure 7. Eastbound Left Lane Travel Time Distribution of Parameter Set 7

Validation with New Field Data

The eastbound maximum queue length between the intersections of Muirfield Lane and Intel County Club Road was used for validation of parameter set 7. The maximum queue length data were collected on a different day, and the input volumes used for the validation process were untried. The maximum queue length observed in the field was compared to the distribution of 100 runs in VISSIM. The field maximum queue length was about the top 90 percent of the simulated distribution, as shown in Figure 8.

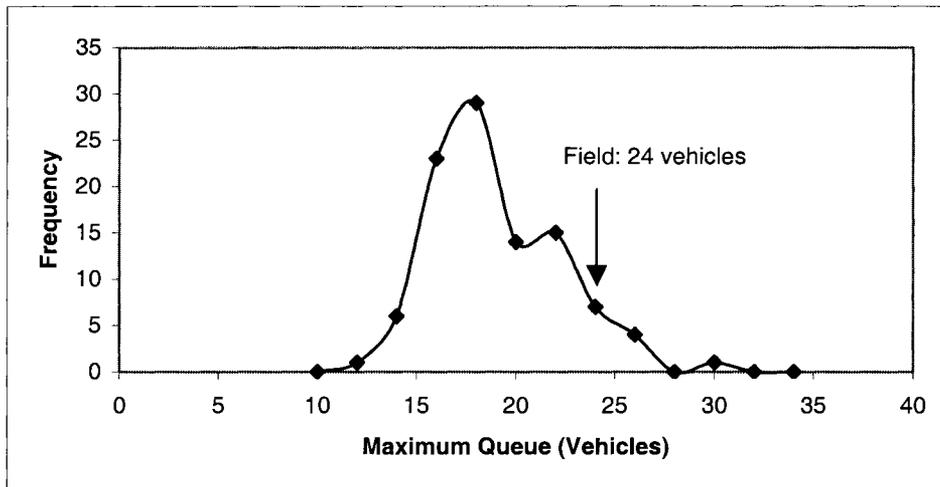


Figure 8. Maximum Queue Length Histogram of Parameter Set 7

Signal Timing Plan Optimization and Evaluation

Signal Timing Plan Optimization

The offsets of 9 of the 11 coordinated actuated signals on Lee-Jackson Memorial Highway were optimized in this study. In order to prevent “boundary effects,” the cycle length and the two entry intersection offsets are not optimized. The green splits could have been optimized, but the decision to optimize only the offsets on test network was made based on two criteria. First, the green splits are commonly determined based on the equal degree of saturation procedure, and their impact on improving the systems’ performance is relatively minimal. Second, the GA-based optimization was an extensive and time-consuming process. The inclusion of green splits optimization would have added significant computation burdens to this process.

The offsets at Rugby Road and Sullyfield Circle, the two intersections at the ends of the test network, were left as their current field offset values so that the new timing plan would not interfere with vehicle progression outside of the test network. Leaving the offsets of the two end intersections as their current values ensured that the new timing plan would not cause progression problems between the test network and the intersections outside of the network. Optimization was performed using traffic volumes collected from the field or the current volumes. Three optimization tools were used as discussed.

Synchro and TRANSYT-7F Optimization

The two existing programs were used to optimize offsets on Lee-Jackson Memorial Highway: Synchro and TRANSYT-7F, both macroscopic optimization programs. The Synchro file obtained from VDOT was updated with new flows, turning percentages, cycles lengths, splits, etc. Synchro was then used to optimize the offsets of the nine intersections in the test network. TRANSYT-7F was used to optimize the offsets of the nine intersections in the test network. Optimization in TRANSYT-7F was performed using GA. Table 4 shows the optimized offsets from Synchro and TRANSYT-7F.

