2008

The Calibration And Verification Of Simulation Models For Toll Plazas

Christopher Russo
University of Central Florida

THE CALIBRATION AND VERIFICATION OF SIMULATION MODELS FOR TOLL PLAZAS

by

CHRISTOPHER S. RUSSO
B.S. University of Central Florida, 2007

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Civil and Environmental Engineering
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Fall Term
2008

ABSTRACT

A great deal of research has been conducted on Central Florida toll roads to better understand the characteristics of the tolling operation. In this paper, the development and calibration of a toll plaza simulation models will be analyzed using two simulation programs varying mostly in their modeling theory. The two models utilized are, SHAKER, a deterministic queuing model for vehicles utilizing toll collection facilities, and VISSIM, a globally popular stochastic simulation software. The benefits of simulation models leads to the purpose of this thesis, which is to examine the effectiveness of two toll modeling programs that are similar in purpose but vary in approach and methodology. Both SHAKER and VISSIM toll plaza models have the potential to work as a tool that can estimate the maximum throughput and capacity of toll plazas. Major operational benefits resulting from developing these models are to simulate and evaluate how traffic conditions will change when demand increases, when and if queues increase when a lane is closed due to maintenance or construction, the impact of constructing additional lanes, or determining whether or not the best lane type configuration is currently implemented.

Traffic operations at toll plazas include both deterministic and stochastic elements. Therefore, it is unknown which modeling philosophy of the two is more sufficient for simulating traffic operations at toll plazas and there is no definite measure of pre-determining if one of these methods is superior to the other. SHAKER is a deterministic queuing model based on classical physics equations that determines a plaza's maximum hourly throughput by assigning vehicles to lanes based on one of the following lane conditions: the shortest queue, the least amount of vehicles, the lane that provides the required service but minimizes the waiting-time-in-the-queue, or the lane with the fastest moving queue. Conversely,

VISSIM models traffic and behaviors stochastically, which means it simulates by assigning vehicle operations based on randomly assigned processes that include probability functions.

To effectively calibrate any model available site data must be used to compare simulation results to for model validity. In an effort to correctly calibrate the SHAKER toll plaza tool and VISSIM model, an extensive field collection procedure was conducted at four Florida Turnpike operated toll facilities located in Central Florida. Each site differed from the others in terms of number of lanes, lane configuration, toll base fee, highway location, traffic demand, and vehicle percentage. The sites chosen for data collection were: the Lake Jesup Mainline Plaza along the Seminole Expressway (SR-417), the Beachline West Expressway Toll Plaza along the SR-528, the Daniel Webster Western Beltway Plaza along SR-429, and the Leesburg Toll Plaza along the Florida Turnpike Mainline SR-91.

From this data, periods of constant queuing were pinpointed and form those periods factors such as demand, throughput, service-time, etc. were extracted. It has been determined that service time, and lane type greatly affects the capacity at a toll plazas. Using the extracted factors, SHAKER was then calibrated and validated to estimate the capacity of the observed toll plazas. The actual calibration and validation procedure is adopted from a widely used method developed by the Federal Highway Administration. To calibrate both SHAKER and VISSIM a base model toll plaza was created. After initial calibration of this model using a specific dataset, it would be validated using an additional data set from another day's worth of traffic. Also, to validate the simulation, the calibration procedure proposed towards the base case was implemented to additional sites to prove the successfulness of constancy in calibration techniques. In order to accurately represent a currently functioning toll plaza the SR-528 Beachline West Toll plaza was used as the base case and additional toll plazas were modeled for validation purposes. Upon completion of calibrating both simulation models, it was demonstrated that the calibration techniques were verified using statistical significance testing.

In addition, a new application was added to SHAKER with the ability to select automatically the best configuration of a toll plaza given that the lane-user remains unchanged. To verify the results of the new SHAKER application VISSIM was coded using the new best configuration geometry and results were compared. The VISSIM simulated capacity matched that of the SHAKER application and thus the best configuration application was verified for approval.

Upon completion of calibration of the two simulation models it is determined that each of the two software are successful in modeling toll plaza capacity and queuing. As expected, each simulation model does possess benefits over the other in terms of set up time, analysis reporting time, and practicality of results. The SHAKER model setup takes mere seconds in order to create a network and input vehicle, another few seconds to calibrate driving parameters, and roughly 10 additional seconds to report analysis. Conversely, setting up the VISSIM model, even for the most experienced user, can take several hours and the report analysis time can take several more hours as it is dependant on the number of required simulation runs and complexity of the network. VISSIM is most beneficial by the fact that its modeling allows for driver variability while SHAKER assumes steady state equilibrium amongst lane choice and queuing. This creates a more realistic condition to observed traffic patterns. Even though differences are prevalent, it is important that in each simulation model the capacity is accurately simulated and each can be used to benefit operational situations related to toll plaza traffic conditions.

Dedicated to

The first of my educators,

the builders of the foundation from which I have grown,

my Mother and Father.

and to

the origin and possessor of all my affection,

the one who brings out the very best in me,

the love of my life,

Kayleen

For their guidance is unparalleled,

Support is outstanding,

and

love is everlasting

ACKNOWLEDGMENTS

I would like to extend my gratitude to everyone who even helped me along my way. Especially I would like to thank Dr. Essam Radwan for providing me the opportunity to further my education at the University of Central Florida. His guidance, support, and trust created a comfortable work environment to conduct my research. Also, I would like to acknowledge Dr. Al-Deek for serving on my review committee and for his interest in my learning. I would like to specifically thank Dr. Vinayak Dixit for all his many lessons in theory and for mostly for getting me started on my research. Also, I would like to specifically extend my gratitude to my coworkers and friends for all their support and suggestions, and for making the past 2 years most enjoyable - Vinayak, Shankar, Rami, Dr. Yan, Noor, Ali, Olivia, Ravi, Jeanne, Jeremy, Abhishek, Jason, Taylor, and John. I would also like to acknowledge my fiancé Kayleen for giving me the freedom to dedicate the necessary time to complete my research and for her continued support throughout the many long days and nights needed for all the model calibration runs.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AAWT	Average Annual Weekday Traffic
ACM	Automatic Coin Machine
AVI	Automatic Vehicle Identification
ETC	Electronic Toll Collection
HCM	Highway Capacity Manual
LOS	Level of Service
MOE	Measure of Effectiveness
NQMT	No Queue Maximum Throughput
OOCEA	Orlando Orange County Expressway Authority
SAIC	Science Application and International Corporation
TPASS	Toll Plaza Animation/Simulation System
VPH	Vehicles per hour

1. INTRODUCTION

The major highways that serve the public in Central Florida are Interstate-4, State Road (SR) 408, SR 417, SR 528, SR 429, and Florida's Turnpike. Of these major highways, all but I-4 are toll restricted facilities. The Florida Turnpike Authority states that "tolls are the most cost-effective way to directly link user fees to specific roads. These roads are self-supporting; freeing highway tax money for other needed road projects"(*1*). Through their repetitive implementation it is evident that toll collection roadways play an important role in traffic throughout Central Florida. Toll collecting roads are unique in operation because they force a percentage of users to make periodic stops to pay tolls while the other users own electronic toll collector (ETC) devices that allow them to pay without making a complete stop and go movement. At tolling plazas each payment type has a specific lane that the user is to use to correctly pay for his or her toll. Common lane payment types along Central Florida toll roads consist of manually collected, Automatic Coin Machine (ACM), reduced speed ETC, and high speed ETC. A great deal research has been conducted on Central Florida toll roads to better understand the characteristics of the tolling operation. A synopsis of previously conducted research will be discussed in the literature review section of this document. In addition, researchers have developed programs and simulation models to aid in understanding tolling elements such as the throughput, capacity, level of service, optimum lane configuration, etc. In this thesis, the development and calibration of a toll plaza model will be analyzed using SHAKER, a queuing model for vehicles utilizing toll collection facilities, and VISSIM, a globally distributed stochastic simulation software. The significant difference between these two models is in their modeling logic. SHAKER models vehicle operations deterministically, meaning that the model assumes no variability in driver-vehicle characteristics and all interactions are defined by exact relationships, such as mathematical, statistical, or logical. Conversely, VISSIM models stochastically, which means it simulates by assigning vehicle operations based on randomly assigned processes that include probability functions. Traffic operations at toll plazas include both deterministic and stochastic elements. It is unknown which modeling philosophy of the two is more sufficient for simulating traffic operations at toll plazas and there is no definite measure of pre-determining if one of these methods is superior to the other. SHAKER modeling is deterministic; therefore, it has been proposed that it generally has a shorter computation time when compared to the microscopic simulation models. However, it is also anticipated that the stochastic methodology, like that of VISSIM, allows for better representation of the variability amongst drivers *(2)*. These implications, and many other benefits and complications of each program, will also be addressed in this study. To summarize, this research is to contain an extensive review of the development and calibration of two differing toll plaza simulation models and to study their uses and limitations.

2. PROBLEM STATEMENT

In 2006 The Orlando Orange County Expressway Authority reported that their most traveled toll plaza exceeds an Average Annual Weekday Traffic (AAWT) greater then 125,000 vehicles *(3)*. The Expressway Authority has also estimated that during peak conditions 60-70% of users make use of their electronics toll collector to pay the toll. Which means 30-40% of the traffic is temporarily interrupted in order to manually pay the toll. It is expected that traffic complications and backups will arise when several drivers experience this forced bottleneck. This relatively large number of drivers using the manual and automatic lanes and the increasing bottlenecks in the latter during peak hours motivates engineers and planners to look into possible countermeasures to mitigate delays, queues and high concentration of automobile emissions at toll facilities *(4)*. One such opportunity resulting from this need is the conception of traffic modeling and simulation.

In the past it was difficult for engineers and planners to determine how particular lane configurations would affect the traffic conditions without physically testing the configurations. Site configuration testing in the field would prove costly and varying toll scenarios may be confusing, thus, dangerous to the motorists. Also, employing inefficient toll plaza configurations just for research can cause costly delays to the user. The impact of delays reaches further then just driver frustrations. Delays also cause queues and Klodzinski et al. *(4)* points out that queuing toll plazas are a point source for high concentrations of dangerous automobile emissions. To counter these complications, simulation models are often used to test alternative designs without the cost of real life testing. The benefits of simulation models leads to the purpose of this thesis, which is to examine the effectiveness of two toll modeling programs that are similar in purpose but vary in approach and methodology. Both SHAKER and VISSIM toll plaza models have the potential to work as a tool that can estimate the maximum throughput and capacity of toll plazas. There is no research on whether the SHAKER or VISSIM model would better represent toll facility operations, therefore lies the need for the research conducted in this thesis. It is unknown which modeling philosophy of the two is more sufficient for simulating traffic operations at toll plazas and there is no definite measure of pre-determining if one of these methods is superior to the other. Toll plaza operations include both deterministic and stochastic elements. According to Astarita drivers' behavior variability is stochastic by nature, while other elements of the toll booth traffic are generally deterministic *(5)*. Therefore, upon completion of the two models, the study results will be compared so as to investigate the effectiveness and efficiency of each program's ability to replicate real life toll facility traffic conditions. Major operational benefits resulting from developing these models are to simulate and evaluate how traffic conditions will change when demand increases, when and if queues increase when a lane is closed due to maintenance or construction, the impact of constructing additional lanes, or determining whether or not the best lane type configuration is currently implemented.

As the population in the Central Florida area continues to grow, simulation is one tool which assists transportation planners in predicting and preparing for the impacts that continued growth will have on the traffic system. In summary, the objective of this study is to develop, calibrate, and analyze two toll plaza simulation models, whose theory differs by their magnitude of meticulousness, in order to reach a better understanding of toll plaza capacity. After a better understanding of toll plazas' operation/ processing time is established, results may be used for future simulation model calibration. Moreover, planners, designers, and toll road agencies may take necessary measures to enhance processing times, capacity, and operations at toll plazas.

3. RESEARCH OBJECTIVES

The objectives and tasks of this research are, but not limited to:

1. Using collected 2007-2008 toll plaza data develop default values for driver characteristics to input into simulation models.
2. Develop an isolated microscopic toll plaza model in VISSIM
3. Using data analysis from step one, calibrate the VISSIM model and SHAKER program to replicate observed capacity and throughput values.
4. Verify the results and performance of the VISSIM and SHAKER model for existing conditions and hypothetical scenarios, such as a lane closure or increased demand.
5. Develop an evaluation on the performance ability of each model and evaluate the effectiveness of each program to complete the task at hand.

4. LITERATURE REVIEW

This literature review covers multiple aspects of toll plaza operations. The first section covers previous exclusive toll plaza models and calibration techniques that have been developed to better understand toll plaza operations through the use of simulation. The second portion consists of a discussion regarding plaza capacity studies that have been conducted and their contribution to the field. Next is two sections based primarily on the SHAKER model. The first of which covers the methodology that SHAKER uses to calculate the throughput and capacity. The second segment gives a brief overview on how to use the SHAKER software and discusses some of the software's potential.

4.1 Toll Plaza Simulation Models and Calibration

4.1.1 Microsimulation Model Calibration

Traffic simulation software has become increasingly popular as a traffic analysis tool used in transportation analyses. One reason for the increase in simulation use is the need to model and analyze the operation of complex transportation systems under congested conditions. There are microscopic and macroscopic traffic flow models for simulation. Microscopic simulation models use numerous independent parameters to replicate traffic control operation and traffic flow characteristics by modeling the state (position, velocity, and sometimes additional information) of every vehicle. In contrast, macroscopic models define variables of state in terms of averages, such as the average speed, volume, and density that describe the system or parts of the system. Before using macro- and micro- simulation models they inevitably must first be calibrated before they can accurately estimate traffic conditions. 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 *(6)*. Calibration is necessary for these models because no single model can be expected to be equally accurate for all possible traffic conditions *(7)*. In general, microscopic simulation models contain default values for each variable, but they also suggest that users to input a range of values for the parameters that better suit their unique condition. Changing these parameters for calibration should only be done so when based on field measured conditions and all changes must be justified and defensible by the user *(8)*. The difficulty in basing calibration on field observations is that many of the parameters used in simulation models are difficult or sometimes impossible to measure in the field, yet they can substantially impact on the model's performance. Park and Schneeberger expressed difficulty in observing particular variables, such as: start-up lost time, queue discharge rate, car-following sensitivity factors, time to complete lane change, acceptable gaps, and driver's familiarity with the network *(8)*. Hellinga *(9)* describes the basic guidelines of a seven component calibration, but provides no direct procedure for conducting calibration and validation. The steps for calibration are listed in order as:

i) defining study goals and objectives

ii) determining required field data

iii) choosing measures of performance

iv) establishing evaluation criteria

v) network representation

vi) driver routing behavior

vii) evaluation of model outputs

Calibration is often referred to as an art, an inexact science, so not all calibration steps will have specific technical significance. That is why Park and Schneeberger also indicate that when using microscopic simulation models the importance of user visualization cannot be over emphasized. The goal of the microscopic simulation model is to represent field conditions as closely as possible. Therefore by nature a model cannot be considered calibrated if the animations are not visually realistic. Even if a parameter set produces statistically acceptable results but the animations are not realistic then the model cannot be considered calibrated *(8)*.

4.1.2 Exclusive Toll Plaza Models

One of the first animated toll plaza simulation software, which was designed in 1992, was the Toll Plaza Animation/Simulation System (TPASS) (10). TPASS is a discrete-event toll plaza model developed by Science Application and International Corporation (SAIC). TPASS allows the user to experiment with various toll plaza configurations and traffic characteristics in order to determine the resulting queuing, wait times, and toll revenue. This models combination of simulation and animation provides the user to make quantitative comparisons of experimental data sets with visual friendly animations presenting information to aide in the evaluation of the simulated scenario. It is stated that the most useful output parameter for calibration of the TPASS model is the total number of vehicles in queue *(10)*.

Toll Plaza Simulation Model (TPSIM) was developed at the University of Central Florida with the purpose to simulate toll plaza operation. TPSIM is a stochastic discrete-event microscopic simulation model that divides the toll plaza into three zones for analysis. The three zones are the approach zone, the transition zone, and the toll zone. TPSIM is a stochastic object oriented discrete-event microscopic simulation model that was coded using Microsoft Visual

Basic 6.0 and interfaces with Windows98/NT. Toll plazas with up to 5 approach lanes and up to 10 toll lanes in each direction can be modeled using TPSIM. The model contains algorithms for car-following, lane-changing, and toll-lane selection and provides output for measures of effectiveness (MOE) which include throughput, average queuing delay, maximum queuing delay, and total queuing delay. The TPSIM model was calibrated with data from the Holland East Plaza in Orlando, Florida and validated for use of different toll plaza configurations and ETC lane uses *(11)*. Klodzinski et al. verified that the TPSIM model has the capability to accordantly model any toll plaza scenario and is transferable to other toll plazas with different configurations using the following measures of effectiveness: throughput, average queuing delay, maximum queuing delay, and total queuing delay *(12)*. For calibration it was found that the service time was determined to have the most significant impact on the simulation model *(12)*. Klodzinski also states that to successfully calibrate and apply TPSIM, calibration data must be chosen carefully. If multiple days are selected for calibration, they must have similar characteristics (plaza configuration) and have a service time that is not significantly statistically different *(13)*. Similar to the SHAKER model the TPSIM model is limited to simulating isolated toll plazas. It can not be used to assess an entire network consisting of several toll plazas and/or intermediate sections between each.

TOLLSIM, a stochastic simulation model developed by Wilbur Smith Associates, is used primarily to analyze the toll operation at the toll plaza approach. To be effective TOLLSIM is programmed with traffic data and lane type configuration, ramp approaches and the storage length of each lane are required. The model produces simulation analysis results in both graphical and numerical format and lists a number of measures of effectiveness, such as delay per lane, overall delay, and queue length. However, TOLLSIM does not have the capability to analyze the traffic operation downstream from the toll plaza *(14)*.

4.1.3 Adaptation of Traffic Simulation Models for Toll Plazas

In addition to using exclusive toll plaza models, microscopic simulation programs have been used to develop toll plaza models. Using Paramics, Nezamuddin *(15)* modeled a toll plaza network and Ozbay *(16)* modeled an integrated freeway and toll plaza. Nezamuddin's capacity and delay models were based on calibrating the simulation with five key parameters: queue gap distance, queuing speed, mean target headway, mean reaction time, and minimum gap *(15)*. Nezamuddin determined if his simulation results were within an acceptable range of values using the GEH statistic. The GEH Statistic is used to compare observed volumes with those obtained from simulation results. The GEH statistic is a modified Chi-squared statistic that incorporates both relative and absolute differences. Ozbay et al modeled a non-mainline toll plaza and found that the best way to fine tune calibration was with trial and error method, and fine tuning with location of sign posting *(16)*.

CORSIM is a simulation model regularly used for highway corridors but it can be used in conjunction with TOLLSIM to analyze the traffic operations on ramps and local roadway systems downstream from the toll plaza. Although CORSIM can theoretically simulate operations at a toll plaza it is not a straightforward modeling effort because it requires declaration of inherent complicated operational data typically found at toll plaza *(14)*.

VISSIM is another wide-ranging simulation model that can be adapted for toll plaza performance analysis. For toll model development and calibration it requires the same input data as listed above under TOLLSIM. Similar to most traffic simulation programs, calibration is required to match existing field observed toll operations. One advantage of VISSIM is that the user can seamlessly analyze the interactions between the highway leading to the plaza, downstream from the plaza, and at the plaza *(14)*.

Ceballos ran queue analysis models at parking exit toll plazas using the VISSIM software *(17)*. Ceballos applied VISSIM because of the capabilities of the software in developing toll plaza simulations based on dynamic assignment of vehicle paths, priority rules, service time distribution, speed reduction zones, and driver behavior *(17)*. Ceballos' contribution to the research comes from his analysis of the difference between using multi-server queuing models and traffic simulation techniques. The study results indicate that multi-server queuing models could be used as an initial, mostly conservative, tool in early stages of planning, but simulation should be used for advance planning, design, operation and management of toll and exit plaza facilities. This research is however limited to applications associated with toll exit plazas at airports.

In VISSIM Park et al. *(18)* calibrated and validated a microscopic simulated freeway work zone network using an eight step procedure developed for calibrating and validating microscopic simulation models. The eight step method primarily focuses on identifying key calibration sets and using genetic algorithms to optimize the parameters used to match field conditions. Their research contributes that genetic algorithm calibration provides a more statistically significant calibrated model than does the best-guessed method or the VISSIM suggested default parameters. However, it is noted by Chitturi that implementing this procedure requires extensive knowledge of numerous microscopic parameters, their ranges, and their significance *(19)*. In addition, this method requires running several hundred simulation runs before the genetic algorithms can determine the optimal parameter set. Chitturi indicates that because of the disproportionate time and resource constraints associated with genetic algorithm design, a regular user of VISSIM may not be able to use this procedure *(19)*.

Using VISSIM Lownes and Machemehl *(20)* studied the sensitivity that the driver behavior parameters have on corridor capacity. VISSIM was calibrated to simulate the US 75/SH-190 interchange, just north of Dallas, Texas. Following initial calibration, researchers studied how capacity was affected when modifying only one parameter at a time for four levels. Statistical analyses were performed after each level to determine if the changes in capacity were statistically significant or not. The contribution of this study is the identification of parameters which could significantly affect the capacity in VISSIM. However, not addressed is the issue of how to choose the values of the parameters that had a significant effect on capacity.

While both macro and micro simulation techniques are reviewed above, it is important to also consider the comparison of the usefulness and effectiveness of both. Festa, et al.*(21)* conducted a comparison between two different motorway traffic models with the purpose to operate a comparative evaluation of potentialities and limits in the two different approaches by applying the models on a large scale motorway network with a complete traffic data base. The two different models used were a model based on microscopic traffic theory and a macroscopic stochastic experimental model. In order to perform a comparison between these two models, they have been applied to simulate the behavior of the real system during the morning of a working day. The time was initially broken down into 15 minute periods to avoid transient periods. The mean relative error, mean square relative error, and mean relative error were used to evaluate the differences between the observed and simulated flow rates. For this experiment the macroscopic approach resulted in smaller errors of all types then did the microscopic method. Other significant findings in the literature are uncovered in terms of effectiveness of the simulation models. In terms model time, the macroscopic approach only few seconds are required, while the microscopic model requires more than 100 per run. This time however depends upon the total number of vehicles on the extension of the network itself and on the number of output required. The length of the computational time makes micro-simulation modeling difficult because the calibration procedure requires a very high number of simulations. In conclusion, Festas contribution to the field is found through the comparisons of the two models. As discovered, the microscopic approach allows tracking of space-time trajectories for every vehicle from its origin to the final destination. It allows better reproduction of the traffic dynamics, but, also it imposes the necessity to calibrate a high number of parameters, causing an increase in computational times. In contrast, the macroscopic approach benefits are that it conducts to an aggregate traffic representation, and it has a very fast execution and low memory occupation *(21)*.

4.1.4 Toll Plaza Capacity Studies

The 2000 Highway Capacity Manual *(22)* defines capacity as the maximum hourly rate at which persons or vehicles can be reasonably expected to traverse a point or a uniform section of a lane or roadway during a given time period under prevailing roadway, traffic, and control conditions *(22)*. The 2000 HCM also states that reasonable expectancy is the basis for defining capacity. Reasonable expectancy meaning that the stated capacity for a given facility is not the absolute maximum flow rate observed at a facility, but a flow rate that can be achieved repeatedly for peak periods of sufficient demand. The HCM uses the concept of Level of Service (LOS) for all kinds of traffic facilities, but still does not provide any standard way to define LOS for toll plazas. Based on field research and data analysis, Klodzinski and Al-Deek *(23)* recommend that delay be the most credible measure of effectiveness to determine the LOS of a toll plaza. Klodzinski and Al-Deek also developed a hierarchy of LOS groups to represent different levels of delay. A vehicle arriving at the toll plaza via no queue represents the best scenario because it only experiences delay caused by the transaction time. The TPSIM computer model was used to verify with 95% accuracy the delay results that were observed in the field and used for LOS development *(24)*.

Aycin developed a manual calculation methodology to determine the capacity, queuing patterns, and delays of toll plazas by considering the approach roadway conditions and traffic demand characteristics. The goal of the research is to improve planners understanding of toll plaza operations and to provide a means of evaluating similar toll plaza simulation results *(25)*.

A review of the literature of toll plaza capacity studies suggests that there is no exact capacity for any type of toll booth lane type. Three major studies, each with different results, are provided to show the uniqueness of the subject.

Pietrzyk conducted a study of multiple types of toll lanes and found the average capacity for the different toll plaza lane types are as follows *(26)*:

- Manned – 350 veh/hr/ln
- Automatic – 500 veh/hr/ln
- Mixed AVI(Automatic Vehicle Identification) – 700 veh/hr/ln
- Dedicated AVI – 1,200 veh/hr/ln
- Express AVI – 1,800 veh/hr/ln

These average capacities for non-dedicated AVI lane types were derived from individual capacity data records provided by toll agencies including the Florida Turnpike, New Jersey

Turnpike, and the Dallas North Tollway. The estimated average capacities for the dedicated and express AVI lanes were based on average speeds and vehicle spacing and headways.

Zarrillo et al. collected field data that showed processing rates for different customer-groups at the Holland East Plaza, located on SR-408 in Orlando Florida, and Interchange 11A, located on the Massachusetts Turnpike 90 *(27)*. The results propose that on OOCEA toll facilities:

- Manual service (M) can process 8.3±0.8 veh/min (498±48 vph)
- Automatic Coin-Machine (ACM) Service lanes (no semi-trucks permitted and no gate present), can process 10.3±0.5 veh/min (618 ± 30 vph)
- Truck Manual (T) service consisting of derives of semi-Trucks can process 2.3+1.3 veh/min (138 ± 78 vph)
- ETC Service (E15) using AVI technology to automatically record the toll amount and drivers are limited to speed limits of 15 mph, can process 15.0±2.0 veh/min (900 ± 120 vph).
- ETC Service (E35) with drivers limited to speed limits of 35 mph, can process 23.0±2.0 veh/min (1,380 ± 120 vph).
- ETC Service (E55) with drivers limited to speed limits of 55 mph, can process 32.0±2.0 veh/min (1,920 ± 120 vph).

Woo and Hoel (28) conducted a study on the Richmond-Petersburg Turnpike in Virginia using synchronized video cameras to determine the service times and capacity of 4 different toll plazas. They determined the service time for trucks ranged from 12.87 to 14.88 seconds for trucks, and from 5.11 to 5.47 seconds for automobiles. They make no distinction between the service time for manned booths versus ACM booths because according to their study ACM service time is shorter or exhibit little or no difference than manned booths. Their capacity study uses an automobile equivalent for trucks that ranges from 2.39 to 2.91 passenger cars per truck. The results of their research are as follows:

- A general toll booth capacity ranges from 650 to 705 passenger cars per hour
- An exact change toll booth without a lifting barrier is between 645 and 665 passenger cars per hour
- An exact change toll booth with a lifting barrier has a capacity of 600 passengers cars per hour.

The dissimilar capacity results published in multiple literatures suggests that there is no one value recommended for all toll plazas and there appears to be no general agreement among traffic engineers as to its precise value. Table 1 below summarizes different toll plaza studies and the capacities discovered in each.

Table 1: Summary of Toll Plaza Capacity Research Results

	Differing Capacities in Research, by Author (veh per hour per lane)		
Lane Type	**Pietrzyk***	**Zarrillo****	**Woo*****
Manned	350	498	675
ACM	500	618	655
Truck	--	138	--
ETC (35 mph)	1200	1380	--
ETC (high speed)	1800	1920	--

* New Jersey Turnpike, Florida Turnpike, and the Dallas North Tollway
**Holland East Plaza, located on SR-408 in Orlando Florida, and Interchange 11A, located on the Massachusetts Turnpike 90
***Richmond-Petersburg Turnpike in Virginia

Other factors that may contribute to a varying capacity are that under light traffic, the toll collector's performance may be reduced due to the lack of need to work fast as the case when pressured with a queue. When toll collectors are under greater pressure from a growing queue, they tend to process transactions faster *(28)*. Service times have a strong influence in the toll plaza capacity. They are important parameters to consider in the design of these facilities. They are also essential information for operational decisions that involve the definition of work shifts and number of opened booths *(29)*. Service time per vehicle is greatly affected by the number of bills and/or coins that must be processed by the toll booth collector or ACM. Astarita et al. suggests that every toll booth type is characterized by its own service time distribution, with an average value and a standard deviation *(5)*. In the same literature it was determined that drivers' behavior variability is stochastic by nature, while other elements of the toll booth traffic are generally deterministic. Manned toll booths charging exact bill amounts tend to have higher capacities than ones that do not *(29)*. ACM booths capacity decreases with the increase of number of coins needed to make payment. Other factors that have the potential to influence the processing time are the experience and pace of the toll collectors, the use of toll gates, the methods of toll collection, and the presence of drivers with exact fee amounts. The Florida Turnpike uses gates on its ACM lanes to indicate when the payment is fully received by the machine. According to Pietrzyk, who compared the Florida Turnpike system to the Tampa Crosstown Expressway, the New Jersey Turnpike, and Dallas North Tollway, the average capacities of ACM lanes with gates are typically reduced by 10 to 20 percent of that of un-gated lanes *(26)*. The traffic distribution can also affect the plaza capacity as manual booths with high truck percentages typically have lower service volumes than those that server primarily automobiles.

Traffic congestion levels also affect the service time per vehicle for the reason that while waiting in queues motorists experience extra time to search for money before they make their final stop in the queue to pay the toll transaction *(28)*. Oliveira also discovered that a joint analysis of the maximum and minimum times model shows that the variability between maximum and minimum service times decreases as flows increase at the toll plazas. Thus, the system becomes more stable at high traffic flows *(29)*. When traffic congestion occurs at plazas with insufficient queue storage lengths there exist the potential for the capacity of the plaza to be drastically reduced. The plaza dimensions and layout upstream of the toll booths is thus another factor that governs the plaza's overall capacity. According to Astarita et al. "When the vehicle arrival rate exceeds the corresponding service rate, slowed vehicles directed towards over-saturated booths (usually the manual ones) can cause a cut-off in the flow of other vehicles, which are destined to a non-congested booth" *(5)*.

While there are numerous studies that focus on the calibration process of simulation models and as many studies that focus on determining the capacity of a toll plaza, there is no evidence in the literature of a study using plaza capacity to calibrate both a deterministic and stochastic toll plaza model with research objectives rooted in the evaluation of the efficiency and exactness of each models ability to match such an extensive field data collection study.

4.2 SHAKER Review

Zarrillo and Schmitt *(2, 30, 31, 32, 33)* completed extensive research in developing the SHAKER tool to simulate vehicle queuing behavior at toll plazas. SHAKER is a deterministic queuing model based on classical physics equations that determines a plaza's maximum hourly throughput by assigning vehicles to lanes based on toll lane queuing conditions *(2)*.