Table 4. Lee Jackson Memorial Highway Offsets

Intersection	Base Conditions	Field Conditions	G.A. Optimized	Synchro Optimized	TRANSYT Optimized
Rugby Road	0	48	48	48	48
Highland Oaks	97	15	93	48	58
Muirfield Lane	90	0	107	25	173
Intel Club	97	157	87	4	92
Stringfellow Rd.	110	189	94	182	124
Lees Corner Rd.	83	162	70	163	121
Galesbury Lane	71	121	96	124	97
Chantilly Road	110	98	93	91	46
Metrotech Drive	111	94	76	92	66
Centreville Rd.	106	76	74	77	91
Sullyfield Circle	107	83	83	83	83

Note: Offsets calculated from the beginning of green of Phase 2

Genetic Algorithm Optimization

A GA was linked to the calibrated VISSIM model of Lee-Jackson Memorial Highway to optimize offsets. A weighted average of travel times, eastbound and westbound, was used as the fitness value for the GA. Offset optimization was performed twice using the GA. The first approach used four VISSIM runs per timing plan to assign fitness values to a set of offsets. Figure 9 shows the convergence of the GA. As the number of generations increase, the weighted average travel time approaches the minimum travel time. The final fitness value produced a weighted average travel time of 538.8 seconds.

The second GA optimization used eight VISSIM runs per set of offsets to determine a fitness value. This approach was considered because of the variability in VISSIM travel times. Because of the variability in travel times, using four runs to assign fitness values may not be adequate. This small sample could quite easily be at either extreme of the distribution, thus assigning a fitness value that is either too large or too small. Naturally increasing the number of runs would help but is time-consuming when using the GA. Figure 10 shows the convergence of the GA using eight runs to determine fitness values. The minimum average travel time was 528.8 seconds, 10 seconds less than when four runs were used to determine fitness values. Table 4 shows the optimized offsets from the GA optimization.

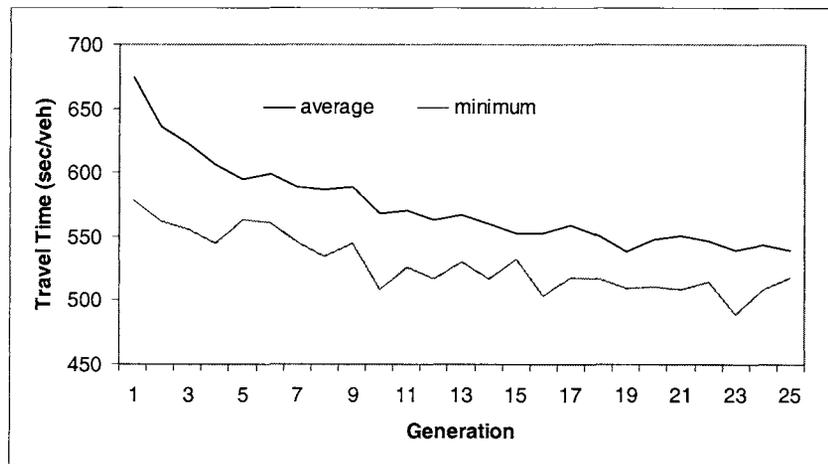


Figure 9. GA Travel Time Convergence (4 VISSIM runs per timing plan)

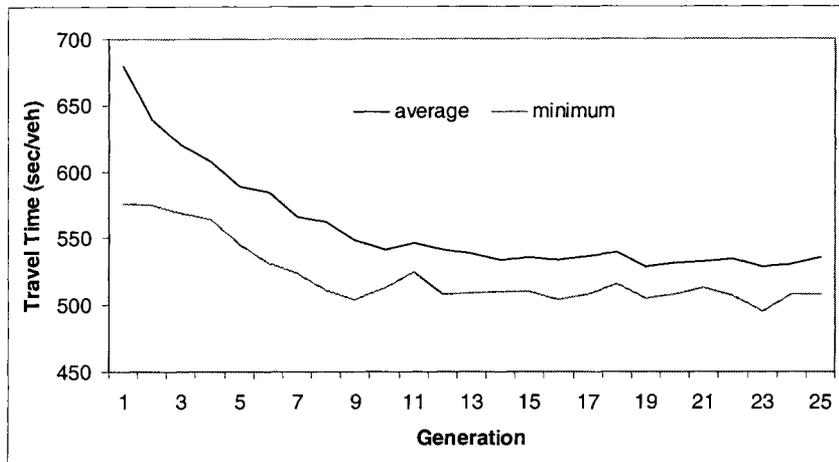


Figure 10. GA Travel Time Convergence (8 VISSIM runs per timing plan)

Signal Timing Plan Evaluation

Five timing plans were under evaluation in this study: (1) the base timing plan, (2) the field timing plan, (3) the GA optimized timing plan, (4) the Synchro optimized timing plan, and (5) the TRANSYT-7F optimized timing plan. Each plan was evaluated using VISSIM. The offset values for the five timing plans are shown in Table 4. The base timing plan was taken from VDOT’s Synchro file. This timing plan was used by VDOT in the field prior to its current timing plan. The timing plan has a cycle length of 180 seconds. VDOT recently implemented its current timing plan, denoted as the field timing plan in this report. The cycle length of this timing plan is 190 seconds and it contains different green splits than the base timing plan. The three optimized timing plans have the same cycle length and green splits as the field timing plan but contain new offset values.

Two measures of performance were used to evaluate the five timing plans. The first was travel time on Lee-Jackson Memorial Highway. The weighted average of travel times for westbound and eastbound traffic was used. Travel time collection points were set at the beginning and end of the test network at the intersections of Route 50 and Sully Road and the Fairfax County Parkway. Travel time was collected from VISSIM through *.rsr files for each run. The *.rsr files contain every completed travel time measurement event in chronological order. Because multiple runs were conducted, resulting in 100 *.rsr files, a Rexx program was created to ease the computational burden. The Rexx program calculated the mean and standard deviation for each of the 100 runs and outputted a summary file containing eastbound and westbound throughputs, travel times, and standard deviations.

The second performance measure was the total system delay. Unfortunately, VISSIM does not calculate total system delay as one of its outputs. In order to find the total system delay, multiple delay collection points were created. Delay segments are similar to travel time collection points. Two points on the test network are chosen, and delay is determined as the difference between actual travel time and the travel time if the vehicle was driving unobstructed at its desired speed. In order to determine the total system delay, delay segments were set up for

every possible combination of vehicle inputs and outputs in the network. For example, for every place a vehicle entered the network, a delay segment had to be created for every combination of where the vehicle might exit the network. The total system delay was determined as a weighted average of all of the delays. The total system delay was obtained through *.vlz files from VISSIM.

VDOT's former and current timing plans were evaluated in VISSIM to determine if their new timing plan (field) performed better than their former timing plan (base). The mean travel time on Lee-Jackson Memorial Highway decreased from 625.9 to 518.9 seconds, a 17.1% reduction. The improvement in travel times can better be seen using the histogram depicted in Figure 11. The field histogram is shifted to the left, or toward shorter travel times. The field timing plan also produces a tighter distribution than the base timing plan. The tightness of the distribution means that the majority of the VISSIM runs produced travel times close to the mean. The base timing plan shows larger variations in travel times.