4.2.1 SHAKER Methodology

In SHAKER hourly throughput for a lane under queuing conditions is calculated using the linear equations of motion. However, the lane-percentages or the relative frequency of occurrence must be known and used as input to these equations. Therefore, a method for distributing the approaching traffic into the available lanes is required. Thus, the “shaking” method is incorporated. The “Shaking” process moves around vehicles from one lane to another until a “correct” distribution is established. The determination of the “correct” distribution is based on the stability of an outcome measure, such as hourly throughput, queue length or delay. The “shaking” process has a set of conditions, constraints or rules that must be obeyed; for instance, the model only allows vehicles to be placed, “shook” or queued in a lane in which their category has available service. If there is more than one lane available as a possible choice, then SHAKER uses one of four types of criteria upon which drivers may base their decision:

1. drivers may prefer lanes that have the smallest number of remaining vehicles in the queue
2. drivers may prefer lanes that have the smallest remaining queue length
3. drivers may prefer lanes that have the shortest wait time in the remaining queue
4. drivers may prefer lanes that have the fastest moving remaining queue

Criteria 1 and 2 appear very similar because both are rooted in the queue length but they differ because of the potential varying length of vehicles in the queue. For instance, the difference is better understood in the example of a driver approaching a toll plaza where one lane has three vehicles in one lane and one 18-wheeler in the other. Due to their respective lengths the remaining queue may be similar lengths but there is a different amount of remaining vehicles in each queue. Criteria 3 and 4 are correlated because the calculation for the determination of each is the inverse of the other. To keep results consistent in this research only criterion 1 is used.

The "shaking" process continues as long as the output measure changes in value. Once stability is reached the "shaking" process stops. SHAKER uses steady state equilibrium assumptions to obtain the maximum throughput for a toll plaza and estimate the optimal booth configuration. SHAKER's throughput calibration is accomplished by adjusting the vehicle-properties for five categories so that the model's output for the hourly throughput matches those measured in the field. The five parameters to be adjusted are vehicle length, acceleration, deceleration, driver reaction time, and processing time. After modeling the queuing at toll collection facilities in SHAKER, the throughput of the entire plaza may be predicted.

SHAKER ultimately takes the demand and forces the vehicles to pass through the toll plaza using a basis of equilibrium to find capacity. An equilibrium methodology has some drawbacks that are addressed in the model. For example, to account for vehicle types such as trucks using lane types unconventional to design SHAKER splits up the ETC category into two groups; ETC trucks and non-ETC trucks. The model can more accurately reflect the policy that trucks are prohibited from using the ACM lanes. The SHAKER tool categorizes vehicles by their payment method and lane type they choose. This study has adopted the use of initials that represent the different lane type choices; they are:

E_p = Two wheel vehicle who chooses Electronic Toll Collection lane

E_T = Truck that chooses the Electronic Toll Collection lane

A = Two axel vehicle choosing the Automatic Coin Machine lane

M = Two axel vehicle choosing the manned toll booth lane

T = Truck that chooses the manned toll booth lane

4.2.2 SHAKER Methodology to Calculate Lane Throughput

This section elaborates on the methodology used by SHAKER *(2, 30, 31, 32, 33)* to compute the capacity and throughput of toll plaza's lanes. "SHAKER can determine the number of processed vehicles per unit time at toll collection facilities given the total number of vehicles arriving at the plaza and their traffic characteristics1. This is identical with knowing the number of vehicles of each customer type arriving at the plaza" (2). In this model, car-following theory is used to derive a model for mixed lanes. To describe the SHAKER methodology further it is more convenient to break up the procedure into two parts:

1. The determination of the traffic characteristic in each lane of the plaza.
2. The determination of the maximal throughput of each lane knowing the traffic characteristics in each lane.

Each procedure requires the other as an input before it can be solved. Therefore, the model is an iterative feedback algorithm iterating these two parts. There are numerous equations used to determine the maximum throughput for lanes in which drivers have to stop to pay tolls and lanes in which drivers only slow down to pass as their ETC apparatus takes care of the toll, and a mixed lane used to simulate vehicles that own ETC apparatuses but choose to drive the non-dedicated ETC lanes. In the SHAKER methodology these lanes are designated as pure lanes for stopping vehicles, pure lanes for non-stopping vehicles, and mixed lanes for stopping and non-stopping vehicles. Initially SHAKER finds the probability of finding trains of each vehicle type in each lane type to create the traffic composition. Next SHAKER determines the overall throughput (maximal number of vehicles a lane can process per unit time dependent upon the traffic characteristics in this lane) of each lane by referencing the vehicle characteristics of each vehicle type. Because of the complexity of the formulas used in the SHAKER methodology to determine the lane throughput and capacity the reader is referenced to the well detailed work of Zarrillo and Schmitt for further explanation.

4.2.3 Previously Calibrated Version of SHAKER

During calibration SHAKER uses the single service lane processing rates for different categories, S_M, S_T, S_A, and S_E, given by observed field data values listed in Table 2 *(30)*. In other words, except for the ETC category, input of 100% into SHAKER's throughput equations of any one of SHAKER's categories will result in lane throughput values that are equal to the processing rates listed in Table 2. For the case of the dedicated ETC lanes in which a speed limit of 35 mph is enforced, SHAKER is calibrated using a value of 1698 vph for the processing rate of ETC vehicles that are passenger cars and 1060 vph for ETC vehicles that are not passenger cars.

Table 2: Field Measured Processing Rates for the Traffic Categories used in Calibration

Type of Toll Service	**X**	**S^X (vph)**
Manual for passenger cars	M	$S^M = 498 \pm 48$
Automatic Coin Machine Users	A	$S^A = 618 \pm 30$
Manual for vehicles other than passenger cars	T	$S^T = 138 \pm 78$
Mixture of ETC passenger cars and semi-trucks traveling at 35mph	E_P and E_T	$S^E = 1560 \pm 120$

In order to achieve calibration such that the throughput equations accurately reflect field values of the processing rates, it was necessary to manipulate the driver/vehicle property values listed in Table 3 *(2)*. These property values vary from region to region anyway. For instance, in a community with a large population of senior citizens, the driver reaction time property, T_R, may have a value of 2.1 rather than 1.8 seconds. In addition, stop time values to pay the toll may be 2.0 rather than 1.5 seconds. These property values, once inserted into the throughput equations, would result in higher values for the processing times for this community. This would compute to smaller throughputs and would accurately reflect the smaller corresponding processing rates measured in the field.

Table 3: SHAKER's Default Properties of the Traffic Categories at the Toll Facilities Table

	M	**T**	**A**	**E_P**	**E_T**
l_X =Average vehicle Length (meters)	5.8	21	5.8	5.8	21
b_X = Distance between vehicles (meters)	2.0	3.0	2.0	2.0	3.0
a_X = Vehicles' Acceleration (meters/second^2)	2.0	0.25	2.0	2.0	0.25
d_X =Vehicles' Deceleration (meters/second^2)	2.0	0.25	2.0	2.0	0.25
T_R = Drivers' reaction Time (seconds)	1.8	1.8	1.8	1.8	1.8
$t_{stop\ X}$ = Stop–Time at payment (seconds)	1.5	4.7	0.075	0.0	0.0

5. EXPERIMENTAL DESIGN PROCEDURE

Before any efforts towards model calibration can begin the experimenter must define a complete, beginning to end, experimental procedure. To develop a micro simulation model calibration and validation process two well recognized methods have been combined, taking the strengths of each, and are slightly adapted to simulate isolated toll plaza operations. General microsimulation calibration techniques are followed to calibrate both SHAKER and VISSIM because there is no literature specific to calibrating a deterministic model versus a stochastic model. According to the FHWA Traffic Analysis Toolbox, Volume 3 *(7)* the overall process for developing and applying a microsimulation model to a specific traffic analysis situation consists of seven major tasks (also shown in Figure 1):

1. Identification of Study Purpose, Scope, and Approach.
2. Data Collection and Preparation.
3. Base Model Development.
4. Basic Model Error Checking/Initial Evaluation
5. Calibration of model
6. Alternatives Analysis.
7. Final Report and Technical Documentation.

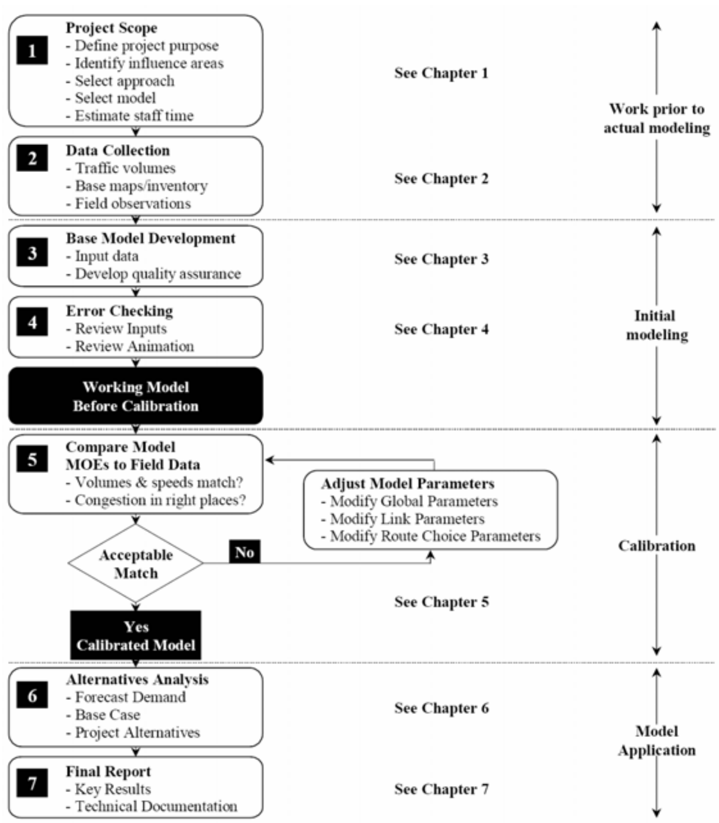

Figure 1: FHWA Calibration Flow Chart (7)

Steps 6 and 7 of this model will not be directly addressed in this research because they are included in the FHWA documentation for clients who wish to use the methodology specifically for planning purposes. The calibration model described by the FHWA follows a sound process to calibrate a microsimulation model, but fails to mention any steps pertaining to model feasibility and validation. To make up for this vacancy, the calibration model proposed by Park and Qi is adopted for its extensive validation procedure *(8)*. The final two calibration steps adopted for this study are: i) the evaluation of final calibrated parameter set and, ii) ensuring statistical and visual validation of calibrated model. The combination of these two calibration procedures results in a more complete experimental design process for calibrating a microsimulation model. It is noted that the calibration process can not continue to the next step unless the previous step in this process is first satisfied. In some case where the step's objectives are not possible, the user may have to reconsider the activity in a previous step in order to move on or in extreme cases restart the entire experiment from step one. Each step is briefly described in the following sections.

5.1 Step 1 Identification of Study Purpose, Scope, and Approach

This step includes how and why the study is to be conducted. In addition, this step includes choosing which sites to use for data collection and which software will be used for model development and evaluation. Step one was partially addressed in the early stages of this paper in the introduction, purpose, and literature review sections. The next phase of the experiment outline consists of defining a particular data collection procedure. The data collection aims at analyzing toll lanes processing times (or service time) and all other the factors affecting the latter. If a better understanding of toll plazas' operation/ processing time is established, results may be used for simulation models' calibration.

5.2 Step 2 Data Collection and Preparation

To effectively calibrate any simulation model site data must be used to compare simulation results to for model validity. Information collected in the field typically consists of three types of data: site geometries, traffic characteristics, and vehicle distributions. Calibration data commonly consists of one or more of traffic characteristics, such as: capacity, demand, travel time, queue length, delay, etc. To avoid collecting useless field data researchers should choose which measures of effectiveness will be used for model evaluation and calibration. For this research a one day pilot study was conducted at a randomly selected toll plaza to test data collection equipment, data collection procedures, investigate site details, and visually observe traffic patterns. A rehearsal data extraction activity was also conducted to gain experience in the process, practice extraction techniques, and learn the limitations of the field collected data. The preparation of the final data is just as important as the collection process. The raw data collected in the field has to be extracted so that it represents the data needed to satisfy the goal of the collection. The data must also be checked for consistency at this step. It is important that field collected data is complete and targets a specific traffic condition. For example, field collected capacities may be checked against the Highway Capacity Manual analysis to ensure there is not large variation. If there is skeptical variation of any classification it is recommended that the data be reevaluated to confirm differences are genuine.

5.3 Step 3 Base Model Development

The goal of the base model development is to accurately recreate the traffic organization, operation, and driver behaviors that existed at the field data collection site. Whether it is visually, like in the VISSIM software, or numerically, as in the SHAKER tool, this step provides verification that the software is compatible to the uniqueness of each site. Because the SHAKER model was designed strictly for toll plazas there is little to no initial model development required. In order to build a toll plaza configuration, the SHAKER model simply requires the user to input number of lanes, vehicle distributions, and demand volumes. In contrast, VISSIM was designed to cater to various transportation applications so the toll plaza and traffic elements pertaining to toll plazas must be entered manually for each and every plaza. In addition to building the network in VISSIM, this is the step where the user would also include all the traffic characteristics into the model; such as: demand, vehicle distributions, and route choices. Also only applicable to the VISSIM model, the user must define an array of random seeds to use for calibration procedures. When using micro simulation programs the user must sometimes be creative when building a model to recreate field conditions because not all facilities are preprogrammed in the software. For example, VISSIM has no dedicated toll plaza features but the stopping behavior of individual drivers can be well represented by modeling payment processing times as stop controlled intersection features, which are available. The specifics on how to develop the models to replicate the field conditions is well described in each program's user manual, thus these will serve as the primary reference for building the initial model. An extensive review of how each model is developed is provided later in the paper.

5.4 Step 4 Basic Model Error Checking/Initial Evaluation

Error checking and initial evaluations of the simulated model should be completed before calibration takes place. If a traffic or network related error is detected after calibration it can potentially cause the entire calibration process to be deemed obsolete. One aspect of basic error checking is visually observing the base case animation model to ensure that general traffic behaviors are observed. The importance of manual visualization checks can not be overlooked because as powerful as the computer is, it does not have the judgment and reasoning skills of the human mind. The computer will undoubtedly produce the optimum model but, without correct user defined parameters, it does so without meaningful knowledge if the simulation specifics are realistic (i.e. vehicles overlapping, unrealistic lane change maneuvers, vehicles disappearing during simulation, etc.). At this point in the procedure is a good step to check that input distributions are functioning as intended. Initial checks may consist of categorizing the throughput results by vehicle type to ensure that the same inputted vehicle distributions were used for the simulation or recording all toll plaza dwell times and comparing them against the programmed dwell time distribution to ensure that the model recreates the intended activity. The later portion of this step is to initially evaluate the base model to check its ability to simulate field conditions. Statistical tests should be conducted to determine if the simulated measure of effectiveness (MOE) are within an acceptable range of the target values. Therefore, a well acceptable range of error must be defined. If the software can initially predict MOEs based on default values alone then there is no need to calibrate the model any further.

5.5 Step 5 Model Calibration

The next step of the experimental design is the actual calibration procedure. According to the FHWA Guidelines for Applying Traffic Microsimulation Modeling Software *(7):*

> "Calibration is the adjustment of model parameters to improve the model's ability to reproduce local driver behavior and traffic performance characteristics. Calibration is necessary because no single model can be expected to be equally accurate for all possible traffic conditions. Even the most detailed microsimulation model still contains only a portion of all of the variables that affect real-world traffic conditions. Since no single model can include the whole universe of variables, every model must be adapted to local conditions."

Before any calibration can take place it is important to determine which MOE will be used for as a surrogate measure to match the model to. Next, influencing parameters must be classified as either directly affecting the MOE or not affecting the MOE. Only simulation parameters that affect the MOE should be reviewed and adjusted in order to reach the optimum parameter configuration. When possible, the parameters in which field data is available should be implemented before any parameters are adjusted because measured values represent justifiable alterations to the base model. Remaining parameters can be used for model calibration in a series of logical, sequential steps. Each time a parameter is adjusted the new model should be evaluated for performance measures and then compared to field conditions. Checking the model results for similarity to the field measured conditions requires use of statistical analysis to improve reliability. The differences between the predicted model outputs, when compared to field measured values, are called the residuals and are used to evaluate the usefulness of the model. Root mean square error, analysis of variance (ANOVA), and/or student t-tests are a few variations of statistical tests used in such analysis *(34)*. In addition to error reduction and residual analysis, a statistically acceptable number of simulation runs will have to be calculated. When using a stochastic microsimulation program multiple runs of the same model are required because results can and will vary based on the simulation seed number. Seed numbers represent the random data source used to give each vehicle their individual properties. To gain model validity, one must compute the minimum number of repeated microsimulation model runs needed to estimate any mean with a certain level of confidence that the true mean of what is being tested actually falls within the target interval. However, the equation used to determine the number of simulation runs requires knowledge of result statistics before the equation can be used; results such as the sample standard deviation, desired length of the confidence interval, and desired level of confidence. This nuance creates an iterative process that usually begins with estimating the equation inputs and then after model run the equation is used as a check to see if enough runs had been conducted already. If it is determined that not enough simulation repetitions were conducted to prove statistical significance the simulation must be repeated and checked one more time. Unlike stochastic models, deterministic models will always produce the same results when repetitions are preformed with the same input data. Therefore, there is no need to calculate how many runs are required when using this type of simulation.

5.6 Step 6 Evaluation of Calibrated Parameter Set

The evaluation of the calibrated parameter set step is to compare the original default parameters versus the calibrated parameters found in the previous step. It is important to compare the results to ensure that the calibrated parameters are justifiable and not just the values that force the calibration to match field conditions.

5.7 Step 7 Validation

The final model calibration step consists of two verification sub steps that are intended to finalize the model and approve of its functionality. The first of the two validation steps requires that the microsimulation model be visually evaluated for reasonability and that ensure the model functions realistically. Visualization checks are important because the computer will always produce optimal models but sometimes does so without knowing if the simulation specifics are realistic by human drivers (i.e. vehicles overlapping, unrealistic lane change maneuvers, vehicles disappearing during simulation). The second of the verification steps is to statistically verity that the calibrated model estimates, within range, similar results when compared to an additional untried data set. Unusually this is done so using data collected at the same site but for a different day or using data collected at a similar site but different location. For this research the validation data will be extracted from the same site but from a different day and to ensure validation, an additional parameter to the original will be examined. These last two steps are very simple compared to the calibration techniques previously discussed, but prove to be extremely important in ensuring that the microsimulation model compromise a useful model.

6. IMPLEMENTING THE EXPERIMENTAL DESIGN PROCEDURE

To complete the objectives of this research, the described methodology had to be implemented for use on both simulation programs. The first two steps of the procedure refer to pre-model development activities, so within these steps there is no need to distinguish between tasks particular for SHAKER and VISSIM. However, once the model development stage had been reached unique techniques that pertain to each program had to be applied. The organization of the experimental design portion of this research follows a similar sequence. The first two homogeneous steps are introduced first and are followed by two sections on the unique steps pertaining to each program's specific calibration procedures.

6.1 Step 1 Identification of Study Purpose, Scope, and Approach

The benefits of simulation models led to the purpose of this study, which was to examine the effectiveness of two toll modeling programs that are similar in purpose but vary in approach and methodology. Both SHAKER and a VISSIM toll plaza models have the potential to work as a tool that can estimate the maximum throughput and capacity of toll plazas. There is no research on whether the SHAKER or VISSIM model would better represent toll facility operations, therefore lies the need for the research conducted in this thesis. Some major benefits from using these models to simulate toll plazas is that one can examine how traffic conditions will change when a lane is closed due to maintenance or construction, or how adding more lanes would improve the plaza operations. Since the population in the Central Florida area continues to grow rapidly simulation will prove to be useful to adjust for future demands. In summary, the objective of this study is to develop, calibrate, and analyze two toll plaza simulation models whose theory differs by its magnitude of meticulousness. To conduct this research necessary data had to be collected at multiple toll plazas along the Florida Turnpike Enterprise network in Central Florida. Ultimately, the process for research consists of using the data collected in the field to serve as the principal dataset in which later developed simulation results will be calibrated to predict.

6.2 Step 2 Data Collection and Preparation

In an effort to correctly calibrate the SHAKER toll plaza tool and VISSIM model, field data was collected at four Florida Turnpike operated toll facilities located in Central Florida. Each site differed from the others in terms of number of lanes, lane configuration, toll base fee, highway location, traffic demand, and vehicle type percentage. The sites chosen for data collection were: the Lake Jesup Mainline Plaza along the Seminole Expressway (SR-417), the Beachline West Expressway Toll Plaza along the SR-528, the Daniel Webster Western Beltway Plaza along SR-429, and the Leesburg Toll Plaza along the Florida Turnpike Mainline SR-91.

Table 4 presents the different elements at each of the toll plaza faculties. If not noted in the table it can be assumed that opposing directions share similar characteristics. Figure 2 and Figure 3 shows an iconic pictorial of a typical toll plaza configuration and how lanes along the Florida Turnpike are color coordinated. Figure 4 shows all 5 toll plaza configurations with number of lanes and shows how lane configurations split from original travel lanes to multiple toll plaza service lanes.

Table 4: Field Data Site Locations - Toll Plaza Characteristics

	base toll fee (two axel)	# of total lanes	# of high speed ETC lanes	# of non high speed ETC lanes	# of manual pay lanes	# of exact change coin lanes	gate regulated lanes
Lake Jesup SR-417	$2.00	4	0	2	2	0	Yes
Beachline Eastbound SR-528	$0.75	6	0	2	2	2	Yes
Beachline Westbound SR-528	$0.75	5	0	2	2	1	Yes
Western Beltway SR -429	$1.00	4	2	1	1	0	Yes
Leesburg Plaza SR-91	$2.50	5	0	1	4	0	Yes

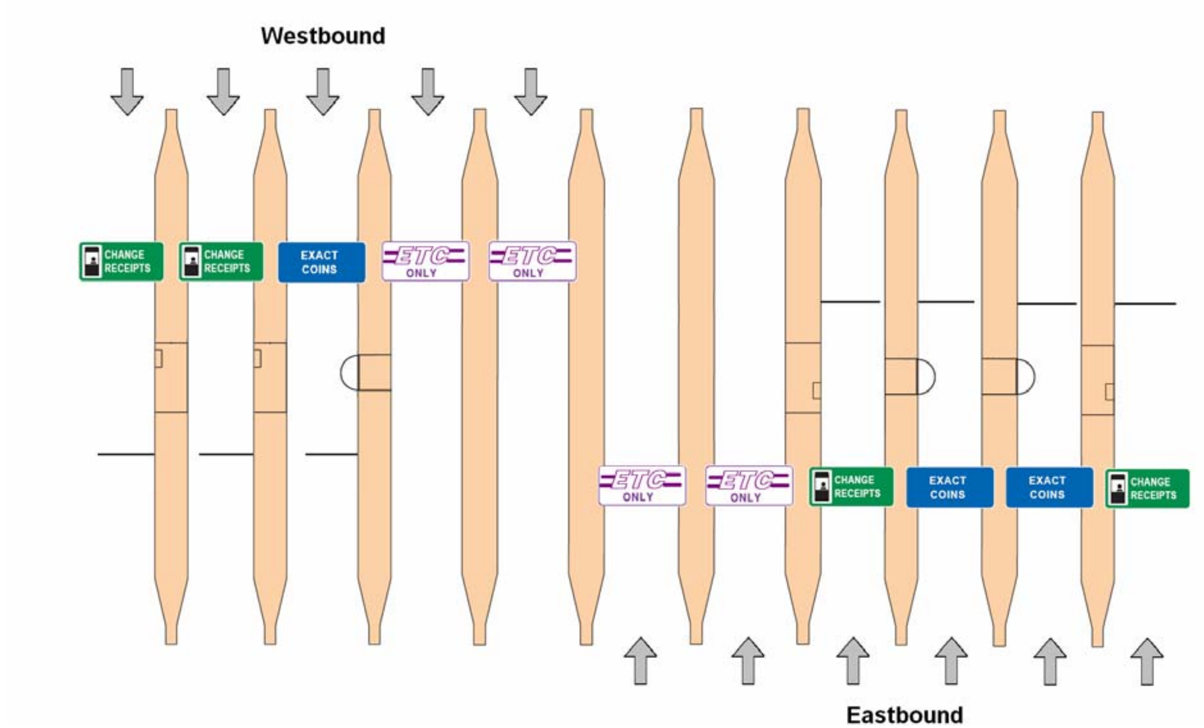

Figure 2: SR-528 Beachline West Toll Plaza Configuration

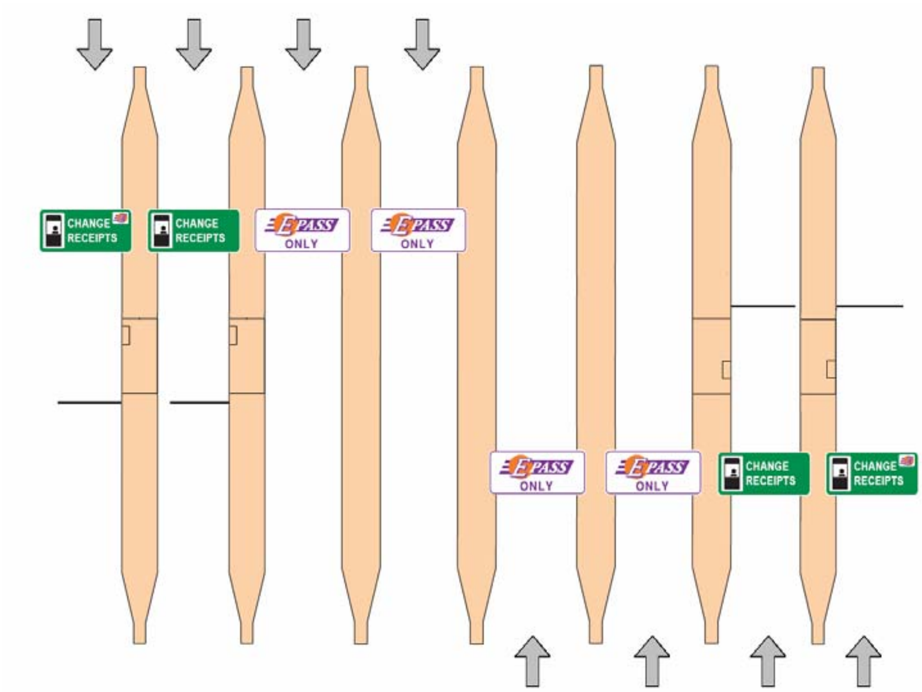

Figure 3: SR 417 Lake Jesup Mainline Plaza Configuration

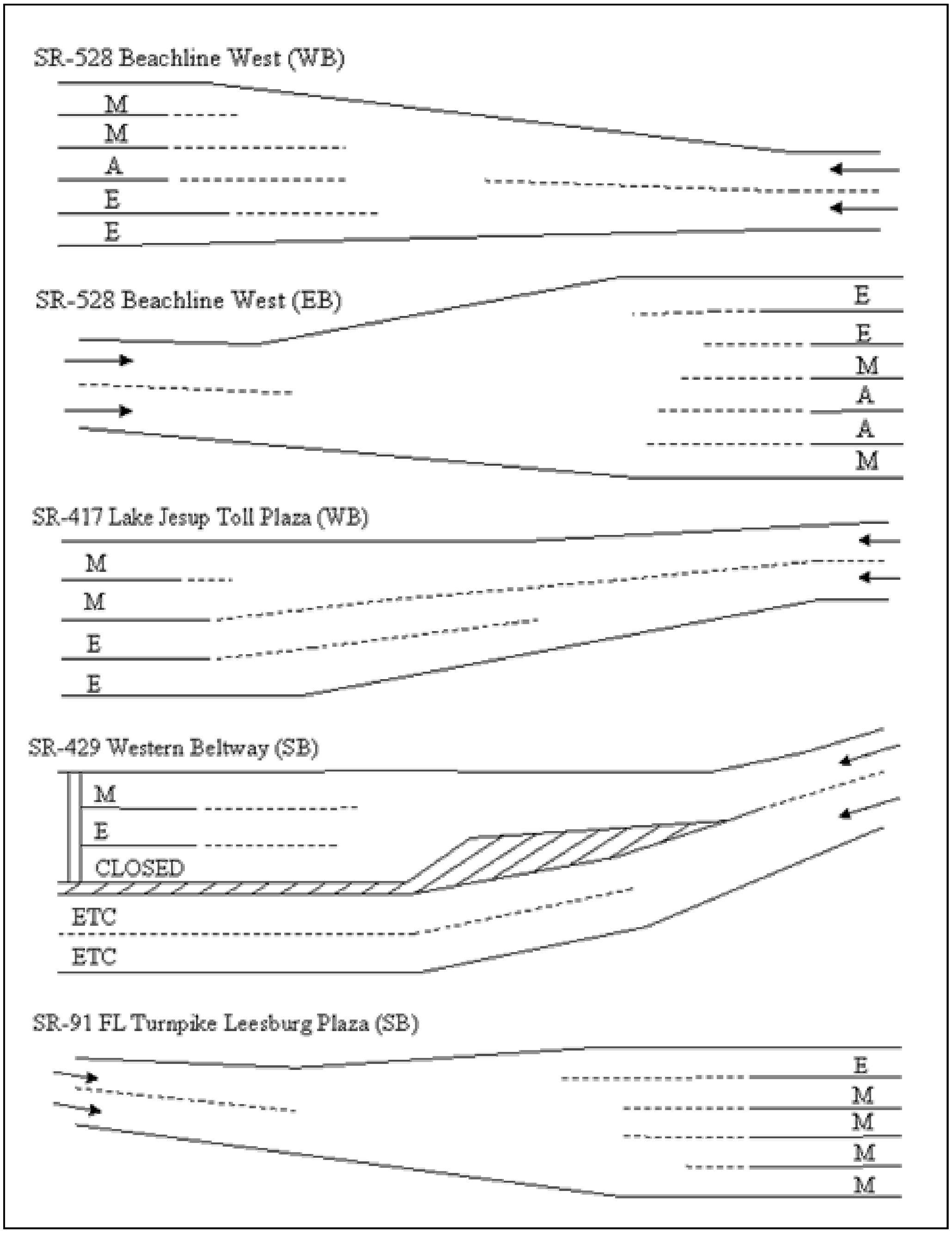

Figure 4: Toll Plaza Configurations.

Field data collected was categorized into three major categories: traffic characteristics, vehicle distributions, and toll plaza characteristics. The traffic data collected at each plaza was volume, demand, throughput, and queue lengths. The individual vehicle data collected was vehicle type, lane choice, processing time, payment type, whether the vehicle arrived during a queue or not, arrival time, departure time, and inter-arrival time between vehicles. The toll plaza data recorded consisted of number of lanes, number of each type of lane, whether the direction of travel was into or out of the metropolitan area, and whether the plaza was observed during the AM or PM. Table 5 is provided to show an inventory of the video data recorded at and the amount of lane-hours at each of the toll plaza sites.

As shown in Table 5, different lane-hours totals were collected for each data collection site. This total varies because each plaza's effective role towards the research goal differed and because of data collection feasibility limitations. Several more lane hours were colleted at the Beachline West and Lake Jesup because this site utilized all possible lane types, it experienced the most diverse vehicle and payment type percentages, and was selected as the plaza to use for primary calibration purposes. Calibration requires several additional hours of data to ensure that enough data points are collected to make an accurate estimate of driver behaviors. Data collected from the Lake Jesup, Western Beltyway, and Leesburg Plaza was primarily used for verification purposes so not as many hours of data were needed. After preliminary analysis of the Lake Jesup plaza it was determined that for model verification the amount of lane-hours required for the Western Beltway and Leesburg plaza could be reduced.

Table 5: Video Data Inventory

Toll Plaza Site	Direction of Travel	Peak Period	Days Collected	Lane-Hours Collected
Lake Jesup SR-417	Northbound	AM	6	24
		PM	6	24
	Southbound	AM	6	24
		PM	6	24
Beachline West SR-528	Eastbound	AM	5	25
		PM	6	30
	Westbound	AM	5	25
		PM	6	30
Western Beltway SR-429	Northbound	PM	1	6
	Southbound	PM	1	6
Leesburg Plaza SR-91	Southbound	AM	1	10
		PM	1	10

6.2.2 Data Collection Equipment Configuration

At each site four cameras are used to capture operations at and upstream of the toll plaza facility. All of the four cameras were started simultaneously and each captures a different condition. Two of the four cameras were used to capture one approach and the other two are simultaneously capturing the opposing direction. Figure 5 shows an aerial view of the Lake Jesup Toll Plaza and is provided to show an example of how the cameras were configured at each site. One camera (Figure 5: NB & SB Camera 1) in each direction was primarily used to capture the throughput, processing times, inter-arrival time, vehicle type, and payment type for each lane. The second camera (Figure 5: NB & SB Camera 2) in each direction was set to capture the demand and queue conditions. A still frame image of the video is provides as an example of both camera set ups in Figure 6 and Figure 7. Similar placed camera arrangements were implemented at the other data collection locations, but exact locations were ultimately restricted by right of way and safety considerations. The video image collection process for each site took place over multiple days and captured two hours during the morning (7-9AM) and afternoon (4-6PM) rush hour peaks.

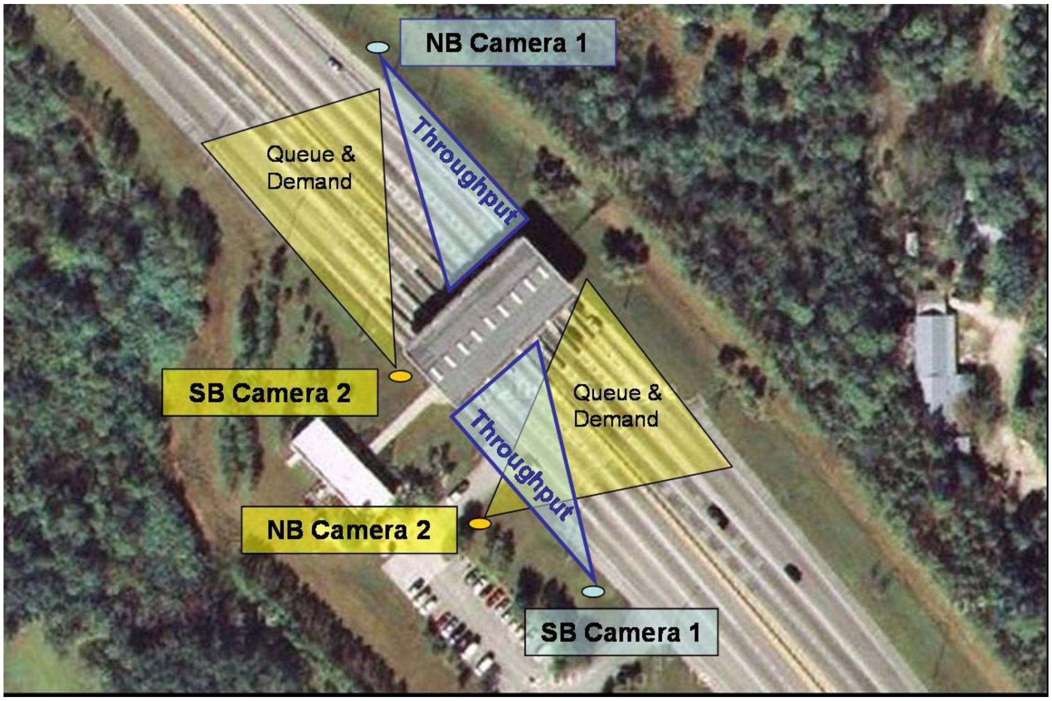

Figure 5: Camera setup Configuration Example

Figure 6: Video Image Example of Camera 1 (Throughput and Processing time)

Figure 7: Video Image Example of Camera 2 (Demand and Queue Length)

6.2.3 Data Extraction Procedure

Once the digital videos were transferred from the camera to the computer, the software Adobe Premier Professional was used to view the files. This program allows the user to study videos to the accuracy of 30 frames per second. Traffic characteristics such as demand, throughput, processing rates, and queue lengths of different toll categories were extracted from the digital video. The throughput was recorded as the number of vehicles that pass through the toll per 15 minutes. The demand is the throughput plus the length of the queue, if present, and is also measured every 15 minutes. In addition, the lane choice and vehicle type (passenger car or truck) was recorded. The following descriptions explain each variable and how it was collected:

- Throughput – recorded as the number of vehicles that pass through the toll plaza within a period of time. Each vehicle is classified by an arrival time and departure time. Recording specific arrival times allows for throughput to be determined for any time frame desired.

- Demand – is the throughput plus the length of the queue, if present. This was measured by counting the number of vehicles in the queue at any given time and adding that value to the throughput for the same period.
- Processing Time – is the calculated difference between the arrival and departure time. The arrival time is the instant that the vehicle makes a complete stop within the toll collectors range. It was observed that a number of drivers attempt to offer their payment to the toll collector while their vehicle is still slowly crawling. In this case the arrival time is classified as the instant the individual begins the transaction with the toll collector. The departure time is recorded upon the onset of acceleration following the payment. The processing times can be very short so it is important to be as precise as possible; therefore, the arrival and departure times are extracted from the video with $1/30^{th}$ of a second accuracy.
- Move-ahead-time – The elapsed time between the lead vehicle departure time and following vehicles arrival time. Calculated as the Inter-arrival time minus the Processing time.
- Queue Length – is measured as the number of vehicles building up in each lane who are waiting to be served by toll attendant. Queue length was measured by simply counting the number of vehicles in the queue. Through the benefits of video review this can be recorded at any point in time.
- Arrival on Queue – When a vehicle approaches the toll plaza it is either faced with a queue resulting from previous toll plaza delay or is faced with no queue. For this research arriving during a queue means they are at least the third vehicle in line. A vehicle

arriving as the first or second vehicle at the toll plaza is not considered arriving during queue.

- Vehicle Type –two categories of vehicles were recorded, either a vehicle was recorded as a passenger car or a truck. A truck is considered any vehicle having or towing with more than two axels touching the pavement.
- Lane Type – 4 types of lanes were observed. They are Electronic Toll Collection lanes (ETC), High Speed ETC, Manual Attendant Assisted lanes and Automatic Coin Machine lanes,
- AM/PM – Video analysis of periods occurring in the morning peak (7-9AM) were classified as AM and periods occurring in the afternoon peak (4-6PM) were classified as PM.

Each vehicle passing the toll plaza during the analysis period was considered one data point. Every non ETC vehicle was observed for above criteria and data was recorded in a table that exceeded 20,000 entries. Vehicles using the ETC lanes were only observed for vehicle type, demand, and throughput. By design, ETC vehicles do not stop so in their case the other parameters proved no practical significance. An example of the individual vehicle data extraction table is provided in Table 6. All vehicles not using the ETC lane use either the ACM or manned lanes. These vehicles are subject to all extraction information data. An additional table is used to organize all vehicle extraction data for volumes, demands, queue lengths and, vehicle type percentages. An example of that table is shown in Table 7 below. All vehicles not using the ETC use either the ACM or manual lanes and they were subject to all processing information data. The number of non-ETC vehicles and lane-hours analyzed for each toll plaza location is listed below in Table 8.

Table 6: Example of Individual Vehicle Data Extraction Table

Veh ID	site	Lane #	Time AM(0), PM(1)	metro out(0) into(1)	lane type m(0) a(1)	Vehicle Type (M or T)	Arrival Time Min	Arrival Time Second	Arrival Time Frame	Departure Time Second	Departure Time Frame	Processing Time (sec)	inter-arrival time (sec)	queue present(1), no queue(0)	volume (veh/15min)
1	6	1	1	1	0	M	0	0	0	7	17	7.57	4.33	1	84
2	6	1	1	1	0	M	0	11	10	12	15	1.17	3.77	1	84
3	6	1	1	1	0	M	0	16	6	18	22	2.53	3.70	1	84
4	6	1	1	1	0	M	0	22	25	26	18	3.77	4.10	1	84
5	6	1	1	1	0	M	0	31	13	44	10	12.90	4.83	1	84
6	6	1	1	1	0	T	0	47	11	57	29	10.60	3.03	1	84
7	6	1	1	1	0	M	1	4	2	9	21	5.63	6.10	1	84
8	6	1	1	1	0	M	1	13	23	16	10	2.57	4.07	1	84
9	6	1	1	1	0	M	1	20	14	23	16	3.07	4.13	1	84
10	6	1	1	1	0	M	1	28	15	30	19	2.13	4.97	1	84
11	6	1	1	1	0	M	1	35	8	40	14	5.20	4.63	1	84
12	6	1	1	1	0	M	1	45	14	47	14	2.00	5.00	1	84
13	6	1	1	1	0	M	1	49	14	52	13	2.97	2.00	1	84
14	6	1	1	1	0	M	1	56	16	58	15	1.97	4.10	1	84
15	6	1	1	1	0	M	2	2	8	6	1	3.77	3.77	1	84
16	6	1	1	1	0	M	2	8	22	11	12	2.67	2.70	1	84
17	6	1	1	1	0	M	2	15	7	19	0	3.77	3.83	1	84
18	6	1	1	1	0	M	2	22	28	24	24	1.87	3.93	1	84
19	6	1	1	1	0	M	2	27	23	36	2	8.30	2.97	1	84
20	6	1	1	1	0	M	2	40	28	43	6	2.27	4.87	1	84
21	6	1	1	1	0	M	2	47	21	50	13	2.73	4.50	1	84
22	6	1	1	1	0	M	2	55	0	57	4	2.13	4.57	1	84
23	6	1	1	1	0	M	2	59	24	64	29	5.17	2.67	1	84
24	6	1	1	1	0	M	3	8	27	13	8	4.37	3.93	1	84
25	6	1	1	1	0	M	3	16	14	21	9	4.83	3.20	1	84
26	6	1	1	1	0	M	3	25	1	28	22	3.70	3.73	1	84
27	6	1	1	1	0	M	3	33	8	37	12	4.13	4.53	1	84
28	6	1	1	1	0	M	3	42	18	44	24	2.20	5.20	1	84

Table 7: Example of Traffic Data Extraction Table

Date/Time	15 min period	Passenger Car Throughput veh/per lane (from out to inner lane)					Truck Throughput veh/per lane (from out to inner lane)					Through-put Total	Que Length (per 15 min)			Demand (veh/hr)
		M	M	A	E_p	E_p	T	T	A	E_T	E_T		M	M	A	
2/25/2007 4:00 -5:00 PM	0-15	87	95	93	113	132	4	0	0	17	4	545	10	12	8	575
	15-30	87	88	89	137	144	2	2	0	6	3	558	1	2	2	563
	30-45	77	86	82	118	158	4	1	0	6	4	536	0	2	0	538
	45-60	75	88	74	141	149	1	0	0	12	3	543	15	10	15	583
Hourly Total		326	357	338	509	583	5	1	0	18	7	2182	-	-	-	2222

Table 8: Data Analysis Inventory

Data Acquisition Location	Vehicles Procesed	Lane-Hours Analyzed
Lake Jesup SR-417		
Northbound	3,297	16
Southbound	3,904	16
Beachline West SR-528		
Eastbound	5,404	24
Westbound	5,619	19.5
Western Beltway SR-429		
Northbound	219	2
Southbound	313	2
Leesburg Plaza SR-91		
Southbound	1,694	16
Total	20,450	95.5

6.2.4 Data Investigation

To calibrate both the SHAKER and VISSIM models the capacity was selected as the measure of effectiveness to first evaluate, hence field observed capacity was compared to the model estimated capacity. To measure capacity in the field, the FHWA recommends observing locations where queues persist for at least 15 consecutive minutes and then measure the flow rate at the point where the queue discharges. The resulting flow rate is the field-measured capacity *(7)*. Therefore, only periods under queuing conditions were used in the calibration and validation process. The following tables (Table 9 and Table 10) show the data used for calibration and validation of the SHAKER model. Each of the periods listed was under constant queuing during the data extraction period. The calibration periods of both the manned and ACM lanes was observed along the Westbound Beachline West SR-528 toll plaza on Feb 25, 2008. The verification data is compromised of a mixture of data from the same toll plaza but from the Westbound approach on Feb. 25, 2008 and the Eastbound and Westbound approach on Nov. 13, 2007.

Table 9: SR 528 Data for Calibration and Validation of Manned and ACM Lanes

		Period	Demand (vphpl)	Capacity (vphpl)	Queue Length (vehpl)	% Trucks	% non Trucks
Manned Lanes	Calibration Data	1	339	336	3	0.036	0.964
		2	374	364	10	0.044	0.956
		3	357	356	1	0.022	0.978
		4	402	344	58	0.012	0.988
		5	415	332	83	0.036	0.964
		6	435	352	83	0.034	0.966
		7	351	348	3	0.011	0.989
		8	392	380	12	0.000	1.000
		9	362	360	2	0.022	0.978
		10	350	348	2	0.011	0.989
		11	362	352	10	0.000	1.000
		12	438	380	58	0.000	1.000
		13	435	352	83	0.011	0.989
		14	451	368	83	0.011	0.989
	Validation Data	15	372	364	8	0.033	0.967
		16	366	364	2	0.022	0.978
		17	329	328	1	0.101	0.899
		18	374	364	10	0.032	0.968
		19	385	376	9	0.031	0.969
		20	395	388	7	0.020	0.980
		21	382	380	2	0.021	0.979
ACM Lanes	Calibration Data	1	388	376	12	0.0	1.0
		2	380	372	8	0.0	1.0
		3	358	356	2	0.0	1.0
		4	338	328	10	0.0	1.0
		5	370	312	58	0.0	1.0
	Validation Data	6	369	364	5	0.0	1.0
		7	333	328	5	0.0	1.0
		8	319	316	3	0.0	1.0
		9	369	364	5	0.0	1.0
		10	345	340	5	0.0	1.0
		11	379	376	3	0.0	1.0
		12	394	392	2	0.0	1.0

Table 10: SR-417, SR-429, FL Turnpike Queuing Periods

Additional Toll Site Data Used for Validation of Calibration technique

	Period	Capacity (vphpl)	% Trucks	% non trucks
SR 417 Manned Lanes	1	360	0.000	1.000
	2	376	0.000	1.000
	3	342	0.018	0.982
	4	360	0.033	0.967
	5	384	0.000	1.000
	6	372	0.000	1.000
	7	344	0.012	0.988
	8	388	0.031	0.969
	9	348	0.046	0.954
	10	342	0.035	0.965
SR 429 Manned Lanes	1	420	0.0	1.000
	2	380	0.0	1.000
FL Turnpike Manned Lanes	1	204	0.0	1.000
	2	200	0.0	0.970

To determine the capacity it was also important to filter the data to eliminate potential unfavorable traffic conditions. The test to run on the data was to determine if any of the saturated periods are performing statistically different than any of the other periods is described. To test this, the individual vehicle's processing time and inter-arrival time from each group was statistically compared to the same parameters in the other group by use of ANOVA statistics. The importance of this check is to determine if there were any unaccountable errors in the traffic makeup that were not detectable from simple observation. Also, in order to investigate different periods, the service times must not be significantly different; otherwise the periods will not have analytical value as a data test set to be used for calibration *(9)*. If the processing times are significantly different from one group to another it suggests that the period at question experiences unique conditions that could be attributed by factors other then the traffic; for example, the speed of toll plaza operator, a slower release gate, congestion downstream, etc. If the inter-arrival time distribution of one period was statistically different from the next period it suggests that there is not as constant of a queue as expected, thus capacity is not reached. When a queue is present the spacing of vehicles is assumed to be generally the same distance thus the inter-arrival time should be generally consistent.

The statistics used for this test was an Analysis of Variance (ANOVA). ANOVA is similar to regression in that it is used to investigate and model the relationship between a response variable and one or more predictor variables. However, analysis of variance differs from regression in two ways: the predictor variables are qualitative (categorical), and no assumption is made about the nature of the relationship (that is, the model does not include coefficients for variables). In effect, analysis of variance extends the two-sample t-test for testing the equality of two population means to a more general null hypothesis of comparing the equality of more than two means, versus them not all being equal.

The output from an ANOVA study is arranged in the tables below (Table 11). The table consists of the sources of variation, their degrees of freedom, the total sum of squares, and the mean squares. The ANOVA table also includes the F-statistics and p-values. These values were used to determine whether the predictors or factors are significantly related to the response. The following describes the statistical test:

Null Hypothesis (Ho): the lanes data are not significantly different

Alternative Hypothesis (Ha): the lanes data are significantly different

Test Statistic: (p-value) significance of 95%

The following is a list of the components of the ANOVA tables (Table 11):

- Source - indicates the source of variation, either from the factor, the interaction, or the error. The total is a sum of all the sources.
- DF - degrees of freedom from each source. If a factor has three levels, the degree of freedom is 2 (n-1)
- SS - sum of squares between groups (factor) and the sum of squares within groups (error)
- MS - mean squares are found by dividing the sum of squares by the degrees of freedom.
- F - Calculated by dividing the factor MS by the error MS; one can compare this ratio against a critical F found in a table or use the p-value to determine whether a factor is significant.
- P - used to determine whether a factor is significant; typically compared against an alpha value of 0.05. If the p-value is lower than 0.05, then the factor is significant.

From Table 11 the p-value of the processing times and inter-arrival times for each period are $p=0.974$ and $p = 0.108$ respectively. P-values greater then 0.05 leads to failing to reject the null hypothesis and suggests that there is not enough statistical evidence to disprove that there is statistical difference between each period's processing times and inter-arrival times. The conclusion that can be drawn for this analysis is that even though the throughputs per hour in each time frame varies from 328 to 388 vehicles per hour per lane there is no indication that the capacity is affected by exterior elements. Because these two tests show no significant differences they are suitable to serve as the primary targets used for simulation calibration. If a statistical difference did occur in one of the time frames it would suggest that this time frame does not follow one of the constraints of determining capacity, which is that results should be repeatable under common conditions.

Table 11: ANOVA Table SR-528 Processing Times and Inter-Arrival Times of Calibration and Verification Data

One-way ANOVA: Processing Time versus Period

Source	***DF***	***SS***	***MS***	***F***	***P***
Period	20	334.6	16.7	0.48	0.974
Error	1844	64085.5	34.8		
Total	1864	64420.2			

One-way ANOVA: Inter-arrival Time versus Period

Source	***DF***	***SS***	***MS***	***F***	***P***
period	20	52.0	2.6	1.41	0.108
Error	1841	3401.2	1.85		
Total	1861	3453.1			

Individual Period Results

		Processing Time		***Inter-arrival Time***	
Level	***N***	***Mean***	***StDev***	***Mean***	***StDev***
1	336	6.41	11.87	4.10	1.14
3	364	5.55	3.82	4.13	1.09
4	356	5.70	4.48	4.17	1.24
7	344	6.08	5.40	4.26	1.96
8	332	5.87	5.42	4.30	1.21
9	352	5.31	5.58	4.14	0.78
11	348	5.15	4.61	4.59	1.05
13	380	5.02	4.02	4.42	1.11
14	360	5.46	3.74	4.54	1.07
15	348	5.77	4.33	4.43	1.26
16	352	5.73	3.84	4.45	1.15
17	380	5.30	4.22	4.23	0.83
18	352	5.86	5.31	4.33	1.00
19	368	5.49	5.58	4.20	0.93
42	364	5.02	3.93	4.64	1.50
44	364	5.62	10.75	4.25	1.18
45	328	6.64	5.47	4.38	2.55
48	364	5.43	5.56	4.62	1.56
50	376	5.01	6.07	4.61	1.75
51	388	5.62	7.28	4.28	0.99
52	380	5.10	5.09	4.36	1.90

7. THE SHAKER MODEL

7.1 Step 3 SHAKER Model Development

As mentioned previously, calibration of simulation models is necessary if the initial default parameters of the model being used do not result in verifiable traffic measure. However, because even the most detailed model still contains only a portion of all of the variables that affect real-world traffic conditions almost every model will require some form of calibration. Also, since no single software can realistically include each and every variable, every model must be adapted to fulfill the objectives of the study. The objective of the calibration process is to find a set of parameter values for the model that best reproduces local traffic conditions *(7)*. The objective of the calibration of SHAKER is to find a set of parameter values that best reproduces the capacity of the non-ETC payment lanes at toll plazas.

7.1.1 Using the SHAKER Software

The developments of the SHAKER model have occurred fairly recently and the program is not yet as globally distributed as the VISSIM software is; therefore the procedure on how to use SHAKER is also relatively new and seldom used before. This section is provided to show the essential steps in obtaining research goals, toll plaza capacities and throughputs, by means of the SHAKER software.

The SHAKER plaza calculator finds the best configuration for the given lane and the lane properties and finds the throughput of all the configurations and displays the throughput of the best configuration. Nonspecific input values are used for the purpose of this section and are only used for software introduction. When SHAKER is first opened the user is presented with a window such as the one provided in Figure 8. Before moving on, the user is to choose whether to allow the SHAKER software to automatically calculate the best lane type configuration (first choice in Figure 8) or enter each of the number and type of lanes manually (the second choice in Figure 8).

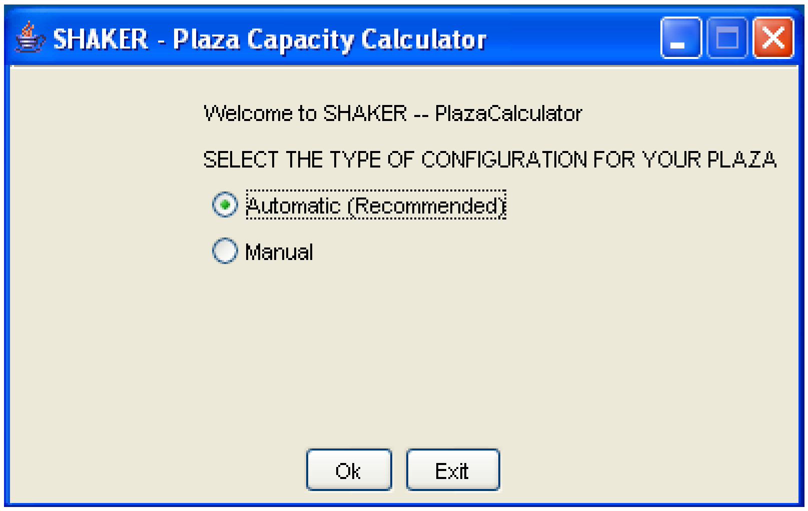

Figure 8: SHAKER opening screen

When using the Automatic option, which is the only portion of SHAKER used for this research, the user is taken to an input screen where has the option to enter the following (as also shown in Figure 9):

1. The number of lanes
2. Toll amount
3. Select if open road tolling or not. This is for the ETC lanes only
4. Enter the total number of arriving vehicles
5. Percentage of users in each of the following categories:
    a. (E_p) two axel vehicles using ETC lanes
    b. (E_t) 2+ axel vehicles (trucks) using ETC lanes
    c. (A) ACM vehicles
    d. (M) Two axel vehicles using manned lanes

e. (T) 2+ axel vehicles (trucks) using manned lanes

Here it is assumed that E-lanes can allow Ep and Et vehicles, A-lanes allow Ep, Et and Auto vehicles, M-lanes allow all type of vehicles.

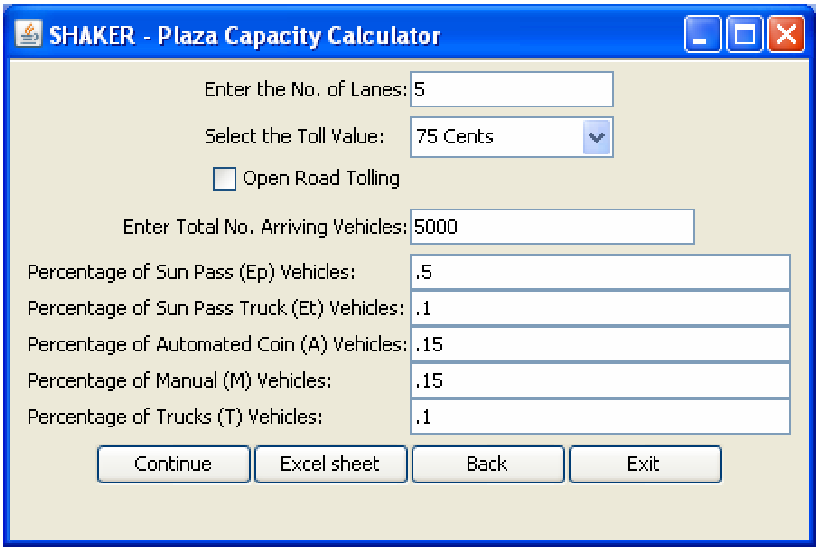

Figure 9: SHAKER Automatic Plaza Configuration Input Screen

After entering these values and when the user clicks 'Continue', the plaza calculator performs its simulation by finding all the possible configurations for the number of lanes entered and the user now is presented with an iconic representation of volumes and capacity per lane (as provided in Figure 10). In this step SHAKER calculates the throughput of all these configurations and compares them to find the best configuration and displays the results for the best configuration. The user can change the number of arriving vehicles or vehicle percentages in this window and press *getThrp* button. This will give the best configuration for updated values. The throughput and capacity is displayed in the *Best Configuration* section located in the lower left section of the window. The throughput, capacity, and queue length of each individual lane is displayed when the mouse is hovered over each lane icon, as shown in Figure 11. The throughputs for each vehicle class and payment type are color coded in this dropdown menu. The throughput for manual paying passenger cars is displayed in yellow and manual paying trucks is displayed in black. The capacity of that particular toll booth is displayed in green and all other overall toll plaza characteristics, such as queue length and time to dissipate queue, are listed below that.

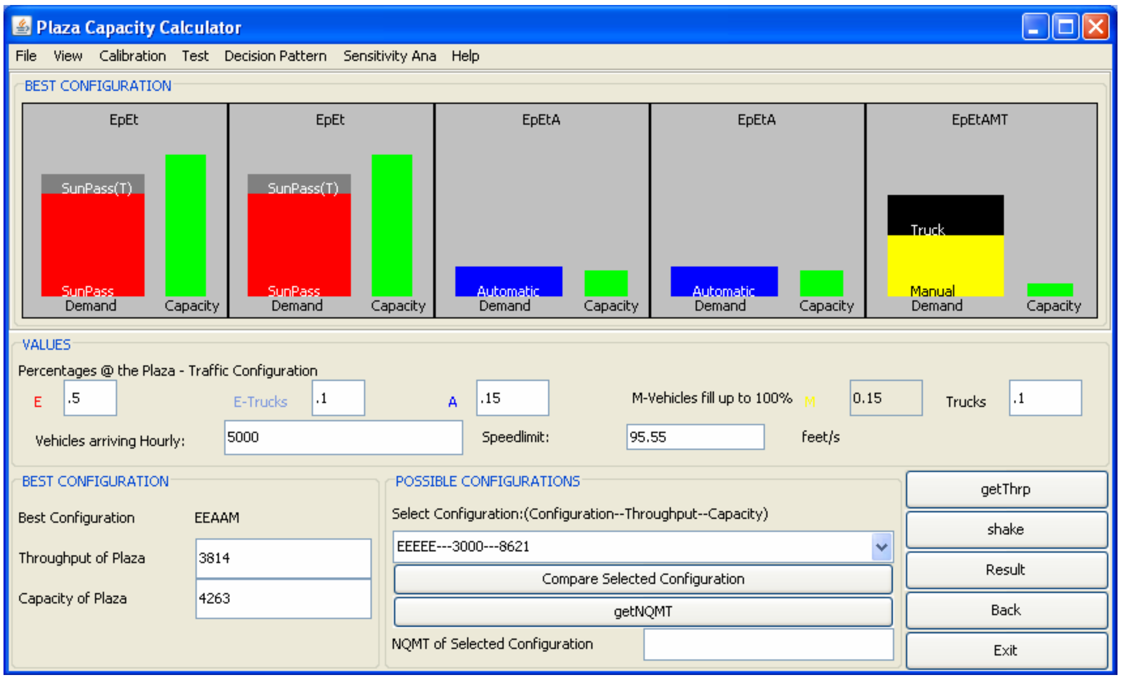

Figure 10: SHAKER Best Configuration Window

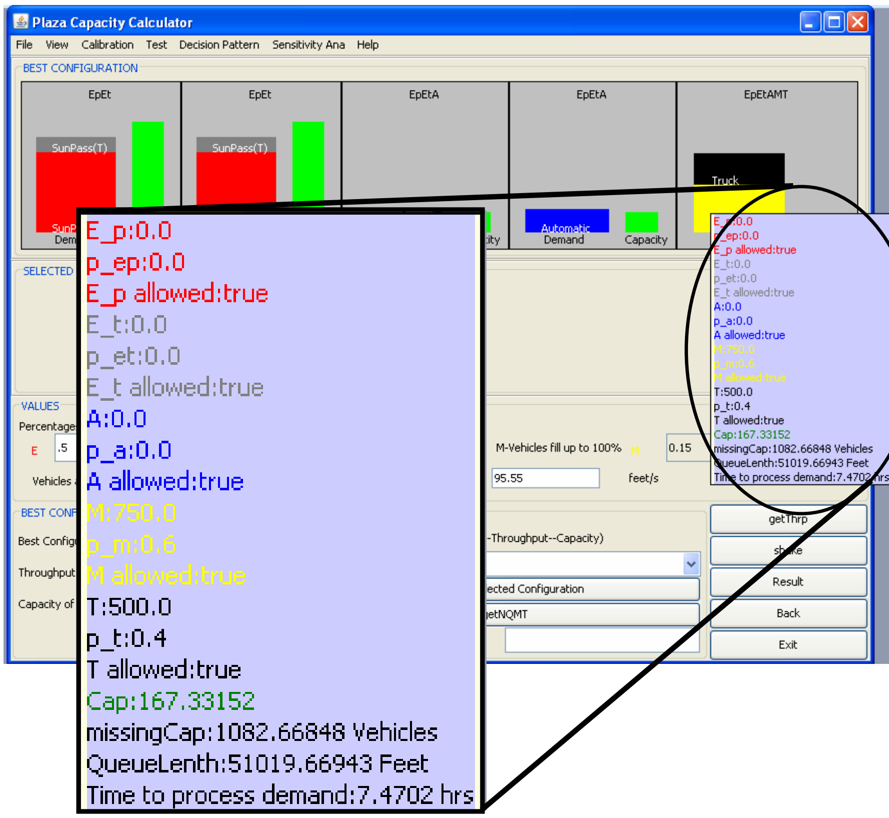

Figure 11: Individual Lane Analysis

To check all possible configurations of the toll plaza the user can select any configuration of lanes from the *Select Configuration* pull down menu. Once the new configuration is selected SHAKER displays the throughput and capacity of each configuration. The user is given the option to compare the original and new configuration by graphical representation of both the best and the selected configurations so that the user can compare the two simultaneously. An example of one such comparison is shown in Figure 12.