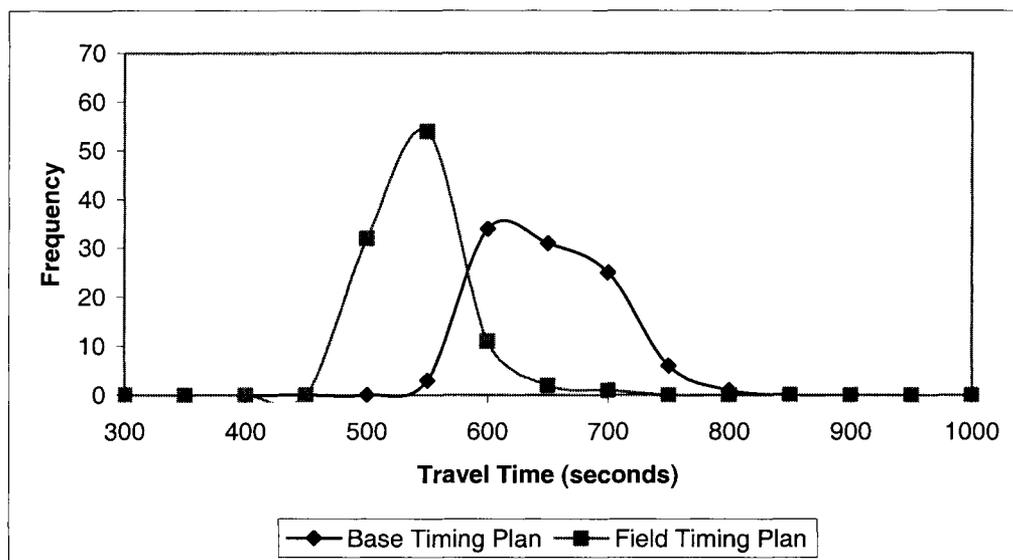


Figure 11. Base and Field: Weighted Average Travel Time Histograms

The total system delay was also improved. The mean total system delay was decreased from 248.0 seconds per vehicle with the base timing plan to 157.2 seconds with the field timing plan, a 36.6% reduction. Evaluation of the two VDOT timing plans shows considerable improvement. Travel times and total system delay are improved with the use of the current timing plan.

Although the current VDOT timing plan outperformed its predecessor, optimization tools were evaluated to determine if they could improve upon the current timing plan. The process used to evaluate the three optimized timing plans was the same as the one used to evaluate the two VDOT timing plans. Travel times on Lee-Jackson Memorial Highway and total system delay were used for evaluation. The three optimized timing plans were compared to each other

as well as to the field timing plan in order to determine which of the four timing plans worked best.

Table 5 shows the weighted average of travel times for 100 runs in VISSIM. As seen in the table, it appears that there is no significant improvement of the mean field travel time of 518.9 seconds. Figure 12 shows the travel time histograms. The field, Synchro optimized, and GA optimized timing plans are strikingly similar. Their mean travel times are also similar. The TRANSYT-7F optimized timing plan is shifted to the right and shows a much larger mean travel time (555.6 seconds) than the other three timing plans.

Table 5. VISSIM Multiple Run Travel Time Results

Timing Plan	Mean Travel Time (sec)	Median Travel Time (sec)	Standard Deviation
Base	625.9	621.2	136.9
Field	518.9	512.5	72.7
Synchro	512.3	511.0	71.9
TRANSYT	555.6	549.7	78.9
GA	520.1	515.4	67.6

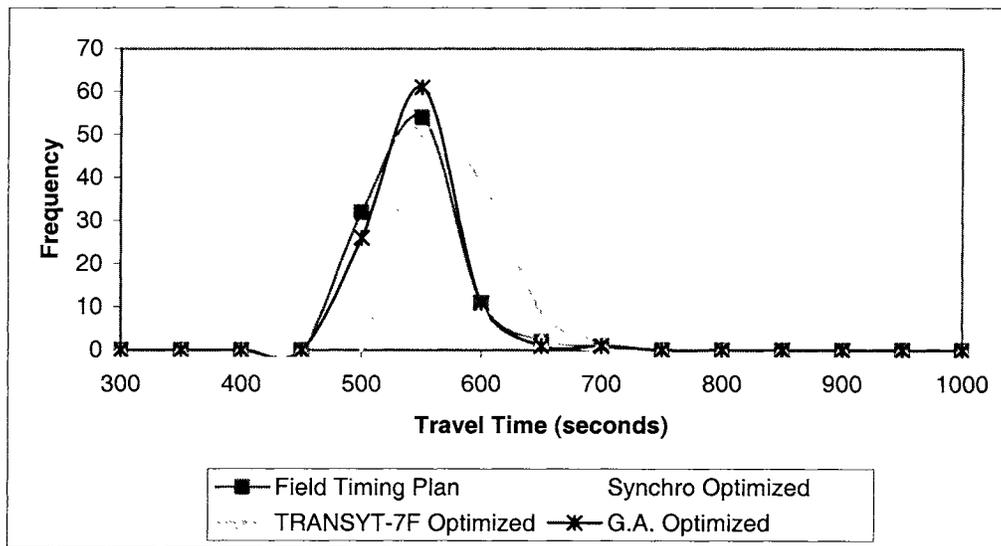


Figure 12. Optimized Timing Plans: Weighted Average Travel Time Histograms

Because the field, Synchro optimized, and GA optimized timing plans gave similar travel time results, a t test was conducted to test if the mean travel times of one timing plan were equal to the mean travel time of another timing plan. Null and alternative hypotheses were set for each combination. For each case, the null hypothesis was that the two data sets were equal. The alternative hypothesis was that the data sets were unequal. P values were determined for each case. The p value provides an objective measure of the strength of evidence the data supply in

favor of the null hypothesis. It is the probability of getting a result as extreme or more extreme than the one observed if the proposed null hypothesis is correct.