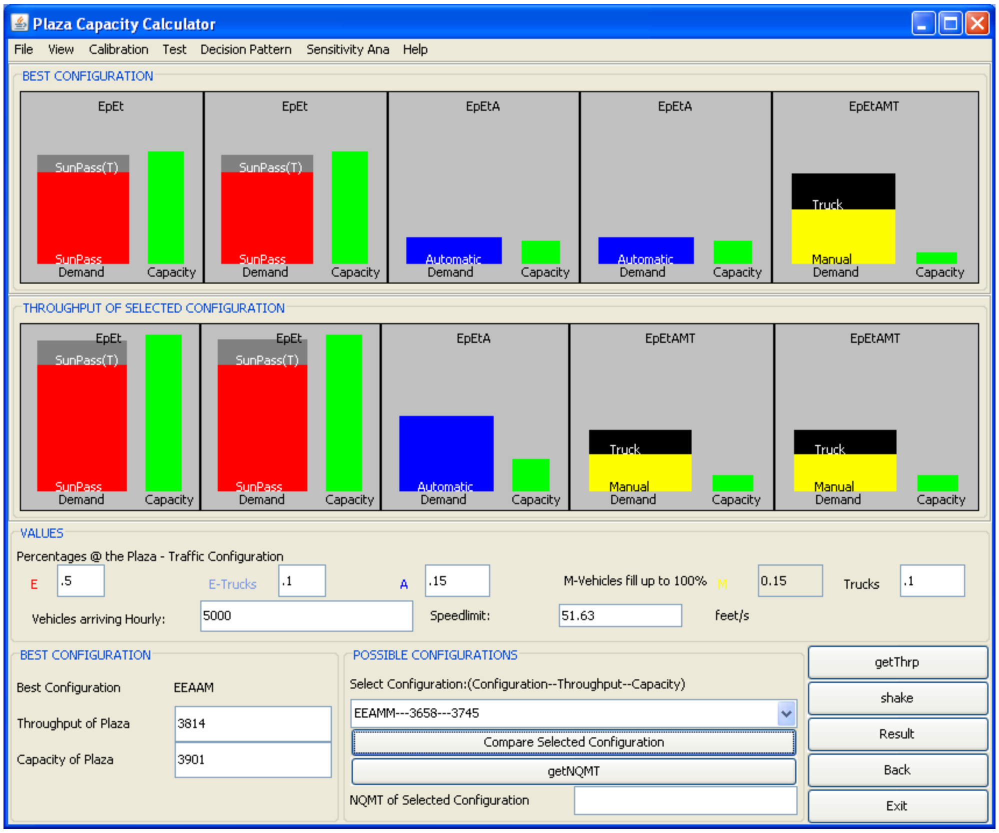

Figure 12: SHAKER - Comparing Configurations example

The SHAKER model also provides the user with the option to conduct the following list of commands:

- SHAKE- If the user wants to monitor the shaking process, he or she can select the configuration and click initialize from the menu Test and click *shake*. The shaking process can be done until the throughput of the plaza becomes stable.
- NQMT- After calculating the plaza, if the user wants to calculate the NQMT (No Queue Maximum Throughput) i.e., maximum throughput of the plaza until a queue is formed in at least one lane, (one can select the configuration from the dropdown menu in automatic

configuration) click *getNQMT* button which will show the NQMT in the capacity of plaza textbox with its label changed and the graph is shown in selected configuration box.

- Calibrate- Calibration is changing the basic properties to meet observed/desired throughput rates of pure lanes. Sets of basic properties can be saved and opened with save Calibration and open Calibration in the Calibration menu in both manual and automatic configurations.

To change basic properties of the calibration the user can alter any number of the provided parameters in this window. Under the sub menu item Customer Type, the user can choose a particular customer type, after which a dialog box appears. The text fields are filled with default values which the user has the option to change. Figure 13 is provided to show this window.

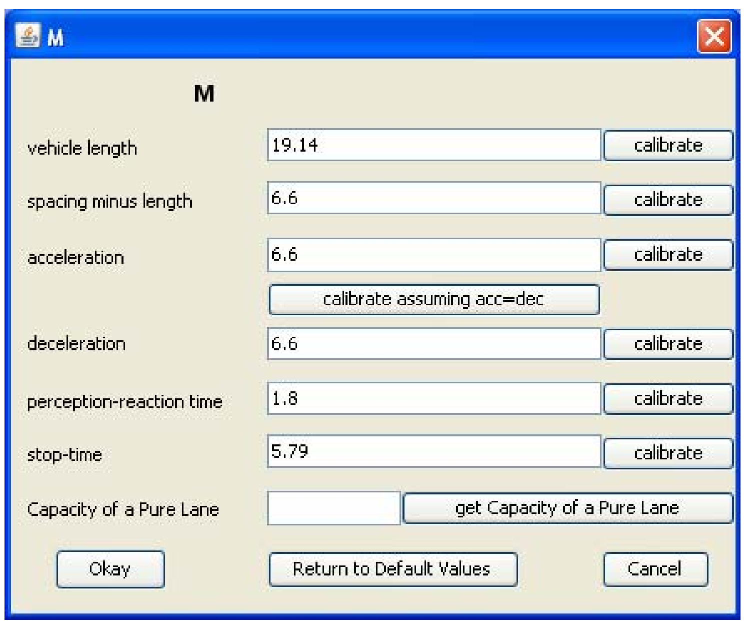

Figure 13: SHAKER - Calibration Window Example

The customer type dialog box also helps the user calibrating the basic properties to meet observed/desired throughput rates of pure lanes. The *getCapacity* button calculates the capacity of a pure lane using the basic properties defined on the sheet and enters it in the capacity text field below. Users may change this value according to their field measurements of the processing rates (vehicles processed per hour) in their region. The calibrate buttons search for a reasonable value of a particular basic property such that the calculated capacity of a pure lane (vph) matches the capacity placed in the capacity text field. All other basic properties stay constant. In the case of a pure *A* lane, pure *M* lane, or pure *T* lane, calibrate buttons appear next to all properties. There is also a calibrate button assuming the vehicle acceleration and deceleration rates are equal. In the case of a pure electronic customer lane, (E_P pure lane or an E_T pure lane), the capacity is only defined by the perception-reaction time and vehicle length. Therefore, all calibration buttons only appear near these two properties. Because the other basic properties influence the mixed lane behavior it is recommended to make each parameter consistent with the basic properties of the corresponding non ETC customer types.

7.2 SHAKER Calibration Steps

The calibration of SHAKER was divided into several steps. First, a particular toll plaza was coded in SHAKER and a measure of effectiveness (MOE) was selected to serve as the index of comparison. Second, an initial evaluation was conducted with SHAKER's default parameter values. Third, if the selected measure of effectiveness was different in simulated and real conditions, an examination of the key parameters was conducted and calibration parameters were determined. Multiple runs with different values of the key parameters were run by trial and error until the calibration part is completed. Fourth, as for the validation part, different toll plazas were coded in SHAKER and the field observed MOE was compared to the simulation MOE.

The overriding assumption for calibrating models is based on simplifying fixed parameters as much as possible. Fixing parameters and/or constraining them to certain intervals helps address the calibration process. Usually the average length of a vehicle type and distance between standing vehicles can be measured rather simply and precisely so that they can be assumed constant in the calibration process. Furthermore, it is a reasonable assumption that manual vehicles, *M*, and passenger cars with transponders using the manned booth lane, E_P, have the same acceleration, *a*, and deceleration, *d*, properties, average length, *l*, and distance between standing vehicles, *b*. The same is true for trucks, *T*, and trucks with transponders using the manned lanes, E_T. The driver's reaction time, t_R, is assumed equal for all customer types. Electronically paying vehicles have no time to pay, $t_{stop} = 0$. ACM users, *A*, should have the same properties as manuals, *M*, except the time to pay, $t_{stop}A$, varies *(2)*. SHAKER was initially calibrated and validated on the OOCEA network and the resulting key calibration parameters are presented in Table 12 and are used by SHAKER as default values.

Table 12: Initial calibration Parameters of SHAKER (31)

	M	T	A	E_P	E_T
l_X =Average vehicle Length (meters)	5.8	21	5.8	5.8	21
b_X = Distance between vehicles (meters)	2.0	3.0	2.0	2.0	3.0
a_X = Vehicles' Acceleration (meters/second^2)	2.0	0.25	2.0	2.0	0.25
d_X =Vehicles' Deceleration (meters/second^2)	2.0	0.25	2.0	2.0	0.25
T_R = Drivers' reaction Time (seconds)	1.8	1.8	1.8	1.8	1.8
$t_{stop\ X}$ = Stop–Time at payment (seconds)	1.5	4.7	0.075	0.0	0.0

To calibrate SHAKER, first the capacity was selected as the measure of effectiveness of the model, hence field observed capacity was compared to the model estimated capacity. To measure capacity in the field, the FHWA recommends observing locations where queues persist for at least 15 consecutive minutes and then measure the flow rate at the point where the queue discharges. The resulting flow rate is the field-measured capacity *(7)*. Therefore, only periods under queuing conditions were used in the calibration and validation process (see Table 9 and Table 10).

7.2.1 Step 4 Initial Evaluation of SHAKER

Second, since the SR-528 toll plaza had all possible lane types in use the initial evaluation of SHAKER was implemented using SR-528 toll plaza, shown in Figure 1. Next, SR-528 was coded in SHAKER, the simulation was run and the simulated toll plaza capacity was determined. The default key parameters used in the first run are shown in Table 12. Field observed capacities were determined for 14 periods (15 minute-intervals) for the manual pay lanes and 5 periods (15 minute-intervals) for the ACM lanes (see Table 9). As shown in Table 14, the initial simulated capacities and the field observed capacities are significantly different (manual lane *p*-value=4.9E-17, ACM *p*-value=2.7E-05).

7.2.2 Step 5 & 6 SHAKER Calibration and Error Checking

After investigating key parameters of SHAKER, it was determined that the model bases the capacity of a lane on the combination of 5 parameters; they are: vehicle length, spacing, acceleration and deceleration, perception-reaction time, and stop-time. In agreement with literature it was determined that of the possible parameters the stop-time was the variable that would vary the most from toll plaza to toll plaza. Therefore, to calibrate the SHAKER model all variables except the stop-time were preset from location to location for each lane group. It should be noted that when using the field data for parameter estimation only the values from the selected periods of queuing should be used for calibration. To obtain the calibration parameters for stop-time, the stop-time mean and mode of only queuing periods were calculated. The approach to use these statistics was based on that the stop-time mode is related to planning evaluation and the stop-time mean is related to operational analysis. Run2 and Run3 were conducted with adjusting stop time (or processing time) while keeping all other parameters fixed. Table 13 shows the parameter values used in each run. For Run2, where the stop time parameter was adjusted using the average field measured stop time, the difference between the modeled and field observed capacities for the manual and ACM lanes were still statistically significant (*p*-value=2.3E-11, *p*-value=0.09 respectively, see Table 14. In Run3, where the mode field measured stop-time was used, the difference between the modeled and field observed capacities for the manual and ACM lanes were not statistically significant (*p*-value=0.315, *p*-value=0.181 respectively; see Table 14). However, also shown in Table 14, the mean relative errors were 2.91% and 6.35% for the manual and ACM lanes in that order.

It was then determined that not only the stop-time parameter, but all 5 parameters should be reevaluated with the extracted data. In earlier calibrations of SHAKER *(2)* all parameters were fixed and estimated capacities were calibrated by forcing the stop-time to result in capacities that matched field data observed by Zarrillo in 1998 *(30)*. It was then proposed to use field data for not only capacities, but for stop-time, reaction times, acceleration and decelerations. First, erroneous decimals in the averages spacing and vehicle lengths were rounded to the nearest whole number to ease future use. The code was originally written with metric units and when converted to SI units superfluous decimals were uncovered. The driver-reaction time was then reconsidered. By nature, the reaction time is difficult to precisely measure, but based on video observation analysis the reaction time was estimated to average 1 second instead of the preprogrammed default of 1.8 seconds. Next, new acceleration and deceleration values were calculated using the field observed inter-arrival times. To determine the time needed for acceleration and deceleration the SHAKER code assumes that the driver accelerates for half the spacing distance and then decelerates for the remaining half. In the field collected data the inter-arrival time was used in conjunction with the linear equation of motion under uniform acceleration to determine an appropriate value (distance traveled equals one half the acceleration times elapsed time squared, $d=\frac{1}{2}\alpha\Delta t^2$) and solved for acceleration, *a*. To use this equation the average inter-arrival time was found. Next, the reaction time was subtracted from the inter-arrival time to give the time when the vehicle is actually moving. This time was then divided by 2 to account for the half acceleration and half deceleration spacing assumption. Next, the linear motion equation was used to calculate the acceleration and deceleration needed to traverse the average vehicle spacing. It was noted that to successfully drive the distance created by the vehicle length and spacing the acceleration and deceleration had to be increased from 6.6 to 9.75 ft/s^2 for passenger cars and from 0.825 to 3.95 ft/s^2 for trucks. Table 13 summarizes the adjusted parameters in Run 4.

Table 13: SHAKER Calibration Parameter Inventory

Run Number	SHAKER Calibration Parameter	Automatic Coin Machine	PC Manual Pay	Truck Manual Pay
Run 1 (Non Calibrated SHAKER)	vehicle length (ft.)	19.14	19.14	69.3
	vehicle spacing (ft.)	6.6	6.6	9.9
	vehicle acceleration, (ft/sec^2)	6.6	6.6	0.825
	vehicle declaration, (ft/sec^2)	6.6	6.6	0.825
	driver reaction-time, (sec.)	1.8	1.8	1.8
	toll stop-time, (sec.)	0.075	1.475	4.68
Run 2 (Average Stop-time)	vehicle length (ft.)	19.14	19.14	69.3
	vehicle spacing (ft.)	6.6	6.6	9.9
	vehicle acceleration, (ft/sec^2)	6.6	6.6	0.825
	vehicle declaration, (ft/sec^2)	6.6	6.6	0.825
	driver reaction-time, (sec.)	1.8	1.8	1.8
	toll stop-time, (sec.)	5.48	5.78	17.58
Run 3 (Mode Stop-time)	vehicle length (ft.)	19.14	19.14	69.3
	vehicle spacing (ft.)	6.6	6.6	9.9
	vehicle acceleration, (ft/sec^2)	6.6	6.6	0.825
	vehicle declaration, (ft/sec^2)	6.6	6.6	0.825
	driver reaction-time, (sec.)	1.8	1.8	1.8
	toll stop-time, (sec.)	3.50	3.00	11.00
Run 4 (All Measured Field Data)	vehicle length (ft.)	19	19	70
	vehicle spacing (ft.)	6	6	10
	vehicle acceleration, (ft/sec^2)	9.75	9.75	3.95
	vehicle deceleration, (ft/sec^2)	9.75	9.75	3.95
	driver reaction-time, (sec)	1	1	1
	toll stop-time, (sec)	5.80	5.56	11.00

In Run 4, the key parameters were adjusted once again. The resulting difference between the modeled and field observed capacities for the manual and ACM lanes were not statistically significant (p-value=0.467, p-value=0.860 respectively; see Table 14). Moreover, also as shown in Table 14, the mean relative errors decreased to 1.03% and 3.76% for the manual and ACM lanes in that order, which are acceptable errors in simulation (< 5%).

Table 14: Calibration and Verification Statistical Results

Calibration: Errors in SHAKER's Estimation of Toll Lane Capacity

	Manned Lane					ACM Lane				
	Observed Capacity (veh/hr)	Modeled Capacity (veh/hr)	*Average Errors[1]*		T test Statistic	Observed Capacity (veh/hr)	Modeled Capacity (veh/hr)	*Average Errors[2]*		T test Statistic
			% Error	*RMSE*				*% Error*	*RMSE*	
Run 1 (Non Calibrated SHAKER)	355	476	34.22	14,922	4.9E-17	361	618	73.12	73,092	2.7E-05
Run 2 (Mean Stop-time)	355	300	-15.55	3,234	2.3E-11	361	321	10.10	1,396	0.090
Run 3 (Mode Stop-time)	355	365	2.91	959	0.315	361	360	3.37	477	0.181
Run 4 (All Measured Values)	355	358	1.03	158	0.467	361	360	0.85	360	0.860

Validation: Errors in SHAKER's Estimation of Toll Lane Capacity

	Manned Lane Analysis					ACM Lane Analysis				
	Observed Capacity (veh/hr)	Modeled Capacity (veh/hr)	*Average Errors[1]*		T test Statistic	Observed Capacity (veh/hr)	Modeled Capacity (veh/hr)	*Average Errors[2]*		T test Statistic
			% Error	*RMSE*				*% Error*	*RMSE*	
SR 528 Validation Group	366	355	-2.88	549	0.169	354	360	1.26	667	0.5986
SR 417 Validation Group	362	367	1.78	268	0.366	--	--	--	--	--
SR 429 Validation Group	400	414	3.76	596	0.611	--	--	--	--	--
SR 91 Validation Group	202	208	3.22	43	0.131	--	--	--	--	--

[1,2] Although very similar, mean relative Average Errors are calculated by first taking the errors associated with each individual test period then calculating the run average, it is not the error in final Capacity Averages
-- ACM lanes were not in use at these locations

7.2.3 Step 7 SHAKER Validation

Next, for the validation process, SR-528, SR-417, SR-429, and SR-91 were coded in SHAKER. The stop time for each toll-lane at each toll plaza was observed in the field and adjusted in SHAKER accordingly. Table 15 summarizes the final field observed stopping time values for each toll plaza, per lane type, and per vehicle type.

As shown previously in Table 14, there were no statistically significant differences between the observed capacities and the simulated capacities for the manned and automatic lanes (SR-528 Manual lane p-value=0.169, ACM p-value=0.5986; SR417 p-value=0.336, SR429 p-value=0.611, SR91 p-value=0.131). Table 14 also shows that the errors are in acceptable ranges (SR528=-2.88%, SR417=1.78%, SR429=3.76%, SR91=3.22%).

Table 15: Stop Times Used for Validation

		Payment Type		
Validation Data Site Location	**SHAKER Calibration Parameter**	**Automatic Coin Machine**	**PC Manual Pay**	**Truck Manual Pay**
SR-528, EB	toll stop-time, (sec.)	5.8	5.56	11.0
SR-417, NB & SB	toll stop-time, (sec.)	*	5.12	12.9
SR-91 FL TPK, SB	toll stop-time, (sec.)	*	12.50	19.0
SR-429, NB & SB	toll stop-time, (sec)	*	4.49	11.0

* ACM lanes were not utilized at these locations

8. THE VISSIM MODEL

VISSIM is a microscopic, time step and behavior based simulation model developed to model urban traffic and public transit operations. VISSIM is a commercially available traffic simulation package developed by PTV AG, Karlsruhe, Germany, and distributed in the United States by PTV America, Inc. The software can analyze traffic and transit operations under user defined conditions, such as lane configuration, traffic composition, traffic signals, transit stops, etc., thus making it a useful tool for the evaluation of various alternatives based on transportation engineering and planning measures of effectiveness *(35)*.

According to the VISSIM User Manual the accuracy of a traffic simulation model is mainly dependent on the quality of the vehicle modeling, e.g. the methodology of moving vehicles through the network *(35)*. In contrast to less complex models using constant speeds and deterministic car following logic, VISSIM uses the psycho-physical driver behavior model developed by Wiedemann in 1974. The basic concept of this model is that the driver of a faster moving vehicle starts to decelerate as he reaches his individual perception threshold to a slower moving vehicle. Since this driver cannot exactly determine the speed of that adjacent vehicle, his speed will fall below that vehicle's speed until he starts to slightly accelerate again after reaching another perception threshold. This results in an iterative process of each vehicle's acceleration and deceleration.

VISSIM simulates the traffic flow by moving "driver-vehicle-units" through a network. Every driver has a specific behavior characteristics assigned to their specific vehicle type. As a consequence, the driving behavior corresponds to the technical capabilities of his vehicle. Attributes characterizing each driver-vehicle-unit can be categorized into three categories, they are: technical specifications of the vehicle, behavior of driver-vehicle-unit, and independence of driver-vehicle-units. More specifically each category includes parameters such as:

- Technical specifications of the vehicle
    - Length
    - Maximum speed
    - Potential acceleration
    - Actual position in the network
    - Actual speed and acceleration
- Behavior of driver-vehicle-unit
    - Psycho-physical sensitivity thresholds of the driver (also known as their ability to estimate thresholds and level of aggressiveness)
    - Memory of driver
    - Acceleration based on current speed and driver's desired speed
- Interdependence of driver-vehicle-units
    - Reference to leading and following vehicles on own and adjacent travel lanes
    - Reference to current link and next intersection

Not every technical specification that VISSIM employs are applicable in toll plaza operations, therefore to reduce model setup and calibration efforts it is important that key specifications be identified as either those that have an impact, or those that do not have an impact on toll plaza modeling. The modeling elements that have a direct effect on toll plaza operations will receive special attention in both the setup and calibration process, while others may not.

8.1 Step 3 Development of VISSIM Model

The process of coding VISSIM consists of a systematic series of programming processes that must be addressed to duplicate an actual traffic situation. Development of a successful model was broken down into three major categories; physical design of the roadway, vehicle characteristics, and driver behaviors. The methodology and process for developing the first two categories is that model characteristics are to remain fixed for all designs while the driver behavior characteristics are reserved as the parameters used for model calibration.

8.1.1 System Layout

To build roadways in VISSIM a series of links and connectors were used to represent the actual geometry of a system. There are two options when deciding to build a system in VISSIM. The decision was whether to build a hypothetical traffic system or to model an existing one. The objective of this research is to use modeling software to recreate toll plaza operations, so it was imperative that the exact geometry of the plaza be represented by the model. However, if there was no reason to suspect lane blockage due to extensive queue lengths or unique upstream lane changing behaviors there should be little difference between modeling a plaza to in field specifications and modeling a generic isolated toll plaza. As mentioned previously by Astarita et al., the capacity of a plaza is also a factor of the approach dimensions and layout. Also mentioned is "When the arrival rate exceeds the corresponding service rate, slowed vehicles directed towards over-saturated booths (usually the manual ones) can cause a cut-off in the flow of other vehicles, which are destined to non-congested booths (usually ETC lanes)" *(5)*. To circumvent the need to study whether or not upstream issues were present at every plaza location and to ensure that the modeling represents reasonable conditions the existing SR-528 Beachline

West toll plaza was simulated using VISSIM. To represent roadways in VISSIM links and connectors must be placed. Links represent a segment of any length of homogeneous (same number of lanes, and width) roadway and connectors were used to connect the multiple segments. Links must be assigned a driving behavior classification of one of the following: urban, freeway, footpath, or cycle path. Links upstream and downstream of the toll plaza were classified as freeway, but links near the plaza were to use the urban settings. Urban settings were used at the toll plaza because they better represent the stopping and queuing conditions that are normally not found on freeways. An aerial image of the Beachline West toll plaza was imported to the background to serve as an overlay to place links (Figure 14). The aerial image was first scaled to match the dimensions embedded into the VISSIM elements.

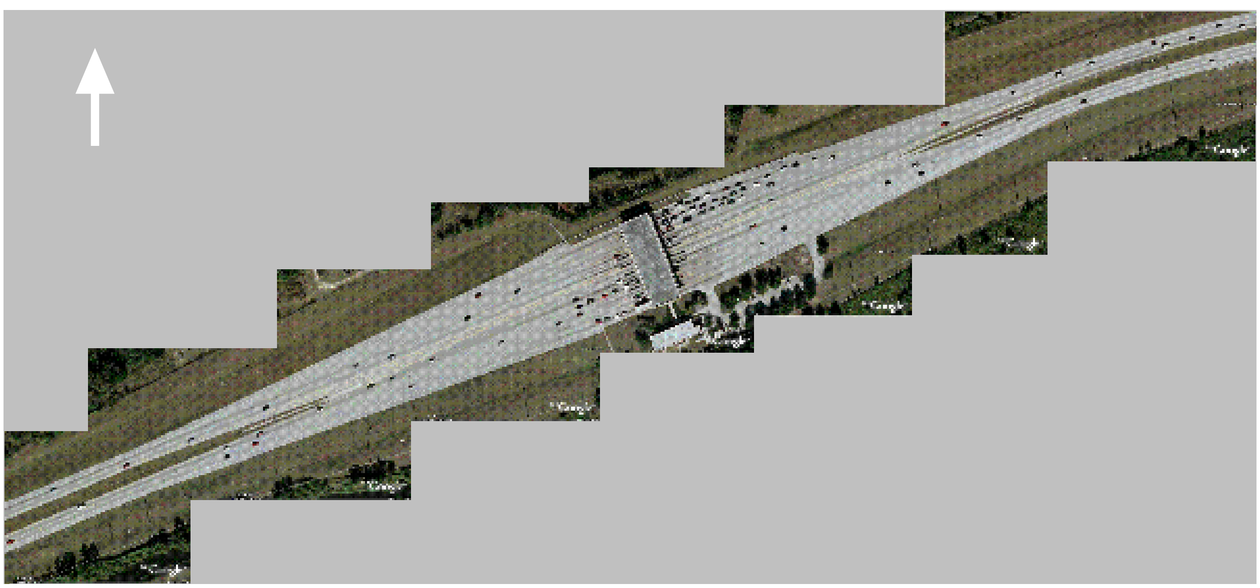

Figure 14: Beachline West Toll Plaza Aerial Image (Source: Google Earth)

Once scaled correctly, the roadway could be laid out by tracing the image with links and connectors until the dimensions and curvatures were correct. When the data was conducted upstream and downstream approaches of SR-528 were two lanes in each direction. The two lanes then gradually turn into 5 lanes in the WB direction and 6 lanes in the EB direction. This was constructed in VISSIM by first laying a 2 lane link and connecting it to a 3 lane link using connectors. This process was repeated until all toll plaza lanes were coded. As it can be seen from the image, there is little storage room for the right most lanes in both directions. This can potentially have a profound effect on the driver interactions if and when the queue length back up into the through lanes. Therefore, in this case it was important to model the roadway dimensions as accurately as possible. The completed plan view of the VISSIM model that represents the toll plaza and approach road segment is shown in Figure 15. When the background image and the roadway links are combined on the same screen, as provided in Figure 16, one can clearly see that the model geometry closely matches that of the existing geometry.

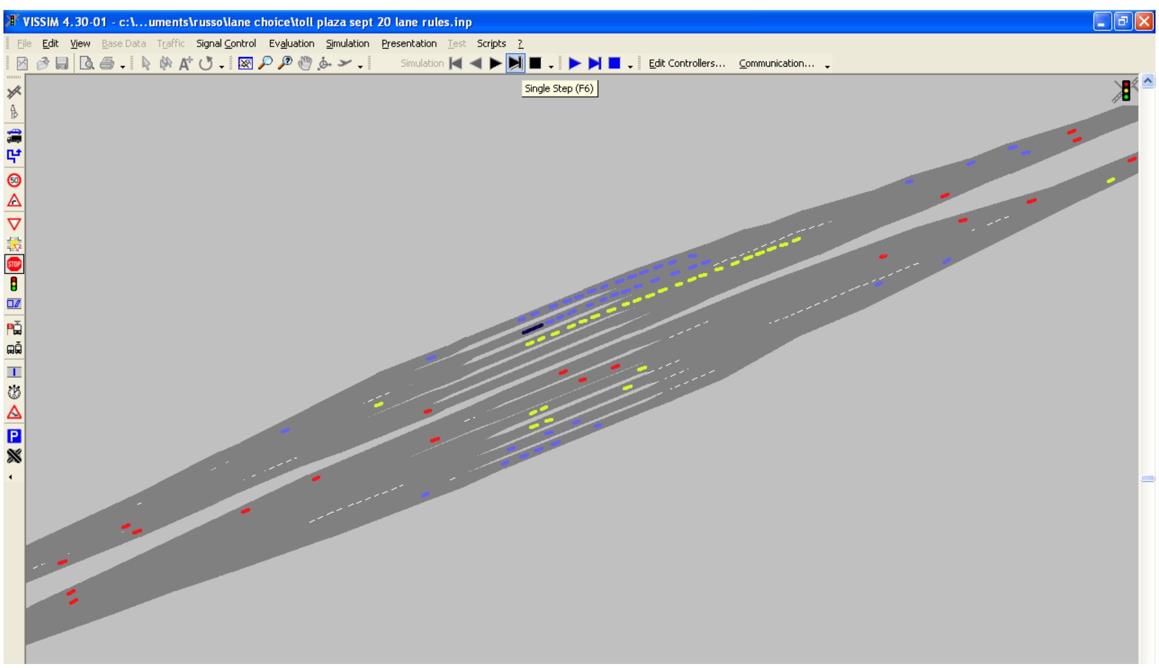

Figure 15: VISSIM Model of SR-528 Toll Plaza

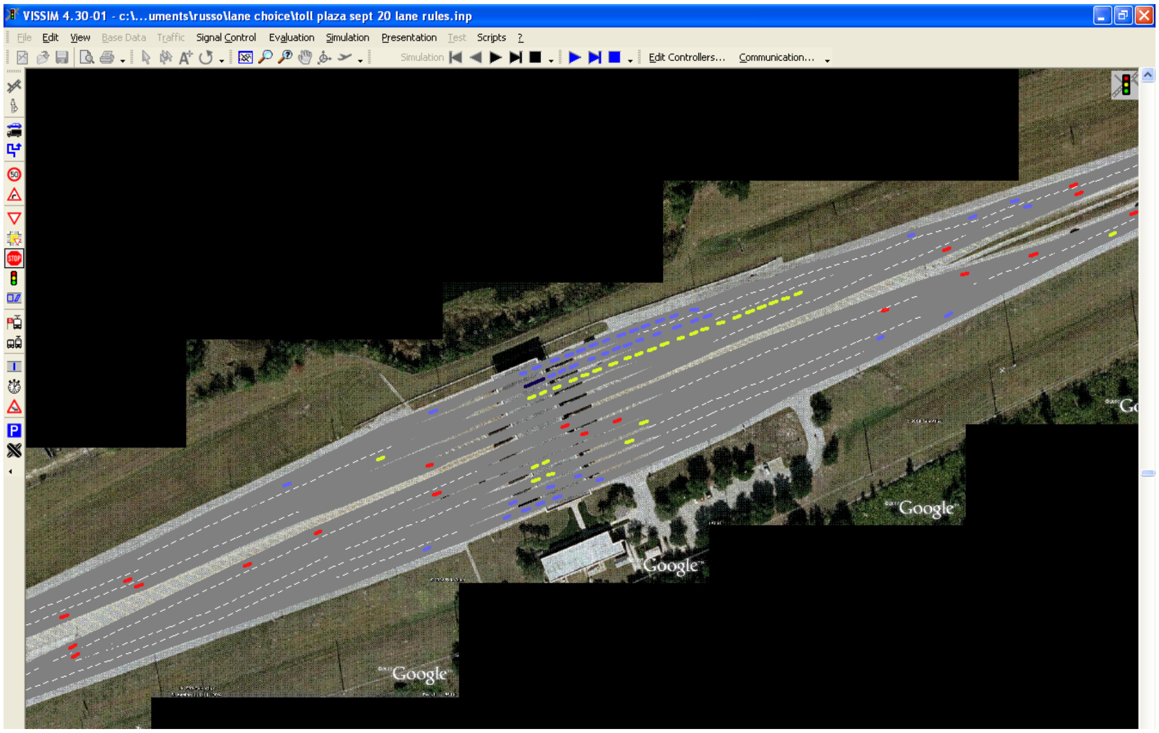

Figure 16: VISSIM Model Overlaid on Aerial Image

8.1.2 Base Data for Simulation

Once the toll plaza and approaches are correctly drawn to match the desired geometry the next step was to define the vehicle characteristics. One major element that makes VISSIM different from SHAKER is its potential to stochastically simulate vehicles through the model. According to the VISSIM Manual, "The stochastic nature of traffic implies the necessity to provide this kind of variability in VISSIM models also" *(35)*. VISSIM implements this stochastic nature by following a set of distributions and functions that represent the variability in driver behavior. Areas in which are modeled by functions and distributions rather than single values are: vehicle acceleration, desired vehicle speed, dwell time, and vehicle type.

8.1.2.1 Acceleration and Deceleration Function

The variability created by acceleration and deceleration was modeled by a function of the current speed of particular vehicle in question. To realistically model change in speed, VISSIM makes use of four functions that together represent maximum and desired acceleration and deceleration. Making the functions stochastic provides the ability to set the minimum, mean, maximum values for each of the functions. The VISSIM default functions were used for each of the acceleration and deceleration categories because there was no more superior information pertaining specifically to toll plazas available.

8.1.2.2 Speed Distribution

The desired speed was modeled using an empirical distribution that is confined by a minimum and maximum. The approach speed is however not predicted to alter the toll plaza capacity. The toll plaza capacity could only be determined once queuing was present, which means every vehicle was required to stop at least once before the final stop to pay the toll. If the speed of the queue was less than the desired speed of the approach the desired speed works only as a bench mark in which vehicles attempt to accelerate and decelerate to. The speed limit of the toll plaza approach decreases to 35 mph, so for model completeness a distribution centered around 35 mph was used here. If not approaching a queued lane, ETC vehicles do not come to a complete stop, but in the field typically reduce their speeds to 35 mph in order to ensure safe passing of narrow toll lanes. To model ETC vehicle deceleration, 35 mph speed reduction zones were placed 500 feet upstream of the toll booths. These zones forced ETC vehicles to temporally reduce their speeds until 35 mph was obtained. Once through the speed reduction zones, simulated vehicles accelerate back to original speeds. A window showing an example of reduced speed vehicle characteristics is provided below in Figure 17.

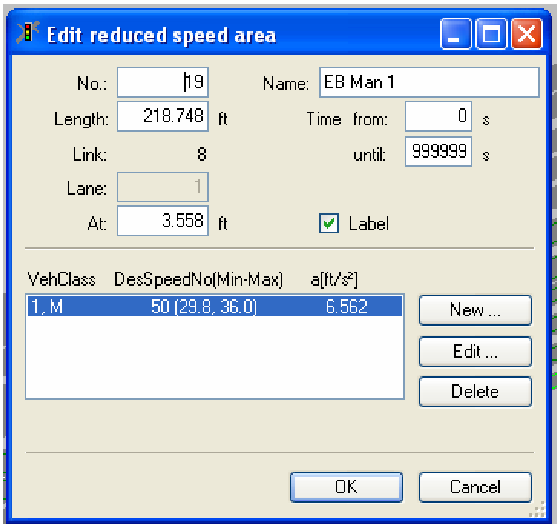

Figure 17: VISSIM Window - Editing Reduced Speed Areas

8.1.2.3 Dwell Time Distribution

The distribution that has the potential to impact the toll plaza model the most is the stopping dwell time. The dwell time distribution utility is typically used in VISSIM to represent the dwell time of parking vehicles, transit wait times, or stopping due to unsignalized intersections. The stopping maneuver observed at a stop sign is similar enough to the maneuver at toll booths so it can be used to represent the operation and act as a dispatch counter. Virtual stop signs were placed at the toll collection location to model the deceleration and stopping maneuver experienced by each driver who chooses to use a non-ETC lane. Attached to each stop sign is the option to designate unique dwell times for each vehicle type and payment method. An example of the VISSIM window where the stop sign dwell times are assigned is provided in Figure 18.

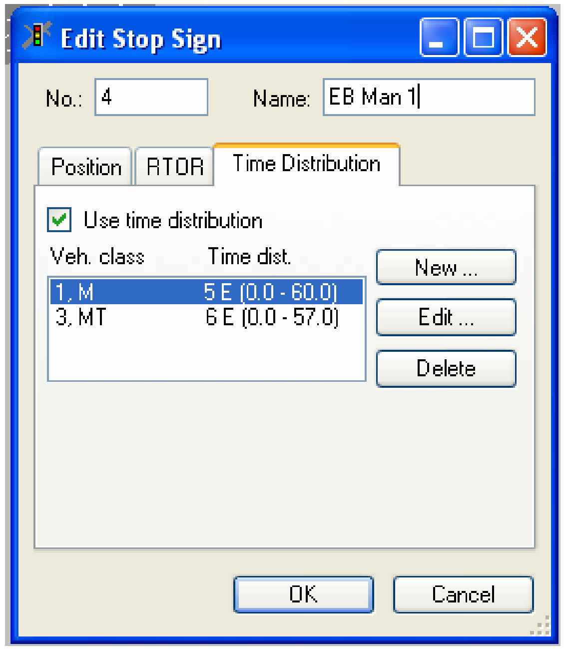

Figure 18: VISSIM Window - Editing Stop Signs

Dwell times can be defined as either normally distributed or empirically distributed. To determine the dwell time the processing time frequencies from the collected data was plotted in a histogram. Three categories are designed to show the differences between vehicle types and payment types. The categories were passenger car using the manned lane, trucks using the manned lanes, and passenger cars using the ACM lanes. The cumulative percentages of the frequencies were then obtained and used to represent the dwell time distributions for each category in VISSIM. The cumulative percentage graph obtained from the data collection is referenced when inserting dwell times distributions in VISSIM. The graphs obtained from the manually collected data and user defined distributions in VISSIM are provided in Figure 19 through Figure 24. It should be noted that the curves in each graph do not visually match because the maximum processing time value used in the spreadsheet generated graphs are reduced to better show the distribution trend. The maximum processing times used in the

VISSIM distributions are the actual observed maximum times while the maximum shown in the histogram charts only represent the 97%.

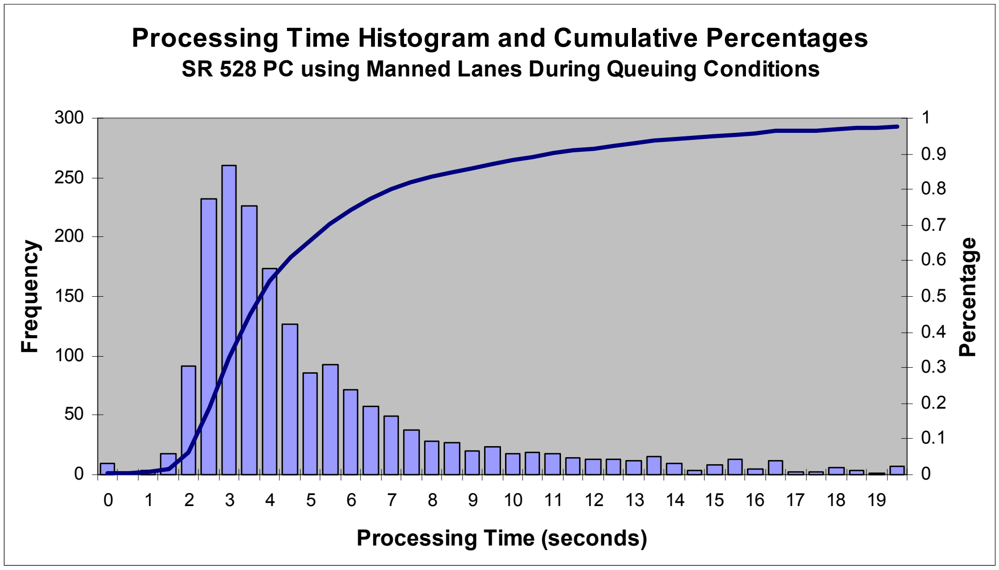

Figure 19: Processing Time Histogram for SR-528 Manned Lanes

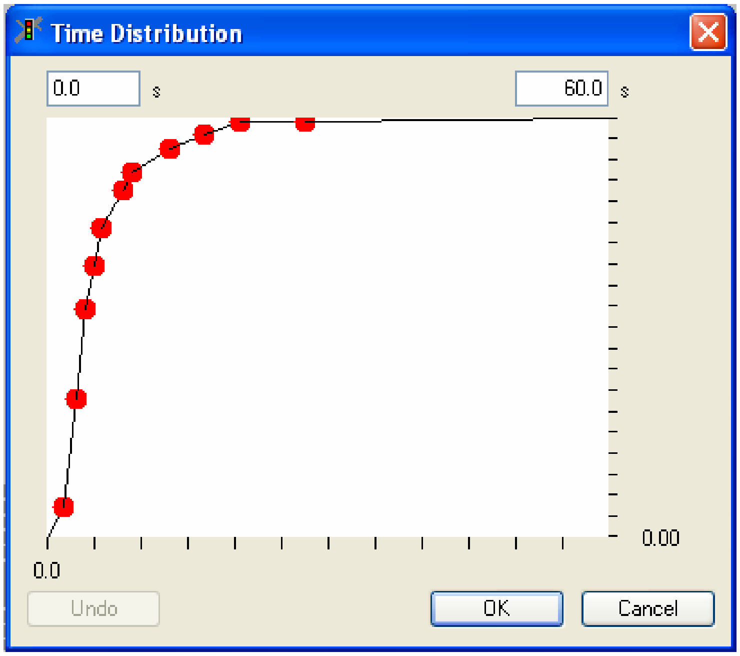

Figure 20: VISSIM coded Dwell Time Distribution for Manned Toll Lanes

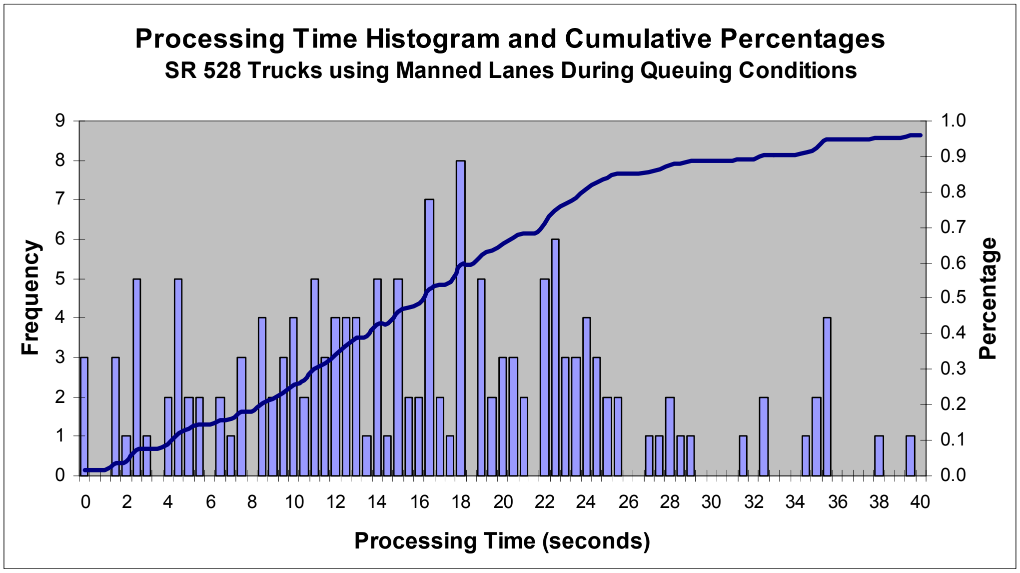

Figure 21: Processing Time Histogram for SR-528 Trucks

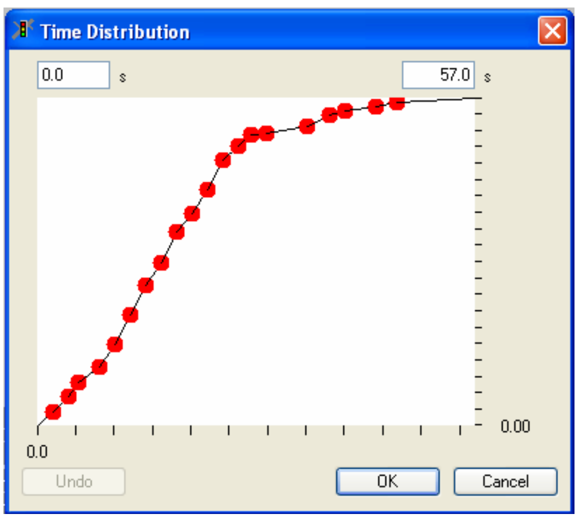

Figure 22: VISSIM coded Dwell Time Distribution for Trucks in Toll Lanes

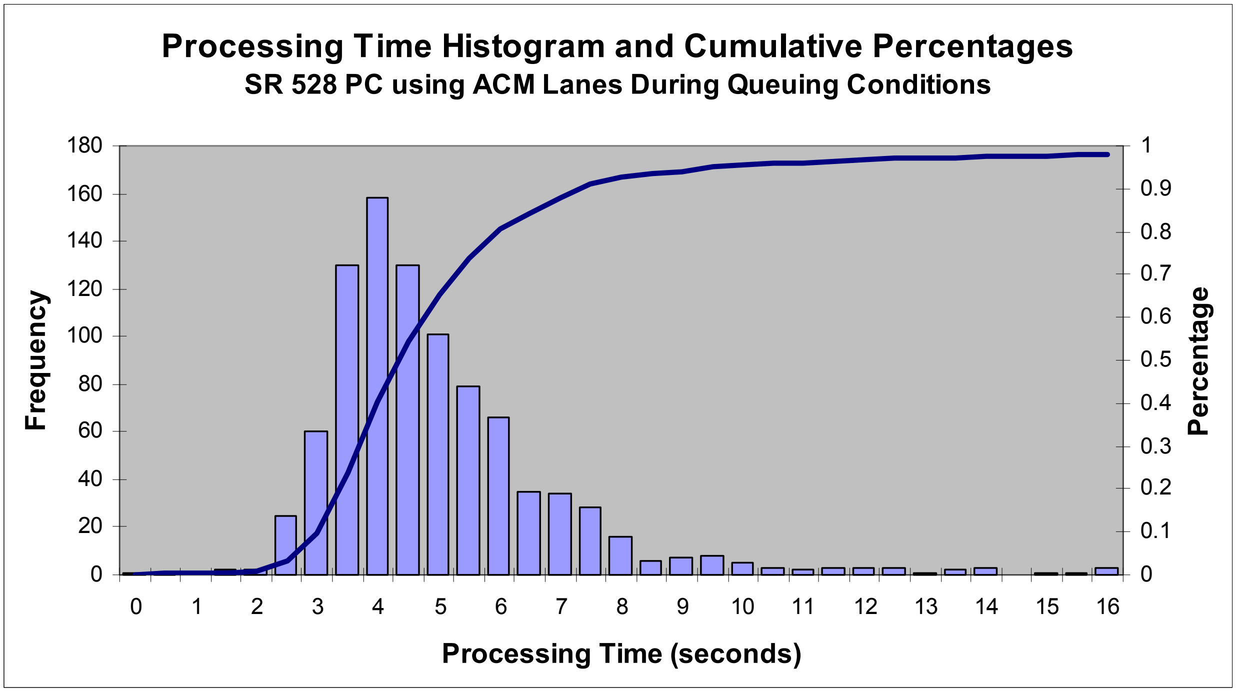

Figure 23: Processing Time Histogram for SR-528 ACM toll lanes

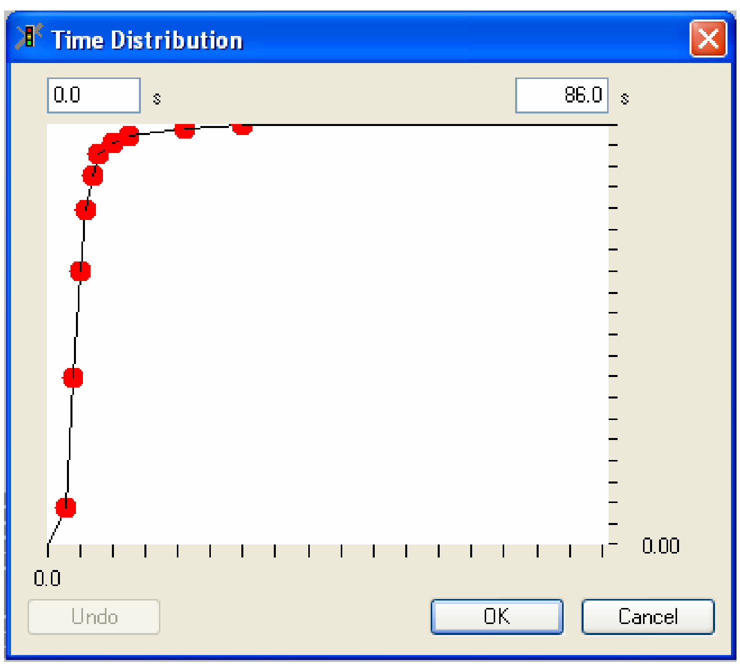

Figure 24: VISSIM coded Dwell Time Distribution for ACM Toll Lanes

8.1.2.4 Vehicle Classification

The next step in developing the model was to create the traffic composition that will be used to test the network. First the different vehicle types had to be defined. For this research five vehicle types were classified to represent the five distinctive payment methods (see Figure 25). Within each vehicle type are two vehicle classes to represent the different vehicle properties of passenger cars versus large trucks. The same set of vehicle properties was assigned to manned, ACM, and ETC lane passenger cars and another set is assigned to manned and ETC lane truck users.

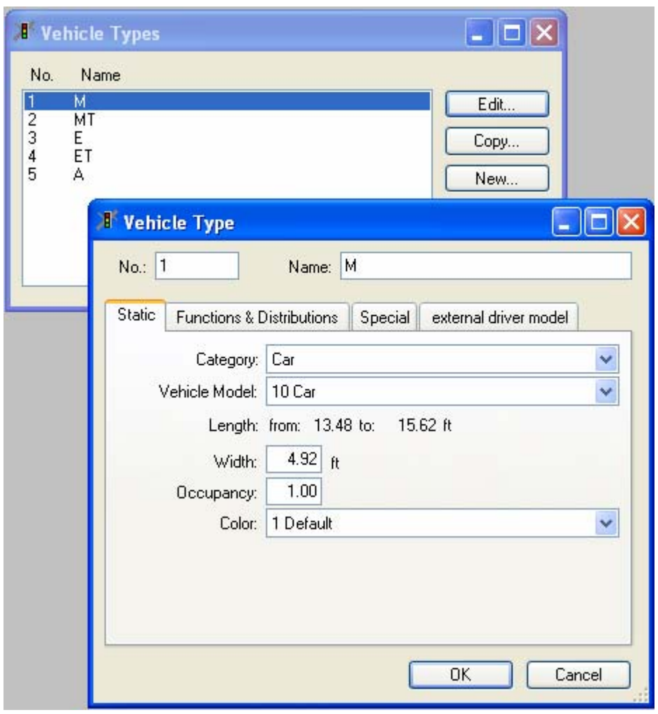

Figure 25: VISSIM Window - Vehicle Type Classification

Next the traffic composition of each lane class was defined. The traffic composition window provided in Figure 26 allows the user to insert the relative flows of each lane class and the desired speed distribution for each vehicle type. The percentage of the flows was determined from the field collected data. The desired speed is an estimation of approach speed upstream of the plaza. The initial desired speed has minimal effects on the plaza operations because speed reduction zones will be implemented near the toll booths. Speed reduction zones forced the incoming traffic to decelerate until their current speed matches that of the temporary speed reduction. In accordance with the 35 mph speed limit near toll plazas, vehicles speeds were reduced to speeds ranging from 25 mph to 50 mph. However, once queues were present in the non-ETC lanes, the reduced speed would be used only as a benchmark that queuing vehicles attempt to obtain but will never achieve. Reducing the speed will have little impact on the plazas capacity because vehicles will ultimately be governed by other factors, such as, dwell times, queuing velocity, and driver behaviors.

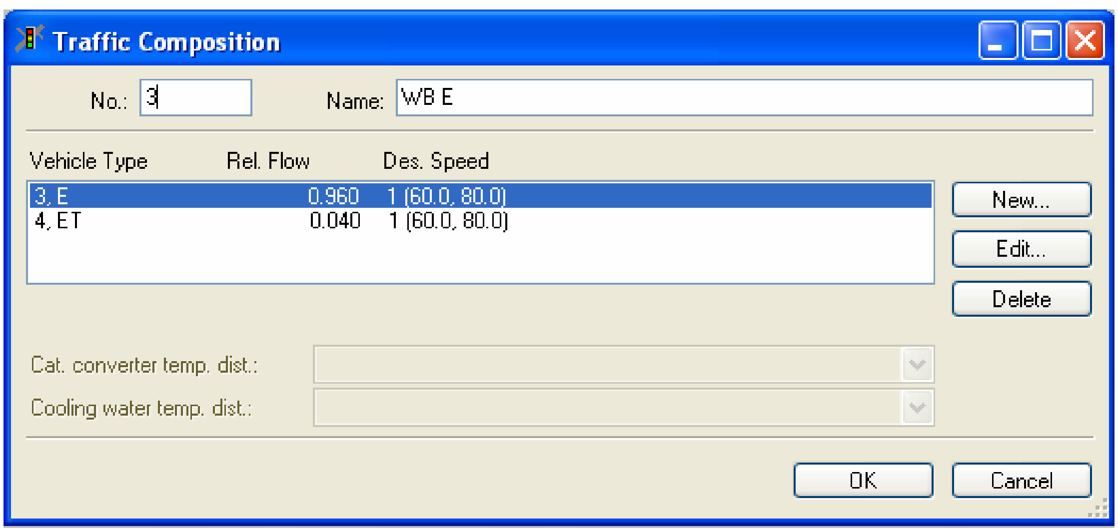

Figure 26: VISSIM Window - Traffic Composition

The vehicle types and distributions have all been entered at this point but no actual volumes have been inserted into the model; that step is next. Assigning traffic volumes was the last step involved in creating the physical attributes of the traffic because volume is the most unstable property observed. In the *Vehicle Input* window, shown below in Figure 27, the user inserts the entering vehicle volumes in vehicles per hour. VISSIM then simulates the vehicles into their respective links by means of a Poisson distribution. Vehicles can fill the network using either exact values or stochastic volumes. When exact volume is enabled VISSIM generates exactly the edited number of vehicles to enter the network as opposed to a distribution. The volumes inserted in this step matched the peak hour volumes observed in the data preparation section and used in the SHAKER calibration. Initially input volumes were slightly overestimated to ensure that queuing conditions occur in each of the ACM and manned lanes.

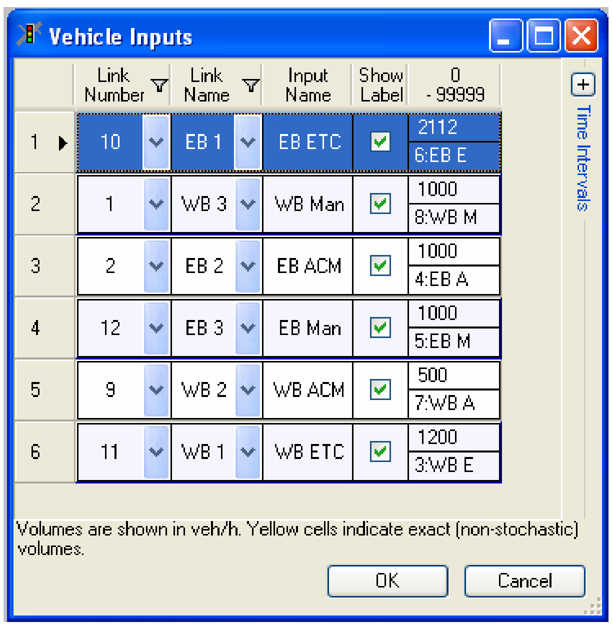

Figure 27: VISSIM Configuration Window - Vehicle Inputs

The vehicle inputs from the previous step indicate that three different originating volumes, each pertaining to a particular payment lane, were used to generate just one approach's traffic. Splitting the approach volumes into lane classes allows for the user to assign certain traffic to a particular route that represents the typical lane choice decisions one might use to arrive at their final lane decision. In VISSIM a route is a fixed sequence of links and connectors that represents the path of a particular vehicle class. There are two available methods for vehicle routing assignment; static and dynamic. Static assignment routes vehicles from a start point to any of the defined destinations using a static percentage for each destination. Dynamic assignment routes traffic by referencing multiple origin-destination matrices on the idea of iterated simulation. That means a modeled network is simulated not only once but repetitively and the drivers choose their routes through the network based on the travel cost they have experienced during the preceding simulations. For purposes of modeling an isolated toll plaza the simpler static routing was sufficient. To begin, the static route originates at a user defined routing decision point and is usually placed at the beginning of the initial link for the route. Each route decision must then be matched to at least one destination point. The user can select and assign the vehicle classes that will be affected by each routing decision. Figure 28 is provided to show an example of how the routing is created along the network. In this example the yellow line represents the practical lane choices that a manually paying driver may use. Routes were then defined for each of the other payment lane types.

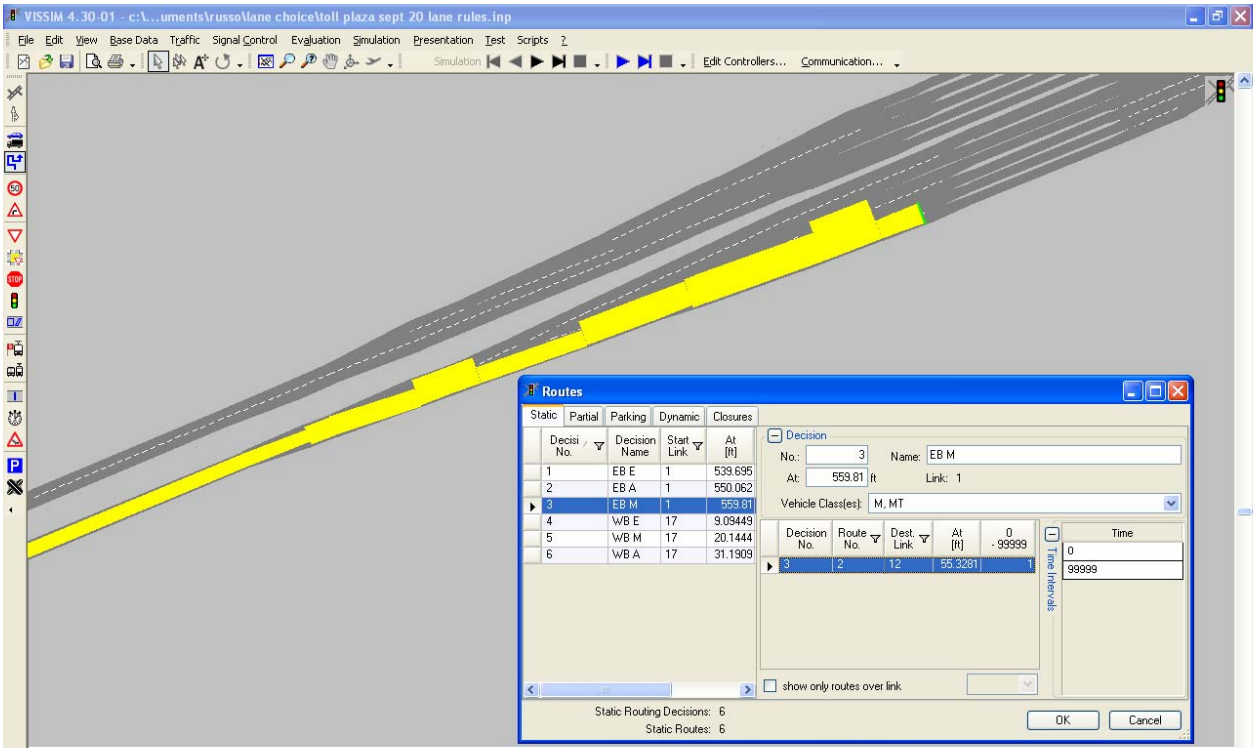

Figure 28: VISSIM Calibration Window - Route Design

8.1.2.5 Driving Behavior

Now that the links are drawn to scale, the routes are defined, and the vehicles classifications and distributions are generated, the complete original model is developed. A functioning model means that without adjusting any additional options the model can successfully simulate some degree of toll plaza operations. The initial model simulation may display correct operational movements but not successfully predict the correct throughput and capacity. The preliminary model will first undergo initial evaluation to determine its performance level. If the model is successful in simulating observed conditions then the model requires no further attention and is considered calibrated. If the model fails to predict field conditions the only category left to adjust is the driving behavior parameter sets. The driving behaviors govern the range of parameters and rules of the car following and lane change models in VISSIM. The driver behavior parameters make up the traffic flow model based on the continued work of Wiedemann. "The referenced traffic flow model is a discrete, stochastic, time step based, microscopic model with driver-vehicle-units as single entities. The model contains a psycho-physical car following model for longitudinal vehicle movement and a rule-based algorithm for lateral movements" *(35)*.

The basic idea of the Wiedemann model is the assumption that a driver can be in one of four driving modes: free driving, approaching, following, or braking. To VISSIM User Manual describes the Wiedemann use of the four driving modes best *(35)*:

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

The driver behavior parameters that make up the psycho-physical car-following model were broken down into four behavior sub categories, they are: following behavior, lane change behavior, lateral behavior, and signal control behavior. For the purpose of modeling exclusive toll plaza operations only the *following behavior* parameters were visited. Within the following behavior adjustment options the user has two choices on which car following model to select; the Wiedemann 74 or the Wiedemann 99. The Wiedemann 74 model is mainly suitable for urban traffic. The Wiedemann 99 model is mainly suitable for interurban (motorway) traffic. It is unknown which model would better suit toll plaza operations, therefore, each model's performance will be evaluated.