Table 6 shows the p values obtained from the statistical test. The table shows the field and GA optimized timing plans to be statistically the same (p value of 0.8). The Synchro and field timing plans also showed a correlation, but not as high. Based on these results, one can conclude that there is no significant difference between the field travel times and the Synchro and GA optimized timing plans.

Table 6. T Test Results: P Values

	Null Hypothesis: H₀	Alternative Hypothesis: H₁	P Value	Decision
1	Base = Field	Base ≠ Field	0.00	Reject H ₀
2	Base = Synchro	Base ≠ Synchro	0.00	Reject H ₀
3	Base = TRANSYT	Base ≠ TRANSYT	0.00	Reject H ₀
4	Base = GA	Base ≠ GA	0.00	Reject H ₀
5	Field = Synchro	Field ≠ Synchro	0.19	Accept H ₀
6	Field = TRANSYT	Field ≠ TRANSYT	0.00	Reject H ₀
7	Field = GA	Field ≠ GA	0.80	Accept H ₀
8	Synchro = TRANSYT	Synchro ≠ TRANSYT	0.00	Reject H ₀
9	Synchro = GA	Synchro ≠ GA	0.09	Accept H ₀
10	TRANSYT = GA	TRANSYT ≠ GA	0.00	Reject H ₀

The total system delays of the three optimized timing plans are depicted in Table 7. The field and Synchro optimized timing plans showed similar results with mean delays of 157.2 seconds per vehicle and 156.7 seconds per vehicle, respectively. The TRANSYT-7F timing plan produced slightly higher delays. Although the GA optimized timing plan gave a similar travel time as field and Synchro, the total system delay for this timing plan was significantly higher with a mean of 245.5 seconds per vehicle. The observations of a few selected simulations indicate that an occasional left-turn overflow (i.e., when heavy left-turn vehicles arrive at the turn bay, at times, left-turn vehicles block through vehicles).

Table 7. VISSIM Multiple Runs Average Total System Delay per Vehicle

Timing Plan	Average Total System Delay (sec/veh)		
	Mean	Median	Std. Dev.
Base	248.0	245.2	28.4
Field	157.2	159.4	13.2
Synchro	156.7	148.3	18.5
TRANSYT	174.6	172.1	15.4
GA	245.5	240.8	29.6

Even though the 100 VISSIM simulations can adequately simulate the day-to-day variations in traffic patterns and driver behavior, the mean number of arrivals at the external nodes remains constant (Cowlshaw, 1990). In reality, these mean arrival inputs will themselves

change over time. Moreover, the estimates of these mean rates are based on traffic counts collected by manual observers and are subject to considerable error. Therefore, the responses of the timing plans to changes in these mean rates were evaluated.

Substantially wide ranges of changes were explored: ($\pm 15\%$ from the base demands at each entry node). Since there were 10 major external input demands (volumes > 1,000 vph), the number of possible demand patterns that can be tested is quite large if only integer percentages are used. An efficient sampling method using a Latin hypercube design was applied to the problem. This is a stratified sampling technique where the input variable distributions are divided into equal probability intervals (Park et al., 2001). It can be considered a deterministic version of a Monte Carlo simulation, but one that can maximally cover the design surface with near zero correlations among input parameters. It requires far fewer samples than simple Monte Carlo methods (Iman and Helton, 1988). A detailed algorithm can be found in McKay et al. (1979).