The Wiedemann 74 model is based on the following parameters:

- Average standstill distance (*ax*) defines the average desired distance between stopped cars. It has a fixed variation of ± 1m.
- Additive and multiplicative part of desired safety distance (*bx_add*) and (*bx_mult*) affect the computation of the safety distance.

The distance *d* between two vehicles is computed using the following formula:

$d = ax + bx$

where,

- *ax* is the standstill distance
- $bx = (bx\_add + bx\_mult * z) * \sqrt{v}$
- *v* is the vehicle speed
- *z* is a value of range [0,1] which is normally distributed around 0.5 with a standard deviation of 0.15. This parameter is automatically determined by the stochastic nature of the car following model.

The setup of the Wiedemann 74 driving behavior parameter set is provided in Figure 29.

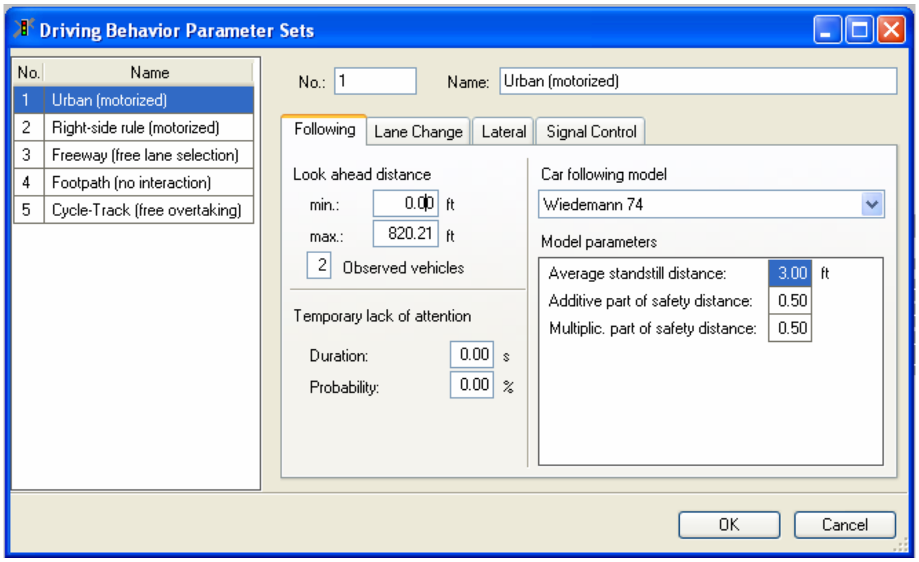

Figure 29: VISSIM - Driving Behavior Parameter Set Window - Wiedemann'74

The Wiedemann 99 model is a much more extensive and complex model. Within this model ten parameters are available for adjustment. A brief explanation of each of the parameters is as follows *(35)*:

- CC0 (Standstill distance) defines the desired distance between stopped cars. It has no variation.
- CC1 (Headway time) is the time (in s) that a driver wants to keep. The higher the value, the more cautious the driver is. Thus, at a given speed the safety distance *dx_safe* is computed to: *dx_safe = CC0 + CC1 * v*. The safety distance is defined in the model as the minimum distance a driver will keep while following another car. In the case of high volumes this distance becomes the value with the strongest influence on capacity.
- CC2 ('Following' variation) restricts the longitudinal oscillation or how much more distance than the desired safety distance a driver allows before he intentionally moves

closer to the car in front. If this value is set to e.g. 10 ft, the following process results in distances between dx_safe and dx_safe + 10ft.

- CC3 (Threshold for entering 'Following') controls the start of the deceleration process or when a driver recognizes a preceding slower vehicle. It defines how many seconds before reaching the safety distance the driver starts to decelerate.
- CC4 and CC5 ('Following' thresholds) control the speed differences during the 'Following' state. Smaller values result in a more sensitive reaction of drivers to accelerations or decelerations of the preceding car, i.e. the vehicles are more tightly coupled. CC4 is used for negative and CC5 for positive speed differences.
- CC6 (Speed dependency of oscillation): Influence of distance on speed oscillation while in following process.
- CC7 (Oscillation acceleration): Actual acceleration during the oscillation process.
- CC8 (Standstill acceleration): Desired acceleration when starting from standstill
- CC9 (Acceleration at 50 mph): Desired acceleration at 50mph

During the calibration process it must be determined which, if any, of these parameters affected the capacity at the toll plaza and which did not. Only the parameters that affected the capacity were adjusted and others were kept at their default values. Driving behaviors were assigned to each link according its link type. As mentioned previously, the link type used for the toll plaza is urban (motorized). The VISSIM window showing where to adjust the Wiedemann 99 modeling driving behavior parameter set is provided in Figure 30.

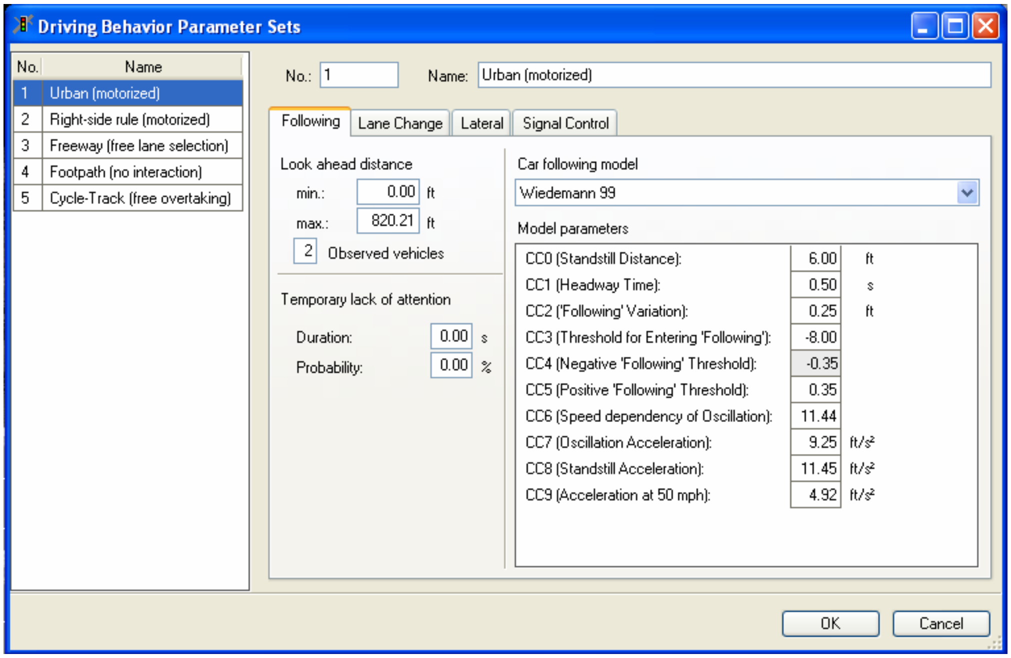

Figure 30: VISSIM - Driving Behavior Parameter Set Window - Wiedemann'99

Once the model was built, traffic data could be collected automatically after each simulation run. In order to set up VISSIM to automatically record data the user must first code evaluation files defining the proper data collection. Throughputs in vehicle per hour, first by lane and second by vehicle type in the specific lane, were setup to be reported by the VISSIM software. For this research vehicles per hour per lane were recorded according to their payment type. Figure 31 is provided to show an example of the data collection file. In this example, eleven total lane throughputs are collected and within each of the eleven lanes five payment types are distinguished. To allow for network saturation (time needed by simulation to allow vehicles to reach the toll plaza and volumes to reach expected arrival rates) the first 500 seconds of each simulation run was completed without data being collected. Simulation runs started on time step 0 and ended at time step 4100, but data was only recorded for the latter 3600 seconds of the simulation.

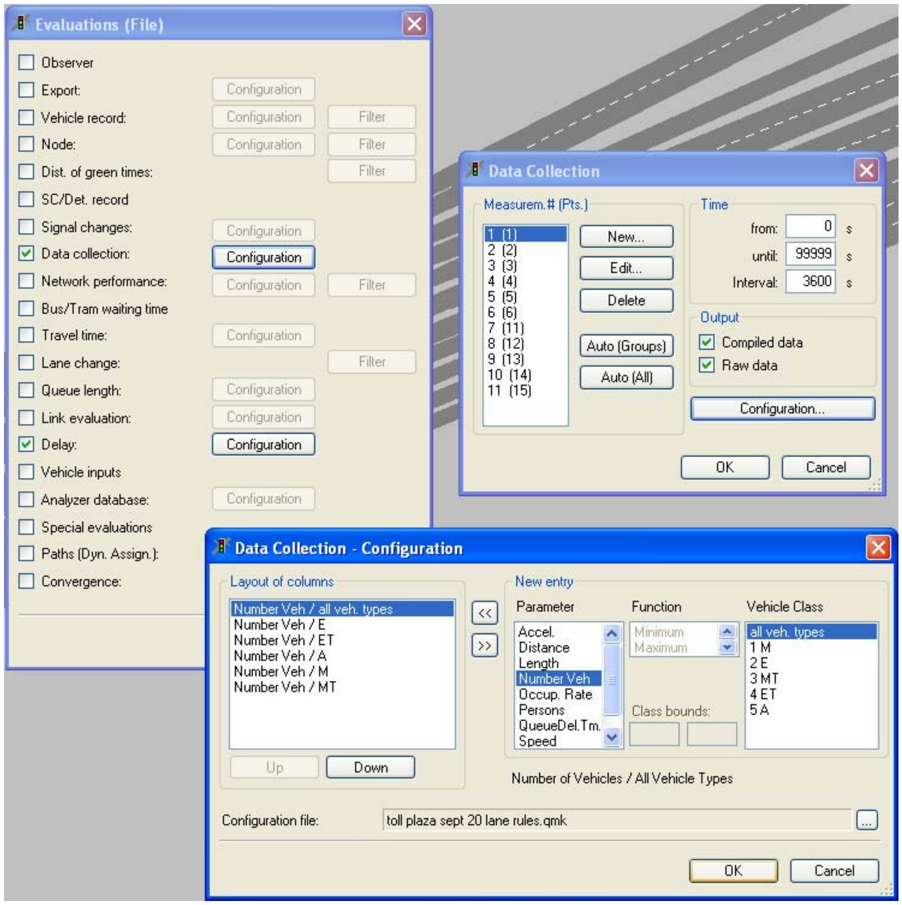

Figure 31: VISSIM Window - Collecting Simulated Data

8.2 VISSIM Calibration Steps

Like SHAKER, the calibration of VISSIM was divided into several steps. First, a particular toll plaza was coded in VISSIM and a measure of effectiveness (capacity) was selected as the index of comparison. Second, an initial evaluation was conducted with VISSIM's default parameter values. Third, if the selected measure of effectiveness is different in simulated and real conditions, an examination of the key parameters was conducted and calibration parameters were determined. Multiple runs with different values of the key parameters were run by trial and error until the calibration is completed. Fourth, for validation requirements, different toll plazas volumes were coded in VISSIM and the field observed MOE is compared to the simulation MOE.

8.2.1 Determining the Required Number of Simulation Runs

Before evaluation of the initial network takes place it must first be determined how many simulation repetitions were required to prove statistical significance. As mentioned previously, microsimulaion models use random seed numbers to perform simulation runs. The random individual vehicle properties are assigned based on the random seed number used for each simulation run. The properties that the random seed number has control over is to generate which vehicle type will enter the simulation next, in which lane they will enter, driver aggressiveness levels, and vehicle interactions once they enter the system. In result of the infinite property variations created by different seed numbers, running multiple simulations can produces as many unique simulation runs as desired. VISSIM simulates vehicle behavior stochastically so every seed number will produce different final simulation results. Due to each runs variance, multiple repetitions of the same model, with different seed numbers, were required to estimate the mean value with a certain level of confidence that the true mean falls within a target interval *(7)*.

The following information is necessary to determine the required number of simulation runs:

- Standard deviation of the sample
- Desired level of confidence
- Desired length of the confidence interval

Unfortunately, the standard deviation of the sample cannot be obtained from the reported data until after the simulation runs have been conducted. That was why a preliminary set of simulation runs was performed to use as an estimate of the standard deviation. This estimate will serve as a preliminary statistic to be used to calculate the number of simulation runs required to make statistical conclusions. Using VISSIM, the parameter chosen to represent the MOE for calibration of the toll plaza model is the throughput and capacity of the toll plaza per hour.

8.2.1.1 Initial Estimation of the Sample Standard Deviation

When estimations of the standard deviation already exist then it could be used as the initial estimate without going through the preliminary simulation procedure. However, since the VISSIM coded toll plaza model was built primarily for this research, preliminary runs had to be completed to obtain a standard deviation estimate of the total plaza throughput. The equation used to calculate an initial sample standard deviation is shown as equation 7.

$$s^2 = \frac{\sum (x - x')^2}{N - 1} \qquad (7)$$

Where:

s = standard deviation

x = variable for which sample variance is desired (i.e. throughput)

x' = average value of the variable produced by the model runs

N = number of model runs

To obtain the information needed to solve for the number of model runs in equation 7 twenty simulation repetitions were preformed. This initial estimate was then revisited later to determine if additional repetitions are required when more precise estimates of mean values are desired.

8.2.1.2 Selection of Desired Confidence Level

The confidence level is the probability that the true mean lies within the target confidence interval *(7)*. The analyst has the liberty of choosing which confidence level he or she thinks will best complement the research. Customarily, 90% or 95% level of confidence is used in micro simulation calibration. The higher the level of confidence usually means that more repetitions are required to satisfy that interval. For this research a 95% level of confidence was preferred.

8.2.1.3 Selection of Desired Confidence Interval

The confidence interval is the range of simulation output values within which the true mean value of the results may lie in order to be accepted. The size of the interval is also at the discretion of the analyst and may vary according to the research. The more precise the interval range is, the more repetitions that will need to be preformed to satisfy the requirements. Confidence intervals are commonly determined from the variance in the data. If the data has a lot of variance then a larger range may be preferred. For this research, the range of the confidence interval corresponds to the same value as two standard deviations above or below the mean, which also represents approximately ad 95% range and a difference less than a 5-percent from the target mean value.

8.2.1.4 Minimum Repetitions Required

Without prior knowledge of how the variation of the simulation data will result, it is impossible to estimate how many repetitions are required to make statistical conclusions without a test. To satisfy this requirement the minimum repetitions were calculated using the initial simulated data in the following equation:

$$CI_{(1-\alpha)\%} = 2 \times t_{(1-\alpha),N-1} \frac{S}{\sqrt{N}} \tag{8}$$

Where:

$CI_{(1-\alpha)\%}$	=	(1-α)% confidence interval for the true mean, where alpha equals the probability of the true mean not lying within the confidence interval
$t_{(1-\alpha/2),N-1}$	=	Student's t-statistic for the probability of a two-sided error summing to alpha with N-1 degrees of freedom, where N equals the number of repetitions.
S	=	Standard deviation of the model results

Applying this process to the data:

An initial set of 20 simulation repetitions using 20 different seed numbers and the toll plaza throughputs from each run were evaluated (see Table 16).

Table 16: Hourly Volume at Different Seed Numbers

Run	Seed #	SR 528 Westbound Lane Throughput (veh/hr)					
		E	E	A	M	M	Total
1	100	555	583	311	307	307	2063
2	115	555	602	311	306	311	2085
3	130	551	605	311	299	307	2073
4	145	559	592	311	306	307	2075
5	160	587	603	311	303	303	2107
6	175	588	597	309	304	308	2106
7	190	590	618	311	302	305	2126
8	205	599	627	311	303	308	2148
9	220	558	582	311	305	303	2059
10	235	595	621	311	300	311	2138
11	250	600	621	311	303	308	2143
12	265	576	630	311	308	303	2128
13	280	591	619	311	301	303	2125
14	295	570	604	311	301	311	2097
15	310	586	628	310	307	307	2138
16	325	584	634	311	299	306	2134
17	340	590	624	311	299	300	2124
18	355	585	618	311	307	303	2124
19	370	602	645	311	305	305	2168
20	385	639	666	311	303	311	2230
Simulation Average							2119.5

These simulation runs were used to determine the mean value of the throughput and its initial standard deviation. Using this information the required number of simulation repetitions was determined. The following process shows the computations involved in the process:

Initial number of runs = 20

Level of confidence = 95%

$\alpha = 1\text{-}0.95 = 0.05$

$t_{(1-\alpha/2),N-1} = t_{(1-0.05/2),20-1} = 2.093$

X' =2119.5

$$s^2 = \frac{\sum (x - x')^2}{N-1} = 1581.9$$

$$S = 39.76$$

To determine if the number of repetitions satisfied the desired confidence interval the target interval was calculated to be 5% of the mean value. Therefore, a confidence interval of 5% of average, which is 106 veh/hr, was the target range to obtain from Equation 8. Solving Equation 8 to determine the required number of simulation runs (N) is an iterative process because it depends on the number of runs in two different locations of the equation. To begin the iterative process, N was set equal to 2 and the corresponding confidence interval was calculated, shown in Table 17. This process continued, increasing N by one each iteration, until the calculated confidence interval was less than or equal to the desired confidence interval.

Table 17: Iterations of Confidence Interval Values Based on Changing N

Number of simulation runs, N	t-statistic, $t_{(1-\alpha/2),N-1}$	Confidence Interval
2	12.706	714.4
3	4.303	197.6
4	3.182	126.5
5	2.776	98.7
6	2.571	83.5

From Table 17 it is apparent that for the simulation to achieve calibration results within the 95% confidence interval chosen for this research that only 5 repetitions was required ($CI_{N=5}<106$). From this analysis it was concluded that the initial approach of 20 repetitions was sufficient and this number can be reduced to any number greater than 5 repetitions for future calibration runs.

Once the number of simulation runs was determined statistically valid conclusions could be made from using the results of the simulation runs. It should be noted that this process for finding minimum number of receptions only serves as an estimate until the final calibration parameters are proposed. At that point Equation 8 will have to be revisited, with final throughput values serving as the data, to serve as a final check to ensure that enough repetitions were in fact provided.

8.2.2 Step 4 Initial Evaluation of network

If an initial evaluation of the network using preprogrammed default values results in a model that can adequately predict the MOE's further calibration is not necessary. In order to test whether the default parameter set provides acceptable and statistically significant results ten replications were run using the VISSIM model. Average lane capacity was recorded as the measure to compare field conditions to VISSIM results. Previously defined acceptable ranges for capacity results are that the simulated capacity must fall within 5% of the field capacity and t-test results must show that the simulation results show statistical significance of similarity to field measured capacity. Ten simulation repetitions, using the default *Driving Behavior* parameters, (see Table 18 below) were run and the capacity results were compared to the field measured capacity. As shown in Table 19, the simulated capacities and the field observed capacities were significantly different (manual lane p-value=4.45E-8, ACM p-value=0.005). The results also show that the percent relative error between capacities is greater than the 5% threshold. Thus, it is determined that further alterations to driver behavior parameters are required for calibration.

Table 18: Default Driving Behavior Parameter Set

cc0	Average standstill distance, ft	4.92
cc1	Desired Headway, s	0.90
cc2	Longitudinal oscillation, ft	13.12
cc3	Start of deceleration process, s	-8.00
cc4	Minimal closing, ft/s	-0.35
cc5	Minimal opening, ft/s	0.35
cc6	Speed dependency of oscillation	11.44
cc7	Oscillation acceleration, ft/s^2	0.82
cc8	Standstill acceleration, ft/s^2	11.44
cc9	Acceleration at 50mph, ft/s^2	4.92

Table 19: VISSIM Initial Evaluation Results

	Manned Lane		ACM Lane	
	Observed	VISSIM	Observed	VISSIM
Measure of Effectiveness				
Average Lane Capacity, veh/hr/ln	355	316	360	344
Statistical Results				
Relative Percent error	11.00%		6.99%	
Significance Value (p value)	4.452E-08		0.005	

8.2.3 Step 5 VISSIM Driver Behavior Parameter Calibration

The overriding assumption for calibrating models is based on simplifying fixed parameters as much as possible. Fixing parameters and/or constraining them to certain intervals helps address the calibration process. Once key simulation driving behavior parameters have been identified and the criterion for optimal calibration has been set, the SR 528 Beachline West Plaza was be calibrated using the adjusted parameters.

8.2.3.1 Calibration of ACM and Manned Lanes

The VISSIM software package provides example networks for demonstration and reference. In the toll plaza demonstration VISSIM recommends classifying roadways with the urban driver behavior parameters. Therefore, the first calibration attempt was to program the network with urban Wiedemann 74 properties and adopt similar driver characteristics. There are only 3 driver behavior parameters within the urban roadway distinction. The default parameters were used to simulate ten runs, each referencing different seed numbers. Initial simulated capacity results indicated that further calibration was necessary. Table 20 shows that the p-value for manned lanes and ACM lanes were well below rejection rejoin ($p>0.05$). Given that default values failed to pass the t-test the VISSIM toll plaza demonstration parameters were adopted for further evaluation. Using recommended toll plaza driver behavior values as a bench mark, multiple simulation runs, each of 10 repetitions, was conducted. The VISSIM recommended urban driver behavior parameters be set so that the average standstill distance is three feet and the additive and multiplicative part of the desired safety distance to one unit each. Table 20 shows the average simulated capacity, relative errors, and t-test statistics for both the manned and ACM lanes for multiple parameter sets. While the table indicates that the average simulated capacity of the ACM lanes was statistically similar to that observed in the field, it also shows that the manned lane average capacity never reaches significant ranges. The default run and test runs were all completed using the same series of seed numbers. Therefore, to verify that a different set of seed numbers also results in unacceptable average capacities Run3 was repeated using ten different seed numbers. Results still indicated that the urban driving behavior parameter set may not be suitable for the simulated toll plaza in this research.

Table 20: Weidman '74 Urban Driver Behavior Parameter Calibration

Base Data Driving Behavior Parameter Sets	Default*	Test Run 1*	2*	3*	Verification**
Average standstill distance, ft	6.6	3.0	3.0	3.0	3.0
Additive part of desired safety distance	2.0	1.0	0.5	0.1	0.1
Multiplicative part of desired safety distance	3.0	1.0	0.5	0.1	0.1
Evaluation of Manned Lanes Capacity					
Average Simulated Capacity, veh/hr	298	314	332	344	338
Average Observed Capacity, veh/hr	355	355	355	355	355
% error	16.06	11.55	6.48	3.10	4.79
T-Test	1E-12	2E-08	3E-05	0.021	0.001
Evaluation ACM Lanes Capacity					
Average Simulated Capacity, veh/hr	315	338	359	359	364
Average Observed Capacity, veh/hr	360	360	360	360	360
% error	12.50	6.11	0.28	0.28	-1.11
T-Test	3E-14	0.014	0.922	0.992	0.705

* Test 1, 2, 3 Seed Numbers: 2000, 2010, 2020, 2030, 2040, 2050, 2060, 2070, 2080, and 2090
** Verification Seed Numbers: 4000, 4010, 4020, 4030, 4040, 4050, 4060, 4070, 4080, and 4090

The other driver behavior parameter set tested for calibration was the freeway link classification. The freeway driver behavior parameters were originally intended by VISSIM to be referenced for free flow freeway operations so it is unknown which parameters should be altered, and how they should changed to model a toll plaza. In order to identify only the driving behavior parameters that have an affect on toll plaza capacity a simplified single lane toll booth was built in the VISSIM program. Thus the only situation being simulated in this model was the queuing condition when the first vehicle in the queue is constrained by a stop time. This simplified model gave insight to which parameters to adjust during calibration because eliminating the effects of upstream lane changing behaviors and lateral vehicle interactions allows the user to identify only those parameters which have a direct effect on an isolated toll plaza queuing conditions. First, the single lane toll plaza was built to the same specifications as the model used for initial calibration. Then each of the ten driver behavior parameters was modified one at a time and simulation repetitions were run for each parameter set. In each case the lane capacity was observed and if it varied from the initial model the modified parameter was identified as important for further calibration. The parameters found to have a significant effect on simulating ACM or manned lane capacities of an isolated toll booth are: average standstill distance, desired headway, longitudinal oscillation, oscillation acceleration, and standstill acceleration identified; also shown in Table 21.

Table 21: Wiedemann 99 Urban Driving Behavior Parameter Significance Table

Driving Behavior Parameter Set		Significant Effect
cc0	Average standstill distance, ft	YES
cc1	Desired Headway, s	YES
cc2	Longitudinal oscillation, ft	YES
cc3	Start of deceleration process, s	No
cc4	Minimal closing, ft/s	No
cc5	Minimal opening, ft/s	No
cc6	Speed dependency of oscillation	No
cc7	Oscillation acceleration, ft/s^2	YES
cc8	Standstill acceleration, ft/s^2	YES
cc9	Acceleration at 50mph, ft/s^2	No

After initial evaluation of key parameters it was discovered that only 5 of the 10 driving behavior parameters affected the toll plaza simulated capacity. Working with the fully developed SR-528 Beachline West toll plaza and nearby approaches the 5 parameters uncovered in the previous step were modified until the simulation produces an average capacity statistically similar to that of field observations. It is proposed that parameters only be adjusted with justification and not just to force the results to match the desired values. The five identified parameters were initially adjusted from the default value to values estimated from field conditions. Similar to the SHAKER calibration, VISSIM calibration test Run2 parameters are adjusted to approximately match the field measured values. Those modifications consist of setting the following: the average standstill distance (cc0) to 6 feet, the desired headway (cc1) to 1 second, the extra distance needed before acceleration in queuing (cc2) was minimal so it is set to 0.5 feet, and just as it was measured, the oscillation acceleration (cc7) is changed to 9.25ft/s^2. The standstill acceleration (cc8) was discovered to have an influence on capacity but the default value of 11.4 ft/s^2 was set to be fixed because larger values tend to be unrealistic. The parameter set tested in VISSIM Run2 was the benchmark for which all further calibration modifications were modified from. As shown in Table 22 the simulated capacities and the field observed capacities of VISSIM Run2 remained significantly different (manual lane *p*-value=0, ACM *p*-value=0.043). The results also show that the mean relative error between capacities is greater than the 5% threshold in the manual lane and just near the threshold in the ACM lane (8.45% and 4.72% respectively). It was proposed that the parameter modification is unfinished until both lanes are simultaneously calibrated. Thus, it was determined that further alterations to driver behavior parameters were required for calibration. To evaluate multiple parameter sets at once, Table 22 is provided to show the parameters referenced and results of multiple runs. Each of the 11 parameter sets utilizes a different configuration of parameters. The results of the error and t-test show that more than one of the 11 tests show statistical similarity to the observed field conditions. The p-values for parameter sets 5, 6, 8, and 10 indicate that there exists insufficient evidence to reject the claim that the average mean of the simulation results are equal to the field measured capacity (p-value>>0.05). Runs 7, 9, and 11 also show p-values greater then 0.05 but sets are neglected because their p-values do not rival in comparison to that of stronger runs.

Table 22: Wiedemann 99 Driving Behavior Parameter Set Calibration

Base Data Driving Behavior Parameter Sets	Test Run Number*										
	1 Default	2	3	4	5	6	7	8	9	10	11
cc0: average standstill distance, ft	4.9	6.0	6.0	6.0	6.0	6.0	5.0	3.0	3.0	3.0	2.0
cc1: headway, s	0.90	1.00	0.25	0.25	0.25	0.25	0.25	0.25	0.50	0.50	0.50
cc2: longitudinal oscillation, ft	13.1	0.5	0.7	0.9	1.0	0.5	0.5	0.5	0.0	0.0	0.5
cc3: start of deceleration process, s	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
cc4: minimal closing, ft/s	-0.35	-0.35	-0.35	-0.35	-0.35	-0.35	-0.35	-0.35	-0.35	-0.35	-0.35
cc5: minimal opening, ft/s	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35
cc6: Speed dependency of oscillation	11.44	11.44	11.44	11.44	11.44	11.44	11.44	11.44	11.44	11.44	11.44
cc7: oscillation acceleration, ft/s^2	0.82	9.25	9.25	9.25	9.25	9.25	9.25	9.25	9.25	9.25	11.44
cc8: standstill acceleration, ft/s^2	11.44	11.44	11.44	11.44	11.44	11.4	11.44	11.44	11.44	11.44	11.44
cc9: acceleration at 50mph, ft/s^2	4.92	4.92	4.92	4.92	4.92	4.92	4.92	4.92	4.92	4.92	4.92
Manned Lanes Capacity Evaluation											
Average Simulated Capacity, veh/hr	316	325	346	346	349	356	351	356	351	342	354
Average Observed Capacity, veh/hr	355	355	355	355	355	355	355	355	355	355	355
% error	10.99	8.45	2.54	2.54	1.69	-0.28	1.13	-0.28	1.13	3.66	0.28
RMSE	2112	1474	517	515	374	358	372	356	443	417	417
T-Test	0.000	0.000	0.045	0.039	0.134	0.773	0.328	0.819	0.295	0.696	0.179
ACM Lanes Capacity Evaluation											
Average Simulated Capacity, veh/hr	334	343	370	370	351	365	375	369	371	360	373
Average Observed Capacity, veh/hr	360	360	360	360	360	360	360	360	360	360	360
% error	7.22	4.72	-2.78	-2.78	2.50	-1.39	-4.17	-2.50	-3.06	0.00	-3.61
RMSE	1120	686	649	657	620	439	790	737	741	573	823
T-Test	0.000	0.043	0.202	0.172	0.253	0.440	0.053	0.218	0.072	0.592	0.078

*All Test Runs Reference Following Seed Number: 2000, 2010, 2020, 2030, 2040, 2050, 2060, 2070, 2080, and 2090

8.2.3.2 Calibration of ETC Lanes

From the data collected in this research it was not possible to observe a true capacity of the ETC lanes. The queuing demand required to calculate capacity was simply not present at any of the four observed sites. Therefore, when modeling the ETC lanes in VISSIM, an estimated value for capacity had to serve as the reference point to evaluate whether and when the model was calibrated. According to Zarrillo *(27)* the capacity of a mixed PC and truck ETC lane is 1,560 vph for a 35 mph posted speed. To prove calibration, multiple VISSIM simulation runs were evaluated until the model could simulate the expected capacity within an accuracy of 5%. An initial test of the VISSIM model indicated that using default driving parameters overestimated the ETC lane capacity by 9.77% (1729 vph). The driving behavior parameter set was then altered until acceptable average capacities were observed. The VISSIM User Manual states that of all driving behavior parameters the average standstill distance and average headway time have the most significant impact on capacity. When using the ETC lane the drivers do not stop so the average standstill distance parameter was kept at default settings. The average headway parameter represents the time a driver wants to keep between his vehicle and preceding vehicle. Thus, for its relevance to ETC tolling and significance to model theory, the average headway was chosen as the calibration parameter. Each simulation test consisted of 10 repetitions, each referencing a different seed number. Table 23 below shows that the model produces acceptable ranges of capacity in run 4 and 5 (run 4=1572vph, Error = 0.72% and run 5=1553vph, Error=0.45%). The simulation was then verified using a different series of seed numbers. It was confirmed in Table 23 that the Run5 parameter set is accepted as a calibrated group of values (Verification Error = 0.71%).

Table 23: ETC Lane Driving Behavior Set Calibration

Base Data Driving Behavior Parameter Sets	Test Run Number* Default	2	3	4	5	Verification of Run 5**
cc0: average standstill distance, ft	4.92	4.92	4.92	4.92	4.92	4.92
cc1: headway, s	0.90	1.20	1.35	1.50	1.70	1.70
cc2: longitudinal oscillation, ft	13.12	13.12	13.12	13.12	13.12	13.12
cc3: start of deceleration process, s	-8.00	-8.00	-8.00	-8.00	-8.00	-8.00
cc4: minimal closing, ft/s	-0.35	-0.35	-0.35	-0.35	-0.35	-0.35
cc5: minimal opening, ft/s	0.35	0.35	0.35	0.35	0.35	0.35
cc6: Speed dependency of oscillation	11.44	11.44	11.44	11.44	11.44	11.44
cc7: oscillation acceleration, ft/s^2	0.82	0.82	0.82	0.82	0.82	0.82
cc8: standstill acceleration, ft/s^2	11.48	11.48	11.48	11.48	11.48	11.48
cc9: acceleration at 50mph, ft/s^2	4.92	4.92	4.92	4.92	4.92	4.92
Evaluation of ETC Lanes Capacity						
Estimated Capacity, veh/hr	1560	1560	1560	1560	1560	1560
Average Simulated Capacity, veh/hr	1729	1680	1639	1572	1553	1549
% error in simulated capacity	9.77	7.14	4.82	0.72	0.45	0.71

**VISSIM Seed #'s: 2000, 2010, 2020, 2030, 2040, 2050, 2060, 2070, 2080, and 2090
**VISSIM Seed #'s: 4000, 4010, 4020, 4030, 4040, 4050, 4060, 4070, 4080, and 4090

The results suggest that there are multiple sets of parameters that appear to have predicted the toll plaza capacity within an acceptable range. These results prove that there is no one particular parameter set that serves as the end all to model calibration. A range of acceptable parameters allows for the user to slightly modify these parameters in future studies where the site characteristics may slightly vary.

8.2.4 Step 6 VISSIM Error Checking

Before the simulation results can be accepted, it is required to determine if enough simulation repetitions were conducted to make any statistical conclusions. To check for simulation repetition requirements, capacity results from Run 5, 6, 8, and 10 were applied to the confidence interval equation shown below as Equation 8. The desired confidence interval was once again defined as less than or equal to 5% of the mean simulated capacity.

$$CI_{(1-\alpha)\%} = 2 \times t_{(1-\alpha),N-1} \frac{S}{\sqrt{N}} \tag{8}$$

Table 24 shows the computations involved in verifying the simulation repetition requirements for each of the proposed parameter sets. A confidence interval of 5% of 350, 360, 362 veh/hr or 17.5, 18, and 18.1veh/hr represent the target interval to obtain from Equation 8. In each case 10 simulation repetition were verified to be sufficient because the target confidence interval is larger then the simulated interval obtained from ten repetitions.

Table 24: Confidence Interval Verification

		Run 5	**Run 6**	**Run 8**	**Run 10**
Mean		350	360	362	350
Standard Deviation		8.38	12.1	9.2	11.5
5% of mean		17.5	18	18.1	17.5
Number of simulation runs, N	t-statistic, *t*(1-α/2),N-1	Confidence Interval per Run #			
2	12.706	150.6	217.4	165.3	206.6
3	4.303	41.6	60.1	45.7	57.1
4	3.182	26.7	38.5	29.3	36.6
5	2.776	20.8	30.0	22.8	28.6
6	2.571	17.6	25.4	19.3	24.1
7	2.447	15.5	22.4	17.0	21.3
8	2.365	14.0	20.2	15.4	19.2
9	2.306	12.9	18.6	14.1	17.7
10	2.262	12.0	17.3	13.2	16.5
Are 10 Runs Sufficient? (if CI < 5%of mean)		Yes	Yes	Yes	Yes

The number of required repetitions relies solely on the evaluation of the variance within the simulated capacity. Therefore, to verify that the results from the proposed calibrated parameter sets of Runs 5, 6, 8, and 10 were reproducible by VISSIM, additional simulation repetitions were conducted using a different series of ten seed numbers. If the statistical results still point to a strong similarity between the field capacity and simulated capacity there was a stronger case that the particular parameter set is calibrated correctly. Table 25 shows the evaluation of the proposed parameter sets. The analysis conducted for verification replicates the evaluation of parameter sets table (Table 23) but as indicated in the last row Table 25 simulation references 10 different seed numbers. The statistical analysis shows that the relative error and t-tests limits are not violated in any of the 4 test runs. Thus, verification suggests that each of the parameter sets 5, 6, 8, and 10 delivers reproducible capacity results.

Table 25: Proposed VISSIM Parameter Set Verification

	Verification Run Number*			
Base Data Driving Behavior Parameter Sets	**5**	**6**	**8**	**10**
cc0: average standstill distance, ft	6.0	6.0	3.0	3.0
cc1: headway, s	0.25	0.25	0.25	0.50
cc2: longitudinal oscillation, ft	1.0	0.5	0.5	0.0
cc3: start of deceleration process, s	-8	-8	-8	-8
cc4: minimal closing, ft/s	-0.35	-0.35	-0.35	-0.35
cc5: minimal opening, ft/s	0.35	0.35	0.35	0.35
cc6: Speed dependency of oscillation	11.44	11.44	11.44	11.44
cc7: oscillation acceleration, ft/s^2	9.25	9.25	9.25	9.25
cc8: standstill acceleration, ft/s^2	11.44	11.4	11.44	11.44
cc9: acceleration at 50mph, ft/s^2	4.92	4.92	4.92	4.92
Evaluation of Manned Lanes Capacity				
Average Simulated Capacity, veh/hr	351	352	360	344
Average Observed Capacity, veh/hr	355	355	355	355
% error	1.29	0.85	-1.41	3.10
RMSE	397	442	369	348
T-Test	0.271	0.379	0.790	0.833
Evaluation ACM Lanes Capacity				
Average Simulated Capacity, veh/hr	359	358	367	358
Average Observed Capacity, veh/hr	360	360	360	360
% error in simulated capacity	0.26	0.56	-1.94	0.56
RMSE	970	997	1137	747
T-Test	0.902	0.783	0.351	0.148

*VISSIM Seed #'s: 4000, 4010, 4020, 4030, 4040, 4050, 4060, 4070, 4080, and 4090

8.3 Step 7 VISSIM Validation

Validation of the VISSIM toll plaza model consisted of two parts. First being, validating the proposed calibration parameters by testing the models ability to simulate a different days worth of capacity data from the same test site. The second part consists of validating the parameter sets by testing their ability to estimate the capacity at other toll plaza configurations. Similar validation methods were used in validation of the SHAKER model so the data sets for each step were already prepared. To recap, the calibration data was obtained from capacity averages observed from the SR-528 Eastbound and Westbound directions on Nov. 13, 2007. The data collected for validation compromised of capacity averages observed on the SR-528 Westbound approach on Feb. 25, 2008. The simulated capacity was then extracted from ten simulation runs from each of the proposed parameter sets and compared to at site capacities. For the parameter sets to remain satisfactory, the resulting simulated capacity must pass the same two tests used in the calibration steps. First, the relative percent error between the observed and simulated capacity must be less then 5 percent. Second, the p-value resulting from the t-test must be above the 0.05 threshold; meaning that a statistical conclusion can be made on the grounds that there is not enough evidence to prove that the data sets are statistically different. Table 26 shows the relative percent error and p-values for the manned lanes and ACM lanes of each of the four proposed calibrated parameter sets. Assessment of the results indicates that only Runs 6 and 8 continued to pass the statistical testing. Run 10 violated the validation investigation because the t-test from manned lanes results are too low (p-value=0.016) and percent error in estimation is greater then 5% (error=6.56%). Run 5 did not directly violate defined thresholds but is not recommended for future use because the manned lane t-test results in a p-value very close to 0.05

(p-value=0.053) and it is suspect to inaccuracy due to marginal error presence. Abandoning Runs 5 and 10 did not hinder the research because acceptable parameter sets still exist in Run 6 and 8. In Run 6 the percent relative errors for the manned and ACM lanes were 3.83% and -1.11% respectively and p-value statistics are 0.22 and 0.645 respectively. In Run 8 the percent relative errors for the manned and ACM lanes are 2.73% and -2.22% respectively and p-value statistics were 0.421 and 0.560 respectively.

Table 26: ACM/Manned Lane Driving Behavior Parameter Set Validation

	Validation Runs *			
Base Data Driving Behavior Parameter Sets	**5**	**6**	**8**	**10**
cc0: average standstill distance, ft	6.0	6.0	3.0	3.0
cc1: headway, s	0.25	0.25	0.25	0.50
cc2: longitudinal oscillation, ft	1.0	0.5	0.5	0.0
cc3: start of deceleration process, s	-8	-8	-8	-8
cc4: minimal closing, ft/s	-0.35	-0.35	-0.35	-0.35
cc5: minimal opening, ft/s	0.35	0.35	0.35	0.35
cc6: Speed dependency of oscillation	11.44	11.44	11.44	11.44
cc7: oscillation acceleration, ft/s^2	9.25	9.25	9.25	9.25
cc8: standstill acceleration, ft/s^2	11.44	11.4	11.44	11.44
cc9: acceleration at 50mph, ft/s^2	4.92	4.92	4.92	4.92
Evaluation of SR-528 Manned Lanes Capacity				
Average Simulated Capacity, veh/hr	349	352	356	342
Average Observed Capacity, veh/hr	366	366	366	366
% error	1.29	3.83	2.73	6.56
T-Test	0.053	0.222	0.421	0.016
Evaluation SR-528 ACM Lanes Capacity				
Average Simulated Capacity, veh/hr	351	365	369	360
Average Observed Capacity, veh/hr	361	361	361	361
% error	0.26	-1.11	-2.22	0.28
T-Test	0.368	0.645	0.560	0.926

* Seed Numbers: 2000, 2010, 2020, 2030, 2040, 2050, 2060, 2070, 2080, and 2090

The next step to model validation was to validate the remaining parameter sets by testing if they still produce acceptable ranges of capacity when used in different toll locations. As done in the SHAKER validation, toll plazas along SR-417, SR-429, and Florida Turnpike Mainline were modeled using the remaining parameter sets (Run6 and Run8). To model each of the different toll plazas three major settings had to be adjusted. First, the roadway configuration was updated to resemble that of the lane configuration of each toll plaza. Second, the vehicle distribution was altered to match that of each site. And third, the dwell time distributions will be altered to match the unique processing times observed due to the different toll amounts. The main setting that remained the same for all sites was the driving behavior parameter sets. At each site both Run6 and Run8 parameter sets were used to determine if both still produce acceptable capacities. It is fair to assume that toll plaza capacity is highly dependant on processing time, and processing time is dependant on plaza location, toll amount, and the toll collector. While, the toll collector speed is out of the control of the research, the other two factors can be accounted for. For each toll plaza model the dwell time distribution were adjusted so that that particular toll plaza processing times is reflected in the model. Therefore, the dwell time distribution for each of the three validation groups was updated in VISSIM. Charts showing the unique distribution of processing rates are provided in Figure 32 through Figure 34. From these charts it is quite simple to identify which plaza should have a higher capacity. The cumulative percentage rate shows which processing time the bulk of the observed values occur. Models were evaluated to ensure that processing time delays matched that of the distribution curves.

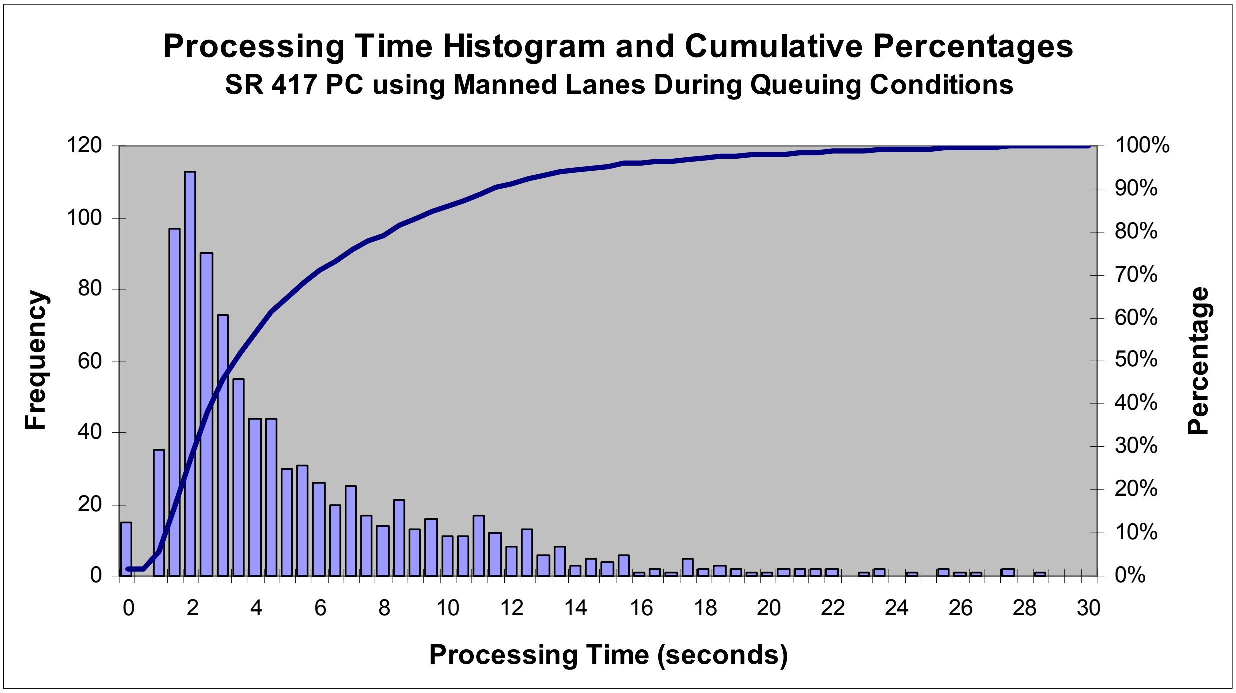

Figure 32: Processing Time Histogram for SR 417 Manned Toll Lanes

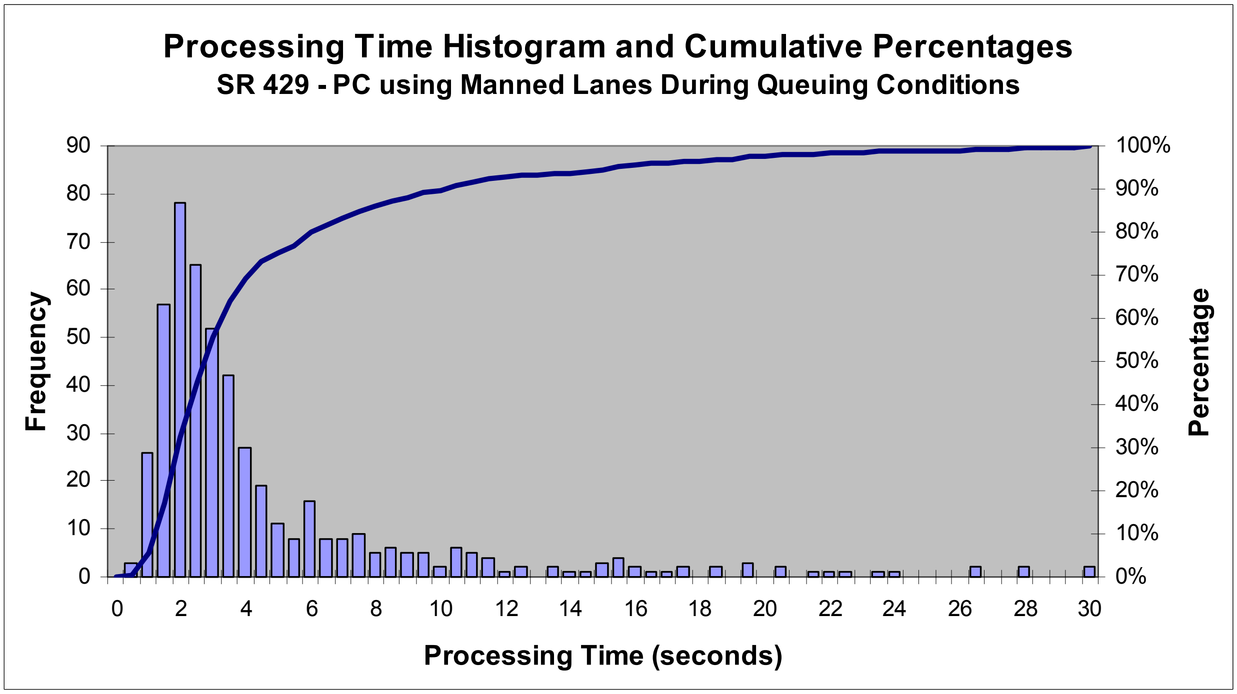

Figure 33: Processing Time Histogram for SR 429 Manned Toll Lanes

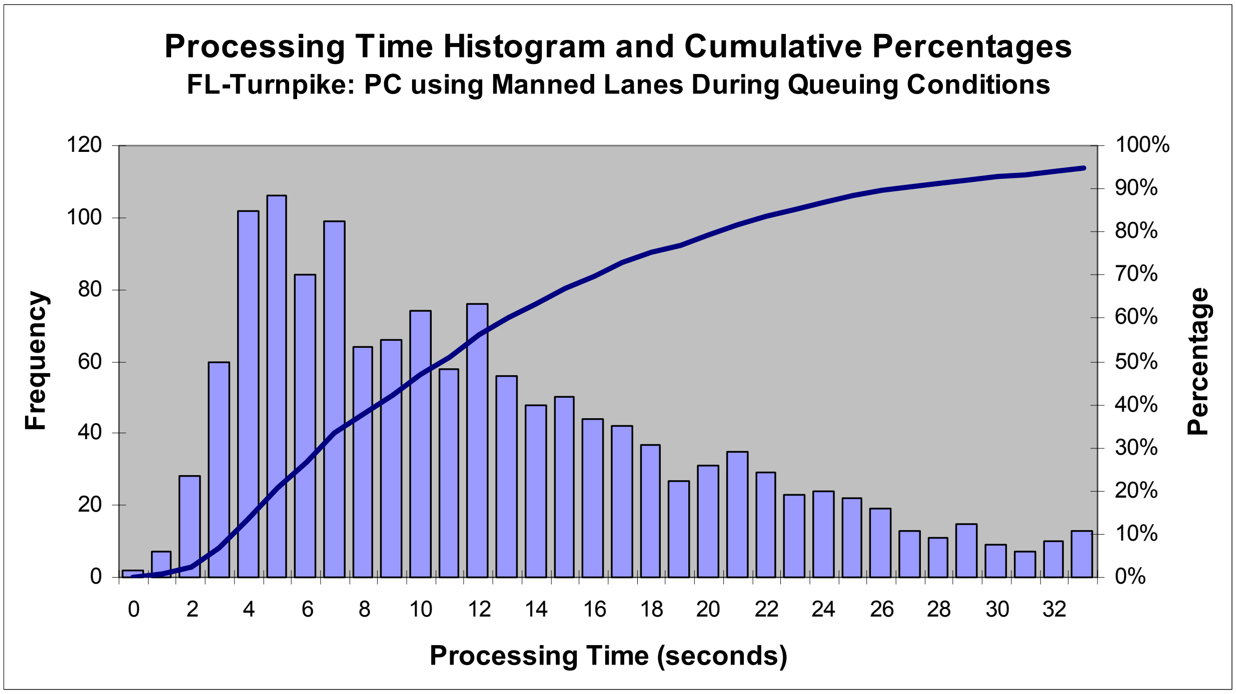

Figure 34: Processing Time Histogram for FL Turnpike Manned Toll Lanes

Once again, the simulated toll plaza capacities were compared to observed capacities by means of relative percent error in capacity and t-tests. As shown in Table 27 the results show that in each of the 3 validation sites the errors in simulation and t-tests indicate that there is similarity between the simulated capacities and the observed capacities. Therefore, it can be concluded with statistical evidence that the parameters used in Runs 6 and 8 provided the VISSIM model with sufficient calibration and validation.

Table 27: ACM/Manned Lane Parameter Set Validation Using Differing Toll Plaza Designs

Base Data Driving Behavior Parameter Sets	**Run 6**	**Run 8**
cc0: average standstill distance, ft	6	3
cc1: headway, s	0.25	0.25
cc2: longitudinal oscillation, ft	0.5	0.5
cc3: start of deceleration process, s	-8	-8
cc4: minimal closing, ft/s	-0.35	-0.35
cc5: minimal opening, ft/s	0.35	0.35
cc6: Speed dependency of oscillation	11.44	11.44
cc7: oscillation acceleration, ft/s^2	9.25	9.25
cc8: standstill acceleration, ft/s^2	11.4	11.44
cc9: acceleration at 50mph, ft/s^2	4.92	4.92

Validation Sites	**Manned Lane Statistics**	
SR-417 Lake Jesup		
Average Simulated Capacity, veh/hr	358	365
Average Observed Capacity, veh/hr	361	361
% error	**0.83**	**-1.11**
T-Test	**0.720**	**0.723**
SR-429 Southern Beltway		
Average Simulated Capacity, veh/hr	395	397
Average Observed Capacity, veh/hr	402	402
% error	**1.74**	**1.24**
T-Test	**0.444**	**0.719**
FL-Turnpike Mainline		
Average Simulated Capacity, veh/hr	206	209
Average Observed Capacity, veh/hr	205	205
% error	**-0.49**	**-1.95**
T-Test	**0.934**	**0.272**

* Test 6 & 8 Seed Numbers: 4000, 4010, 4020, 4030, 4040, 4050, 4060, 4070, 4080, and 4090

9. APPLICATIONS AND COMPARISONS OF TOLL PLAZA MODELS

This section is dedicated to the uses and developments of each of the two evaluated simulation models. First, a case study focusing on a new application within the SHAKER model is observed. Next, SHAKER and VISSIM models are coded to represent a lane closure so as to evaluate their strengths in simulating a special case scenario. Lastly, to establish model efficiency, the two model's results were compared to the initial field observed capacities and queue lengths.

9.1 SHAKER Best Configuration Optimization Case Study

A new application was added to SHAKER that has the ability to automatically select the best configuration of a toll plaza given that the lane-user remains unchanged. After inputting data extracted from observed volumes into SHAKER and running the model, this application, using same input data, generates the optimum lane configuration. The optimum lane configuration was based on increasing capacity, reducing queue lengths, but still providing the lanes required to service all payment and vehicle types. Using data from the SR-528 Eastbound approach, an example of the best configuration outputs is generated and shown in Figure 35. The top row displays the best lane configuration; the bottom row represents the current lane configuration, and below both is where the input data is defined. In the display throughputs and capacities are represented by vertical bars, similar to that of a bar chart. The green color bar denotes the capacity of each lane and the red, blue, and yellow denote the throughputs for the ETC, automatic, and manual respectively. Within this application, SHAKER is not limited to providing the results from the best configuration, but also automatically calculates the capacity, throughput, and queue lengths for every possible lane configuration. The pull down menu under

POSSIBLE CONFIGURATIONS allows the user to select from any of the lane configurations and visually examine the results of each.

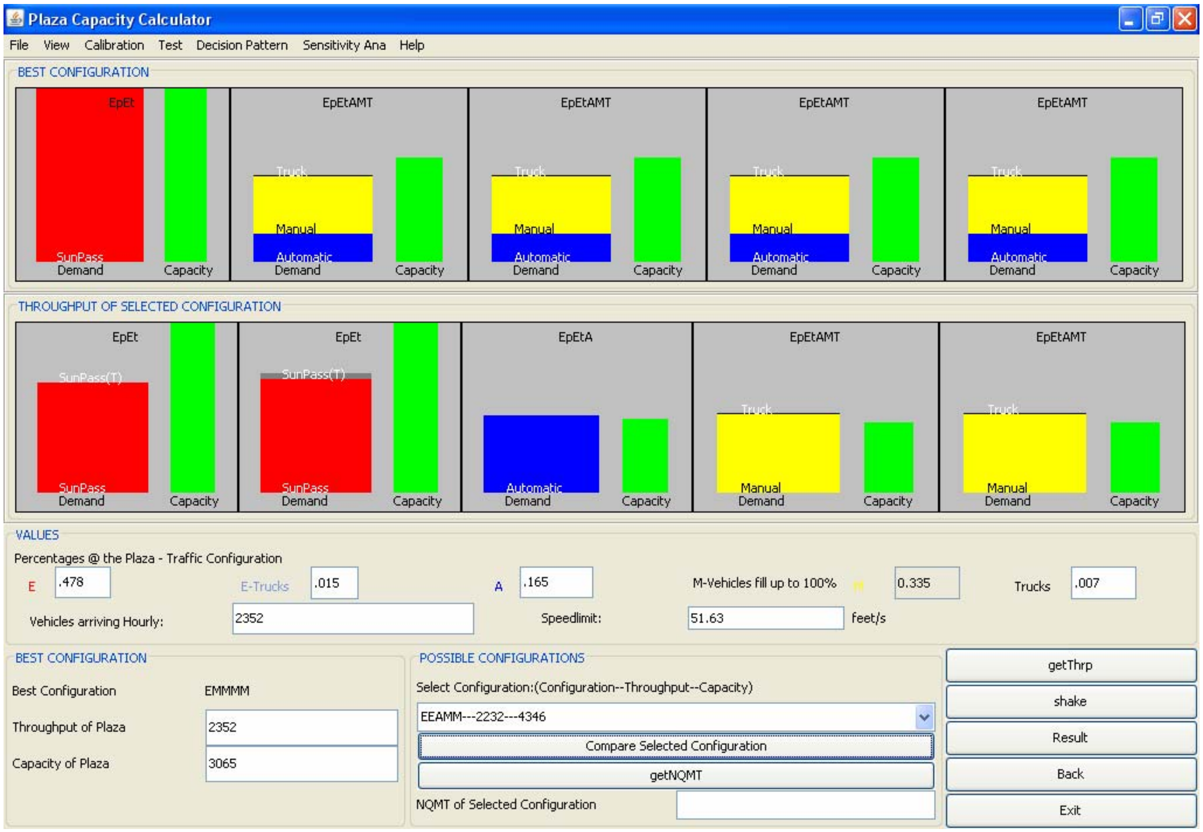

Figure 35: SHAKER GUI for current versus best configuration of SR-528 EB

All four toll plazas (6 different configurations) were coded in SHAKER and the new SHAKER application was run to determine the optimal toll plaza configuration that increases capacity and reduces queues in the manual and automatic lanes. The data used for demand values and vehicle percentages are randomly selected from the capacity periods that were previously identified in this research. The traffic data inputted for this case study are provided in Table 28. Case study results are shown below, as Table 29 summarizes the current and the optimal configurations of the six configurations. For instance, for SR 528 WB, the existing configuration consists of 2

ETC lanes, 1 ACM lane, and 2 Man. lanes and the optimal selected SHAKER configuration is 2 ETC lanes, 0 ACM lanes, and 3 Man. lanes. The best configuration of this toll plaza would result is a queue reduction of 119 veh/hr. As shown in Table 30, in order to increase capacity and reduce queuing SHAKER recommended changing the configuration of all toll plazas except for the FL Tunpike toll plaza.

Table 28: Input Data for Best Configuration Case Study

Site Location	Passenger Car Demand veh/hr/per lane type			Truck Demand veh/hr/per lane type			Demand Total (veh/hr)	Queue Length veh/per lane			Vehicle Type Percentages				
	M	A	E_p	M_T	A	E_T		M	A	E	E_p	E_T	M	A	M_T
528 WB	788	388	1124	16	0	36	2352	88	32	0	0.478	0.015	0.335	0.165	0.007
528 EB	636	728	2000	16	0	52	3432	20	20	0	0.583	0.015	0.185	0.212	0.005
528 EB 2	695	776	2000	5	0	52	3528	20	20	0	0.567	0.015	0.197	0.220	0.001
417 NB	708	-	1412	12	-	80	2212	56	-	0	0.638	0.036	0.320	0.000	0.005
417 SB	752	-	2128	0	-	76	2956	36	-	0	0.720	0.026	0.254	0	0
FL TPK	816	-	500	0	-	115	1431	6	-	0	0.349	0.080	0.570	0	0
429 NB	420	-	428	0	-	20	868	20	-	0	0.493	0.023	0.484	0	0

Table 29: SHAKER selected optimal configurations

	Current Configuration			**Best Configuration**			***Total Queue Reduction (veh/hr)***
	Lane Type (count)			***Lane Type (count)***			
Toll Plaza Site	***ETC***	***ACM***	***Man.***	***ETC***	***ACM***	***Man.***	
SR 528 WB	2	1	2	2	0	3	119
SR 528 EB	2	2	2	1	3	2	20
SR 417 NB	2	0	2	1	0	3	14
SR 417 SB	2	0	2	1	0	3	31
FL TPK	1	0	4	1	0	4	0
SR 429 NB	2	0	1	1	0	2	6

9.2 Simulated Lane Closure– Comparison of Special Scenario Case Study

An important element of traffic simulation models is rooted in their ability to adhere to special situations. Due to special events such as: heavy demands, emergency situations, crashes, lane closures, and unexpected maintenance, traffic often deviates from the expected. A simulation model should have the ability to do so also. To test the two model's ability to adjust to special situations a lane closure will be coded into the base network and results are collected. A lane closure was chosen as the most advantageous scenario to test because it applies to traffic situations on multiple levels. Obviously, a lane closure represents the situation of actually closing a lane due to maintenance or accident at the toll plaza. In addition, if the same input volume is used a lane closure also estimates the effects of an increased traffic demand per lane. Instead of running another simulation to test increased demand, the lane closure forces the same situation upon the model. This procedure will also provide results on whether or not the SHAKER model produces similar results to that of the widely used VISSIM software. The procedure followed for this investigation starts with first using SHAKER to find the best configuration of the toll plaza with a lane closed and recording results on capacity, throughput, and queue lengths. Next, the network built in VISSIM is adjusted to replicate the SHAKER recommended configuration (See Figure 37 and Figure 38). The throughput and queue length results from multiple VISSIM runs will then be compared to the SHAKER outputs.

Using the same demand volumes as in the best configuration case study (Table 28) SHAKER was run with a lane closed. Figure 36 is provided to show an example of the best configuration generated by SHAKER using SR-528 Eastbound Data. After each of the 6 configurations are run the best one lane closed configurations results are tabulated in Table 30.

As shown in the table below, SHAKER suggested that sacrificing an ETC (keeping all other types available) in most of the cases results less queues at the toll plaza compared to closing any other type of lanes. However, for the FL TPK toll plaza SHAKER suggested closing a manual lane since this toll plaza consists of one ETC lane.