A total of 50 demand combinations were created from the Latin hypercube design algorithm. Four simulations were made for each demand combination, resulting in 200 VISSIM runs. Demand variations were explored for all five timing plans. Travel times on Lee-Jackson Memorial Highway were evaluated, and the results can be seen in Table 8. Results show that the fluctuations in demand had little influence on the performance of the timing plans. Mean travel times for the field, Synchro optimized, TRANSYT-7F optimized, and GA optimized timing plans were similar for the base demand and varying demand conditions. The base timing plan is the only timing plan influenced by these demand fluctuations, but it resulted in a better mean travel time for varying demand conditions. Overall, the demand analysis shows that the current and optimized timing plans perform well under fluctuating demands.

Table 8. Varying Demand Travel Time Comparison

Timing Plan	Base Demand (sec)			Varying Demand (sec)		
	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.
Base Condition	625.9	621.9	136.9	585.5	582.0	64.7
Field Condition	518.9	512.5	72.7	524.1	521.2	55.7
Synchro Optimized	512.3	511.0	71.9	523.6	527.1	55.9
TRANSYT Optimized	555.6	549.7	78.9	563.0	565.3	51.2
GA Optimized	520.1	515.4	67.6	532.7	532.9	54.7

CONCLUSIONS

Statistical testing and visualization are the most critical aspects of the simulation model calibration and validation. The statistical testing dealt with when to claim the calibrated model is “equal” to the field data, and the visualization dealt with if the animation looks “real” as seen in the field.

The statistical testing is due to the variability in simulation runs. All multiple runs were not statistically equal to the field distribution. In other words, the simulation output passed the statistical test at different percentiles for each parameter set. Given that the field data are just one realization of an infinite stochastic process, this seems natural. Thus, individual runs may not be applicable for statistical testing. Instead, the simulation results that are most close to the field data can be used. If the data pass this test, it ensures that the field data were represented at least once in the simulation model. In addition, the percentile of field average value at the distribution of the simulation output can be used to determine how the simulation represents field condition.

The importance of visualization cannot be overstated. Although obtaining measures of performance from the simulation close to those observed in the field is important, if the animations are not realistic, the model should not be considered calibrated. The purpose of microscopic simulation models is to represent the real world as closely as possible. Simulation models that generate behavior not exhibited in the field are unrealistic. Thus, parameter sets producing unrealistic simulations should not be considered.

Evaluation of timing plans revealed a significant benefit of using the current VDOT timing plan rather than the one previously used. The new timing plan resulted in a 17.1% reduction in travel time on Lee-Jackson Memorial Highway and a 36.6% reduction in total system delay. It is therefore concluded that VDOT’s current methodology for setting timing plans works well in improving performance.

This 17.1% travel time saving translates to 53,054 hours of time saved per year for all vehicles traveling on Lee-Jackson Memorial Highway in the afternoon peak hour. The delay reduction saves a combined 100,061 hours per year for afternoon peak hour for all vehicles in the system. Delay savings were calculated by taking the difference of average delay per vehicle for the base and field conditions. This difference, which was the time savings for the weekday afternoon peak hour, was multiplied by the total number of vehicles in the system. The result was multiplied by the number of weekdays in a year to determine the total annual cost savings. Assuming time savings can be given a monetary value, the user delay savings is roughly \$515,317. This value is based on the assumption that each hour of time saved is equivalent to Virginia’s minimum wage of \$5.15 per hour.

It should be noted that the time savings and benefits stated in this report are based on the demand conditions collected from the field and used in the simulation model. Even though different demand conditions may result in different time savings and delay reductions, it is expected that the implementation of the new timing plan will result in time savings and delay reductions.

The timing plans from three optimization tools were also evaluated in VISSIM to determine if the performance of the current timing plan on Lee-Jackson Memorial Highway could be improved upon. The GA, Synchro, and TRANSYT-7F optimized timing plans did not yield significant benefits over those for VDOT's current timing plan.