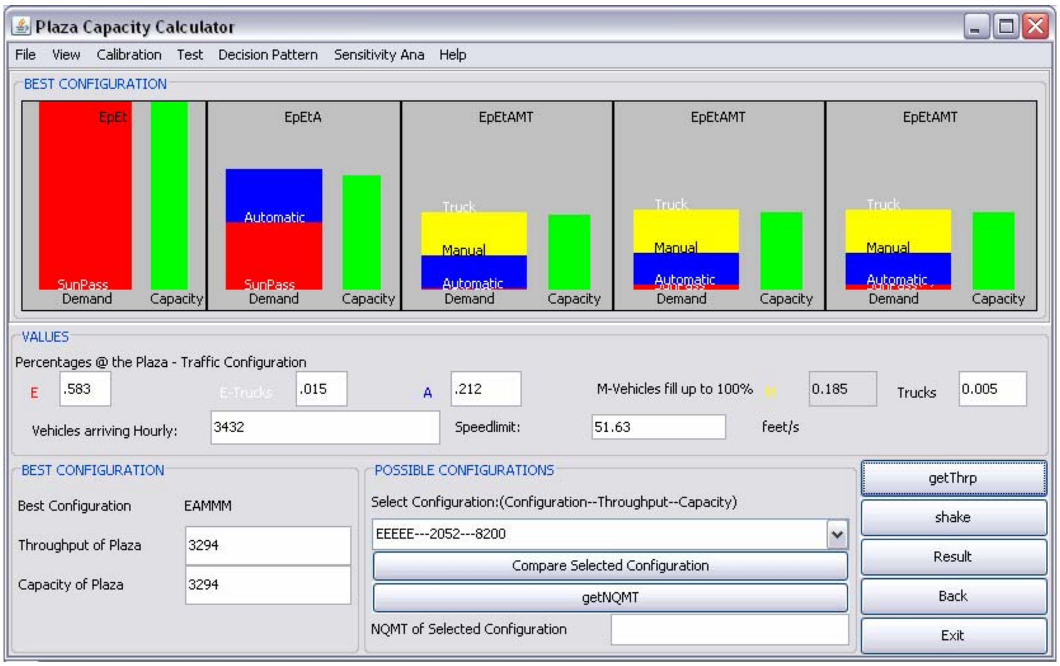

Figure 36: Best One Lane Closed Configuration Example - SR 528 Eastbound

Table 30: Best configuration in case of lane closure

	Current Configuration			One Lane Closed Configuration		
	Lane Type (count)			Lane Type (count)		
Toll Plaza Site	ETC	ACM	Man.	ETC	ACM	Man.
SR 528 WB	2	1	2	1	1	2
SR 528 EB	2	2	2	1	1	3
SR 417 NB	2	0	2	1	0	2
SR 417 SB	2	0	2	2	0	1
FL TPK	1	0	4	1	0	3
SR 429 NB	2	0	1	1	0	1

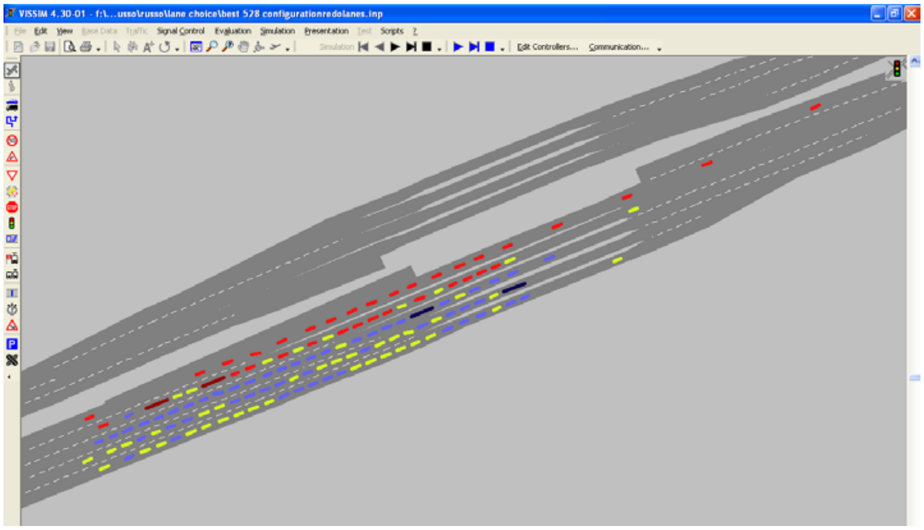

Figure 37: Two Dimensional Model of VISSIM Network with Lane Closure

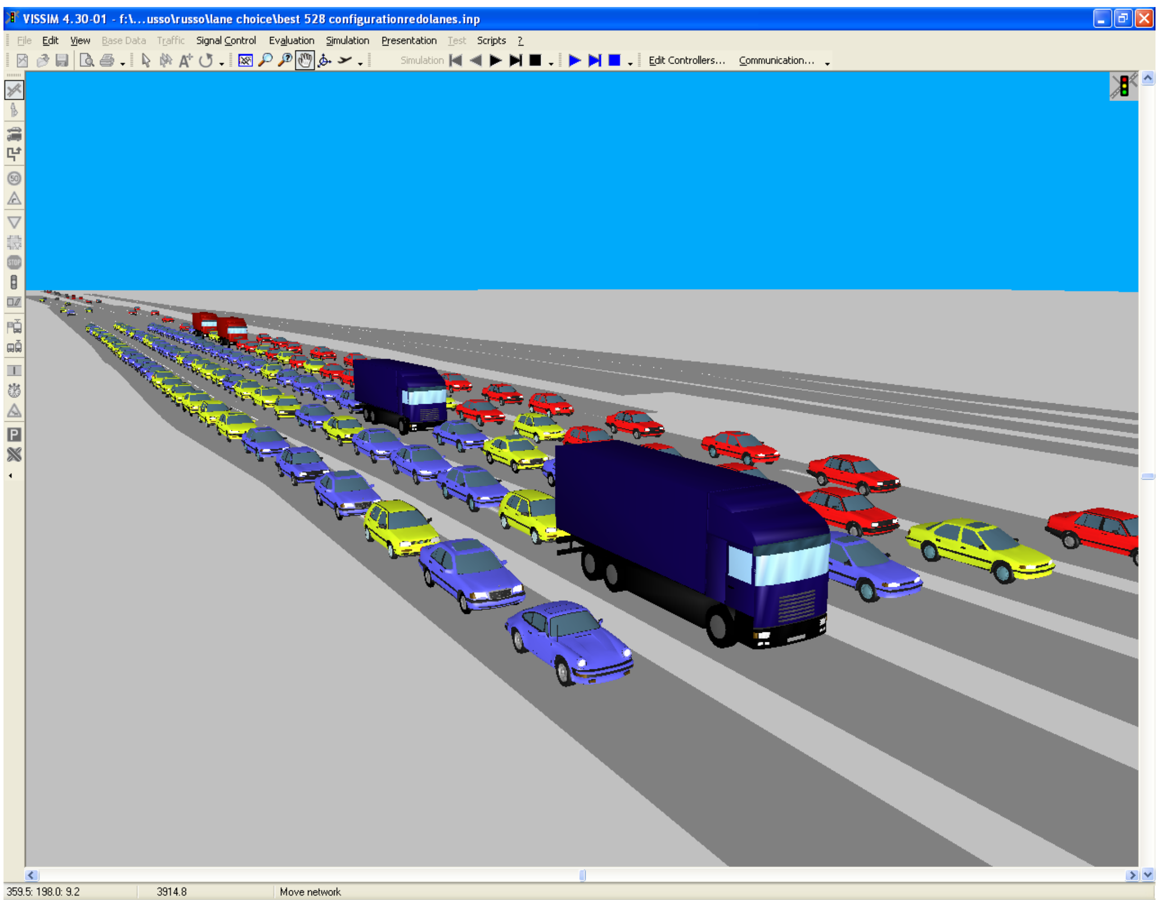

Figure 38: Three Dimensional Model of VISSIM Network with Lane Closure

Next the best one lane close configuration was coded into the VISSIM networks. This was easily done by simply deleting one of the connector nodes from one of the networks previously built for model calibration. One additional change to the network that must be done before realistic results are noticed was that ETC vehicles should be allowed to use any lane, which they will if the queues are short enough, and there needs to be more variation for ACM users to be willing to use manned lanes. This was accomplished by redrawing the route choices so that every vehicle was available to use any lane type; except that cash users can not use ETC lanes. Each approach from the SR-528 plaza and SR-417 plaza was used for this evaluation in VISSIM. Other sites were neglected for this procedure if there was no decision on which lane type to close. For instance, there is only one ETC lane and four manned lanes at the Florida Turnpike Mainline Plaza so in order to maintain the correct services to drivers it is unavoidable that only a manned lane be closed. The opposite holds true for the SR-429 plaza where two ETC lanes and one manned lane are currently in operation. It is also unavoidable that if a lane is to be closed it must be the ETC lane.

The results of the VISSIM generated throughput and queue lengths are provided in Table 31. SHAKER results were consistent and independent of when and how the results are obtained. On the other hand, VISSIM results were dependant on the seed number that corresponds to that simulation run. Therefore, VISSIM results originate from 10 different seed numbers. The error calculated is the fractional difference between the VISSIM model (base case) and the SHAKER Model (comparison case) divided by the base case results. The results indicate that SHAKER model estimates an accurate throughput of the toll plaza within 5% for each of the VISSIM simulated toll plazas. Also, according to the GEH statistic, the SHAKER model and VISSIM model agree on similar queue length estimation. The GEH statistic is used here because percent error would be skewed by the small values associated with the queue lengths and results would not be indicative of the actual strength of model estimation *(19)*. When evaluating the GEH statistic any result less than 5 indicates a suitable match between modeled and observed conditions; a result between 5 and 10 warrants further investigation of the data; and any value greater than 10 is not a suitable match and indicates there is no match between the two data groups.

Table 31: One Lane Closed Configuration Evaluation

Comparison of Simulated Throughput (veh/hr) / **Comparison of Simulated Queue Length (Veh)**

SR-528 Eastbound	E	A	MA	MA	MA	Total	Error*	E	A	MA	MA	MA	Total	GEH***
VISSIM**	1493	778	378	373	374	3396		8	18	26	26	27	105	
SHAKER	1636	551	361	372	373	3293	0.03	24	29	25	29	31	138	2.994

SR-528 Westbound	E	A	M	M	Total	Error*	E	A	M	M	Total	GEH***
VISSIM**	1092	352	351	344	2139		0	28	41	41	110	
SHAKER	1159	369	352	352	2232	-0.043	0	19	50	50	119	0.841

SR-417 Northbound	E	M	M	Total	Error*	E	M	M	Total	GEH***
VISSIM**	1494	352	252	2098		2	13	14	29	
SHAKER	1491	354	354	2199	-0.048	0	7	7	14	3.235

SR-417 Southbound	E	E	M	Total	Error*	E	E	M	Total	GEH***
VISSIM**	1073	1274	354	2701		0	0	387	387	
SHAKER	1102	1103	360	2565	0.05	0	0	390	390	0.152

* Error calculated as percent error in simulation variation using VISSIM as base case and SHAKER as test case

** VISSIM results are averages of 10 simulation runs, each run referencing a different seed number

*** GEH Statistic = square_root[(2(m-c)^2)/(m+c)], where m is modeled value and c is orignal value
a value less than 5 is considered a suitable match between observed and modeled values (19)

9.3 Simulated Values – Comparison of Estimated Values to Observed Values

The final step in the use of the SHAKER model and VISSIM simulation was to evaluate the effectiveness of each to determine which better estimates volumes and capacity closest to that of the observed conditions. Each of the two models has been calibrated and statistically verified to match the observed capacity but no evaluation has been conducted to determine the precision of each. The queue lengths cannot be fully evaluated from model to model in this step because the queue is the average of the difference between the demand and throughput and in the field the queue length is dependant on unpredictable conditions and varies from day to day. Table 32 is provided to show the comparison of the observed and simulated capacities from SHAKER and VISSIM. It is extracted from the table that there is no clear indication of which model better estimates capacity. In each of the sites tested the observed capacity fell between the SHAKER and VISSIM simulated capacity. This made it difficult to determine which model better estimates the capacity. However, when evaluating calibration results from Table 32 it was determined that based on t-test values there is no significant difference between the SHAKER and VISSIM model when compared to the observed conditions. The differences between the simulated capacities do show a slight trend that for each manned lane SHAKER slightly over estimates capacity and VISSIM slightly underestimates capacity, but this trend is considered insignificant to make any sound conclusions from. In either situation or lane type the slight variation is small enough to not completely hinder the performance of the toll plaza. A trend is however observed in the ETC lane capacity as SHAKER estimates a larger capacity than does the VISSIM simulation. The capacity of the ETC lanes were determined from the simulation models but not though observation. It was not possible to observe ETC lane capacity in the field because there was never a period of time where queuing was present for any substantial amount of time. From this analysis there is no clear indication which simulation model better estimates capacity but there is evidence that for manned lanes the SHAKER model is the less conservative approach because the capacity is larger thus meaning the resulting queues and delays will be shorter.

Table 32: Comparing Toll Plaza Simulation Models

Toll Plaza	**Data Category**		
Manned Lanes	**Observed**	**SHAKER**	**VISSIM***
SR 528			
Average Capacity, veh/hr	355	358	352
% error in capacity		-0.85	0.85
T-Test on capacity		0.467	0.222
SR 417			
Average Capacity, veh/hr	361	367	358
% error in capacity		-1.66	0.83
T-Test on capacity		0.366	0.720
SR 429			
Average Capacity, veh/hr	400	414	395
% error in capacity		-3.50	1.25
T-Test on capacity		0.611	0.444
SR 91 FL TPK			
Average Capacity, veh/hr	202	208	206
% error in capacity		-2.97	-1.98
T-Test on capacity		0.131	0.934
ACM Lanes SR-528			
Average Capacity, veh/hr	360	352	365
% error in capacity		0.85	-1.11
T-Test on capacity		0.222	0.440
ETC Lanes SR-528			
Estimated Capacity, veh/hr	N/A**	1587	1559

* VISSIM Test Run 6 is used for this analysis
**Not available. Queues were never present in field observations but through simulation capacities can be obtained

10. SUMMARY AND CONCLUSIONS

This research focused on the development and calibration of two vastly different simulation models. The two models utilized were, SHAKER, a deterministic queuing model for vehicles utilizing toll collection facilities, and VISSIM, a globally popular stochastic simulation software. The benefits of simulation models led to the purpose of this thesis, which was to examine the effectiveness of two toll modeling programs that are similar in purpose but vary in approach and methodology. Both SHAKER and VISSIM toll plaza models have the potential to work as a tool that can estimate the maximum throughput and capacity of toll plazas so that planners and engineers can better develop traffic plans for toll plaza design.

An extensive field study provided valuable processing time and demand data that was used to establish capacities based on lane type, payment type, payment amount, and vehicle type. The capacities resolved were used as the primary measure in calibrating and validating simulation modes. Much attention was put on using the field measured values in as many instances in calibration as possible. It is quite possible that both models could be forced to replicate field results without using field conditions but the methodology would contribute nothing to the field for future research. Upon completion of calibration of the two simulation models it was determined that each of the two software were successful in modeling toll plaza capacity and queuing. The equilibrium based assumptions of the SHAKER model and stochastic route choice decisions utilized in VISSIM both were effective simulation foundations. After each model was validated the uses of the model were investigated. The best configuration application in SHAKER was demonstrated and proven to be effective in suggesting a better configuration of current plaza configurations. The best configuration application was also applied to test the effects that closing a lane or increasing demand has on a toll plaza. As expected, each simulation model possessed benefits over the other in terms of set up time, analysis reporting time, and practicality of results, and potential for adapting to variation.

10.1 Evaluation and Limitations of the Simulation Models

The first element to be compared is the set up and simulation time. The SHAKER model setup takes mere seconds in order to create a network and input vehicle, another few seconds to calibrate driving parameters, and roughly 10 or so additional seconds for simulation to run its course and to report analysis. Conversely, setting up the VISSIM model, even for the most experienced user, can take several hours because the roadway and traffic must be created and defined before any simulation can take place. Also, when using VISSIM the report analysis time can take several more hours as its accuracy is dependant on the number of required simulation runs and complexity of the network. Because of VISSIM's stochastic nature of assigning unique characteristics to each vehicle one by one the simulation time required for VISSIM is exponentially higher than the time needed for SHAKER. The major benefit to having quick programming and reporting times is that in times of unpredicted conditions SHAKER can be setup and run in a matter of minutes; thus allowing decision makers to take quick action when needed. VISSIM on the other hand would take hours to run particular scenarios.

In addition to long set up times, data reporting in VISSIM is also more complicated than in SHAKER. In order to observe capacity results in VISSIM a file type requesting the software to record particular data for each lane and vehicle time must first be created. Then the user must exit VISSIM and open up a text file where results are recorded. In constant, because SHAKER is dedicated for the uses of toll plaza queuing, the capacity and queue reporting is simplistic, available immediately, and no programming of file types is necessary to obtain results. Unlike VISSIM, SHAKER reports capacity and queues in a visually pleasing graphical manner assisting the users understanding of the traffic situation without having to look at specific numbers. This element is beneficial for multiple reasons. For instance, when reporting results to unfamiliar persons the graphical file is quickly understood even with no previous knowledge. Also, its simplicity allows for a user of any SHAKER experience or transportation expertise to quickly grasp how to use the software and what the results mean.

As mentioned multiple times in this research traffic is sometimes unpredictable. VISSIM is most beneficial when referring to unpredictability because it's modeling allows for variability between drivers while conversely SHAKER assumes equilibrium amongst lane choice and queuing. This variability creates a more realistic simulation condition which also more compares to observed traffic patterns. SHAKER does not show the simulation of individual vehicles but it is known that it assumes equal headways and arrival rates throughout the analysis time period. VISSIM on the other hand, inputs vehicles according to a predefined assignment associated to each seed number. Another benefit to the VISSIM model's variability is that the user has the option to model the processing rate stop time as an empirical distribution rather. This differs from SHAKER because in SHAKER the user must enter one processing time that is applied to all vehicles within that lane group and payment type. As shown in the calibration steps in this research the processing time distribution does not follow a normal curve nor is there one distinct processing rate that prevails over the others. Allowing for a distribution of processing rates determines the presence and magnitude of queues. When visually observing the VISSIM simulation it was not rare to see a wave pattern in the queue length as queues grown and shrink according to the variability in arrival rate and processing time; similar wave queuing conditions were observed in the field.

Even though differences are prevalent, it is important that in each simulation model the capacity is accurately simulated and each can be used to benefit operational situations related to toll plaza traffic conditions. While the benefits of each model are important, it is also just as important to identify any limitations of the models. The SHAKER model strictly has the potential to model an isolated toll plaza. If queue lengths are longer than the queue storage bays upstream of the toll plaza then queues may back up into un-congested ETC lanes; thus reducing the capacity of those lanes. Without any network visualization it is difficult to identify if this situation is occurring. A limitation also lies within the distribution methodology in the SHAKER model. When SHAKER faces demands in excess of the lane plaza capacity the program assumes that the queue vehicles follow the same distribution as the throughput distributions entered. However, this limitation can easily be avoided if the overall demand accounts for the make up of vehicles types within the queue. Another limitation of SHAKER that has already been addressed is its deterministic approach to modeling. SHAKER modeling assumes equal headways and processing times. However, field observations show that queuing is more localized and dependant on the arrival rate and heavily dependant on processing times. For instance, one particular vehicle could require over one minute to complete their transaction and during that time multiple vehicles will undoubtedly queue up temporarily until favorable conditions return. This is the very reason that multiple time periods to estimate capacity and queue lengths are needed to more accurately find an average value of these constraints in the field and when using VISSIM.

It should be noted that all four toll plazas evaluated for this research are managed and operated by the Florida's Turnpike Enterprise. At toll plazas managed by different entities could result in different procedures that could influence lane capacity. The capacities and processing times are limited to toll plazas that utilize gates to indicate when proper payments are received. If toll plazas are installed without gates the processing time is expected to be reduced and capacity is expected to increase. If future research is to be conducted in this topic it is recommended that a toll plaza that experiences queues in the ETC lanes as well as cash payments lanes so as to develop a capacity value for all lanes at once.

10.2 Future Research Topics

It has been concluded in this research that the isolated toll plaza models developed and calibrated effectively predict the operations at toll plazas of numerous lane configurations and traffic demand. However, the applications can be further improved from future research on the topic. This section is intended to introduce some ideas that will expand upon the research conducted in this thesis.

To calibrate the two models capacity was chosen as the measure of effectiveness to determine the level of performance of each model. This calibration procedure could be further strengthened by validating the results of each model's ability to estimate accurate delays. Delay is commonly used to report the Level of Service of a facility and is often the most important measure to report to the public. Unlike capacity, delay is dependant on the demand volumes so one cannot simply calibrate based on this parameter but there is potential in this research to strengthen the validity of each model.

Another recommendation for future research is rooted in examining altering model assumptions. In the isolated toll plaza model the SHAKER model assumes perfect equilibrium for lane choice behavior and the VISSIM model attempts to reach equilibrium lane choice conditions. However, the field data collected suggests that undefined factors may influence the lane choice of particular vehicles. Based on observation of the field data it appears that heavy vehicles seemed to prefer the right most manned toll booth lane over other manned lanes. Also, the processing time of this far right lane also resulted in slightly longer processing times and smaller capacity values.

To more properly simulate these anomalies it is suggested that research be conducted on using the VISSIM origin-destination dynamic routing decisions rather then route choice decisions used in this research.

Using the origin-destination dynamic routing decisions leads to yet another topic that should be researched further. The VISSIM model developed here does indeed predict the throughput, capacity, and queue lengths of an isolated toll plaza but does not account for an entire network of traffic. To better understand the integration and relationship effects that toll plazas have on adjacent free flowing roadway sections, nearby on and off rams, and on downstream toll booth conditions, an all encompassing simulated network of the entire Florida Turnpike Network using VISSIM is recommended. Within this network the user can define origins and destinations at the very instant a vehicle enters the simulation model's constraints. Using this modeling technique simulated vehicles will choose lane types based on interests other then just queue lengths. For instance, in actuality vehicles may choose the far right lane, even when longer queues exist, so as to prepare for exiting the mainline just downstream of the toll plaza. This recommended research also has potential in determining if an isolated toll plaza model remains effective when an entire simulated network is added to between separate isolated toll plazas. Thus, the methodology used to develop the processing time simulation and queuing operations in this research can be combined with intermediate sections and ramps to research the benefits of VISSIM's potential to simulate the operation of an entire network.

This same approach can be used with the SHAKER model as well. Research is recommended on how to instantaneously run multiple SHAKER models, each representing a different toll plaza configuration and traffic demand along a network. Instead of having to simulate each toll plaza separately this model would be able to automatically adjust for varying inputs to the network from on ramps and reduction in downstream demand due to network departures. This model would serve promising when modeling the effects that special event conditions such as sporting event, accident, or lane closure have on the entire network. For instance, in the current model a lane closure increases the queue at that particular toll plaza, but it is unknown if special conditions also have a profound effect the operations at plazas downstream. The overall scope of future research should be rooted in focusing not only on plaza operations but how plazas interact with both other plazas and the network itself.

LIST OF REFERENCES

1. *Florida Turnpike - Frequently Asked Questions*. Florida Turnpike Enterprise. Florida Department of Transportation. http://www.floridasturnpike.com/about_faqs.cfm. Accessed Jan. 31, 2008.

2. Schmitt, Daniel T. Modeling Toll Plaza Performance – Decision Support System for Operatiors – SHAKER. Thesis, University of Massachusetts Dartmouth, 2003.

3. *Orlando Orange County Expressway Authority System's Traffic Data & Statistical Manual-Mainline Plaza Characteristics.* Orlando Orange County Expressway Authority. http://www.oocea.com/trafficstatistics/historicaltraffic. Accessed Jan. 31, 2008.

4. Klodzinski, J.G., Al-Deek, H. M., and Radwan, A.E. Evaluation of vehicle emissions at an electronic toll collection plaza. Presented at the 77th Annual Meeting of the Transportation Research Board. Washington, D.C. 1998.

5. Astarita, Vittorio, Michael Florian, and Giuseppe Musolino. A Microscopic Traffic Simulation Model for the Evaluation of Toll Station Systems. IEEE Intelligent Transportation Systems Conference Proceedings. 2001. Oakland, CA., August 25-19, 2001. pp. 692- 697.

6. Milam, R. T., *Recommended Guidelines for the Calibration and Validation of Traffic Simulation Models*. Richard T. Milam. Fehr & Peers Associates, Inc., http://www.fehrandpeers.com/publications/traff_simulation_m.html, Last viewed March 2002.

7. *Traffic Analysis Toolbox- Volume III*: Guidelines for Applying Traffic Microsimulation Modeling Software. Federal Highway Administration United States Department of Transportation. McLean, VA: Research, Development, and Technology Turner-Fairbank Highway Research Center, 2004.

8. Park, Byungkyu and J.D. Schneeberger. Microscopic Simulation Model Calibration and Validation: A Case Study of VISSIM for a Coordinated Actuated Signal System. Presented at the 82nd Annual Meeting of the Transportation Research Board. Washington, D.C. 2003.

9. Hellinga, B. R., Requirements for the Calibration of Traffic Simulation Models, Department of Civil Engineering, University of Waterloo, http://www.civil.uwaterloo.ca/bhellinga/Publications%20Page/Publications/CSCE-1998-Calibration.PDF, Last viewed August 2008.

10. Redding, Robert. T., Andrew J. Junga, *TPASS: Dynamic, Discrete-event Simulation and Animation of a Toll Plaza*, Proceedings of the 24th Conference on Winter Simulation, Arlington, Virginia, United States, December 13-16, 1992, pp. 1292-1295.

11. Al-Deek, H.M., A.A. Mohamed, L. Malone. *A New Stochastic Discrete-Event Micro Simulation Model for Evaluating Traffic Operations at Electronic Toll Collection Plazas*. Journal of Intelligent Transportation Systems; Oct-Dec2005, Vol. 9 Issue 4, p205-219, 15p.

12. Klodzinski, J. G., H.M. Al-Deek, Transferability of a Stochastic Toll Plaza Computer Model. Journal of the Transportation Research Board, No. 1811, December 2002, pp.40-49.

13. Klodzinski, J. Gordin, E., and Al-Deek, H., Evaluation ofImpacts from Deployment of an Open Road Tolling Concept for a MainlineToll Plaza, Journal of the Transportation Research Board, No. 2012, pp. 72-83, October 2007

14. Smith, Richard F. et al. State of the Practice and Traffic Control Strategies at Toll Plazas: Best Practices. Manual of Uniform Traffic Control Devices. Federal Highway Administration, Federal Highway Administration. June 2006.

15. Nezamuddin, N. and Al-Deek, H. Developing a Microscopic Toll Plaza and Toll Road Corridor Model Using PARAMICS, Journal of the Transportation Research Board, No. 2047, pp. 100-110, October 2008.

16. Ozbay, K., Sandeep Mudigonda, Bekir Bartin. Development and Calibration of an Integrated Freeway and Toll Plaza Model for New Jersey Turnpike using Paramics Microscopic Simulation Tool. Proceedings of the 8th International IEEE Conference on Intelligent Transportation Systems. Vienna, Austria, September 13-16, 2005.

17. Ceballos, Gustavo and Owen Curtis. Queue Analysis at Toll and Parking Exit Plazas: A Comparison between Multi-server Queuing Models and Traffic Simulation. VISSIM Library.

18. Park, B. Qi, H. Microscopic Simulation Model Calibration and Validation for Freeway Work Zone Network – A Case Study of VISSIM. Virginia Transportation Research Council and Federal Highway Administration. March 20, 2006.

19. Chitturi, Madhav V. and Rahim F. Benekohal. Calibration of VISSIM for Freeways. Presented at the 87th Annual Meeting of the Transportation Research Board. Washington, D.C. 2003.

20. Lownes, N. E., and R. B. Machemehl. Sensitivity of Simulated Capacity to VISSIM Driver Behavior Parameter Modification. Presented at the 85th Annual Meeting of the Transportation Research Board, Washington D.C., 2006.

21. Festa, Demetrio Carmine, Giovanni Longo, Gabriella Mazzulla, and Giuseppe Musolino. Experimental Analysis of Different Simulation Models for Motorway Traffic Flow. Proceedings from 2001 IEEE Intelligent Transportation Systems Conference, 25-29 Aug. 2001, Oakland, CA, USA.

22. Highway Capacity Manual. TRB, National Research Council, Washington, D.C., 2000.

23. Klodzinski, J. G. and H.M. Al-Deek. New Methodology for Determined Level of Service at Toll Plazas. Journal of Transportation Engineering, March 2002, Vol.128, No. 2, pp.1-9.

24. Klodzinski, J.G. and H.M. Al-Deek. Proposed Level of Service Methodology for Toll Plazas. Presented at the 81st Annual Meeting of the Transportation Research Board. Washington, D.C. 2002.

25. Aycin, Murat F. Simple Methodology for Evaluating Toll Plaza Operations. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1988*, Trzansportation Research Board of the National Academies, Washington, D.C., 2006, pp. 92-101.

26. Pietrzyk M.C. Electronic Toll Collection Systems. *Transportation Research Board Special Report 242, Curbing Gridlock*, National Research Council, Washington, D.C., 1994, Vol. 2, pp. 464-501.

27. Zarrillo, Marguerite L., A. Essam Radwan and Joseph H. Dowd, "Toll Network Capacity Calculator: Operations Management and Assessment Tool for Toll Network Operators", Transportation Research Record 1781, TRB, National Research Council, Washington, D.C., 2002.

28. Woo, T. Hugh and Lester A. Hoel. Toll Plaza Capacity and Level of Service. In *Transportation Research Record: Journal of the Transportation Research Board 1320*, Transportation Research Board, National Research Council, Washington, D.C.,1991, pp. 119-127.

29. Oliveira, Marcelo Leismann de and Cybis, Helena Beatriz Bettella, "An Artificial Neural Network Model for Evaluating Workers' Performance at Tollbooths" (June 2006). 1st International Symposium on Freeway and Tollway Operations, Athens, Greece Available at SSRN: http://ssrn.com/abstract=962730

30. Zarrillo, M.L. *Development and Applications of TPModel: A Queuing Model Describing Traffic Operations During Electronic Toll Collection (ETC),* Ph.D. Dissertation, University of Central Florida, Orlando, FL, 1998

31. Zarrillo, M.L., A.E. Radwan, and J. Dowd, “Toll Network Capacity Calculator: Operations Management and Assessment Tool For Toll Network Operators”, Transportation Research Board Journal, Record Number 1181, May 2002.

32. M. Zarrillo, and A.E. Radwan, “Identification of Bottlenecks on a Toll network of Highways”, 7th International ASCE Specialty Conference on Applications of Advance Technology in Transportation, August 2002, Boston, Massachusetts.

33. Zarrillo, M.L., A.E. Radwan and H.M. Al-Deek. Modeling Traffic Operations at Electronic Toll Collection and Traffic Management Systems. The International Journal on Computers and Industrial Engineering, editor Yeong-Dae Kim, special edition, December 1997, Volume 33, Numbers 3-4, pp. 857-860.

34. Mendenhall, William, and Terry Sincich. *Statistics for Engineers and the Sciences*. 5th ed. Upper Saddle River: Prentice Hall, 2006.

35. *VISSIM User Manual: Version 4.30*. P. T. V. AG., Karlsruhe, Germany, 2007.