RECOMMENDATIONS

1. Evaluation of the test site shows that the current VDOT signal optimization procedure significantly improves the performance of network operations. Thus, it is recommended that VDOT continue to use its procedure for developing new timing plans.
2. Although the current (field) timing plan, which was updated in April 2001, performs well, it is recommended that (1) VDOT evaluate its existing timing plan regularly in an effort to keep the plan up to date, and (2) VTRC conduct a follow-up study to identify when the current signal timing plan should be re-optimized. The timing for this research can be determined by monitoring the trend of traffic patterns, including the changes of traffic demand, turning percentages, vehicle mix, etc., and conducting a cost-benefit analysis between the existing and new timing plans.
3. As this study substantiates that traffic signal optimization is an effective way of improving urban network mobility, it is recommended that VDOT allocate more time and effort toward updating traffic signal timing plans in Northern Virginia as well as the rest of the state.
4. In this study, the proposed simulation calibration and validation procedure used only a single day of field data and two measures of performance. It is recommended that VTRC conduct further research that uses multiple days of field data in order to account for variability and validate the proposed procedure with other test networks. Such research is underway.

REFERENCES

- Beasley, D., D. R. Bull, and R. R. Martin. An Overview of Genetic Algorithms: Part I, Fundamentals. *University Computing*, Vol. 15, No. 2, 1993, pp. 58-69.
- Cheu, R., X. Jin, K. Ng, Y. Ng, and D. Srinivasan. Calibration of FRESIM for Singapore Expressway Using Genetic Algorithm. *Journal of Transportation Engineering*, November/December 1998.
- Cowlshaw, M. F. *The REXX Language*, Second Edition. Prentice-Hall, Englewood Cliffs, NJ, 1990.
- CORSIM User's Manual*. FHWA, U.S. Department of Transportation, Office of Safety and Traffic Operations, McLean, VA, 1997.

- Goldberg, D.E. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Publishing Co., Inc., Reading, MA, 1989.
- Hellinga, B. R. Requirements for the Calibration of Traffic Simulation Models. Department of Civil Engineering, University of Waterloo. <http://www.civil.uwaterloo.ca/b.PDF>. Last viewed March 2002.
- Iman, R. L., and J. Helton. An Investigation of Uncertainty and Sensitivity Analysis Techniques for Computer Models. *Risk Analysis*, Vol. 4, 1988, pp. 71-90.
- Lee, D., and X. Yang. Parameter Optimization for PARAMICS Using a Genetic Algorithm. Paper presented at the *6th International Conference on Applications on Advanced Technologies in Transportation Engineering*, June 2000, Singapore.
- McKay, M. D., R. J. Beckman, and W. J. Conover, A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, Vol. 21, 1979, pp. 239-245.
- Milam, R. T. Recommended Guidelines for the Calibration and Validation of Traffic Simulation Models. Fehr & Peers Associates, Inc. http://www.fehrandpeers.com/publications/traff_simulation_m.html. Last viewed March 2002.
- Paracha, J. Software Choices for Intersection Signals. *Better Roads*. Vol. 69, No. 9. September 1999, pp. 25-27.
- 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, DC, 1999, pp. 133-142.
- Park, B., N. M. Roupail, and J. Sacks. Assessment of A Stochastic Signal Optimization Method Using Microsimulation. *Transportation Research Record 1748*, TRB, National Research Council, Washington, DC, 2001, pp. 40-45.
- Roupail, N. M., B. Park, and J. Sacks. Direct Signal Timing Optimization: Strategy Development and Results. *In XI Pan American Conference in Traffic and Transportation Engineering*, Brazil, November 19–23, 2000.
- Sacks, J., N. Roupail, B. Park, and P. Thakuriah. Statistically Based Validation of Computer Simulation Models in Traffic Operations and Management. *Journal of Transportation and Statistics*, 2003.
- Synchro 4.0 User's Guide*, Trafficware, Albany, CA, 1999.
- TRANSYT-7F's User's Manual*. Transportation Research Center, University of Florida, Gainesville, FL., March 1998.

VISSIM Version 3.6 Manual. PTV Planug Transport Verkehr AG. Innovative Transportation Concepts LLC, Karlsruhe, Germany, December 2001.

Yang, X. K. Comparison among Computer Packages in Providing Timing Plans for Iowa Arterial in Lawrence Kansas. *Journal of Transportation Engineering*, Vol. 127, No. 4, July/August 2001.