

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

**Application of Microscopic Simulation Model
Calibration and Validation Procedure:
A Case Study of Coordinated Actuated Signal System**

by
Byungkyu "Brian" Park, Ph.D.*
Assistant Professor
Department of Civil Engineering, University of Virginia
and
Faculty Research Scientist
Virginia Transportation Research Council
PO Box 400742
Charlottesville, VA 22904-4742
Phone: 434-924-6347, Fax: 434-982-2951
Email: bpark@virginia.edu

Jongsun Won
Graduate Research Assistant
Department of Civil Engineering
University of Virginia
PO Box 400742
Charlottesville, VA 22904-4742
Phone: 434-996-7543, Fax: 434-982-2951
Email: jwon@virginia.edu

and

Ilsoo Yun
Graduate Research Assistant
Department of Civil Engineering
University of Virginia
PO Box 400742
Charlottesville, VA 22904-4742
Phone: 434-924-1420, Fax: 434-982-2951
Email: iy6m@virginia.edu

WORD COUNT : 4,712 + 11 FIGURES/TABLES × 250 = 7,462 WORDS

* Corresponding Author

ABSTRACT

In the application of microscopic simulation models, the importance of model calibration and validation cannot be overemphasized. A recent study proposed a systematic approach of conducting simulation model calibration and validation procedure on the basis of experimental design and optimization, and applied it to an isolated intersection using a VISSIM simulation model.

This study further evaluates the previously developed simulation model calibration and validation procedure using an urban arterial network consisted of 12 coordinated actuated signalized intersections. Both VISSIM and CORSIM simulation models were used. Travel time was used for calibration measure, while maximum queue length was used for validation measure. The study results showed that the calibrated and validated simulation models were able to adequately represent field conditions while default parameter based models were not. As such the previously developed simulation model calibration and validation procedure was proven to be effective for an arterial network under both VISSIM and CORSIM simulation models.

INTRODUCTION

As the surface transportation system becomes more complex and broader in its coverage area, it is very challenging to accurately analyze and evaluate the performance of the system itself before implementing new control strategies and/or physical changes. Obviously, the Highway Capacity Manual (HCM), the most widely used engineering guidebook in the analysis of surface transportation system, cannot be used for a large scale transportation system analysis. One of alternative options is the use of simulation models. Fortunately, current computational technology has advanced such an extent that one can simulate a sizeable network of surface transportation system within a reasonable amount of time. Not surprisingly, microscopic traffic simulation models become one of the most popular tools for traffic engineers to conduct transportation system analyses and evaluations. The microscopic traffic simulation models have been used for various occasions such as evaluating geometry changes, traffic signal timing plan updates, or the benefits estimation of new Intelligent Transportation Systems (ITS) strategies. It is noted that microscopic simulation models are much faster, safer and cheaper than actual field implementations. However, when a microscopic simulation model is used without proper calibration and validation, the results could be misleading and would lose public and/or decision maker's trust. Therefore, the simulation models need to be well calibrated and validated before using it for evaluation purpose.

Most microscopic simulation models require many different types of parameters that can describe traffic flow characteristics, traffic control systems, driving behavior, etc. Basic input parameters such as geometry, number of cars, and traffic signal setting are easy to obtain. However, some of calibration parameters that are related to drivers' behaviors are almost impossible to collect from the field. Thus, it is common practice to use either default parameters provided by the microscopic simulation model developer or minimally calibrated model based on engineering judgments. However, researchers have shown that the simulation models under default calibration parameters may not accurately represent field conditions and possibly produce unreliable results (9, 10).

In recent years, several researchers proposed microscopic traffic simulation model calibration and validation procedures and tested using various microscopic traffic simulation models. Sacks *et al.* (1) identified four key issues on model validation and demonstrated an informal validation process by using CORSIM. They emphasized the importance of data quality and visualization. Milam *et al.* (2) proposed a guideline for the calibration and validation of traffic simulation models. Cohen (3) recommended key parameters be adjusted close to field measured data. Dowling *et al.* (4) proposed a practical, top-down approach. Kim *et al.* (5) proposed non-parametric statistical technique for calibration and applied it to real network by using VISSIM. They claimed that simple metrics could lead to erroneous results.

In the process of microscopic simulation model calibration and validation, genetic algorithm (GA) has been widely applied. Cheu *et al.* (6) applied a GA as an optimization method to find a suitable combination of FRESIM parameter values for a real network. A similar effort for PARAMICS model calibration was made by Lee *et al.* (7) and their evaluation results turned out to be reliable. Kim and Rilett (8) conducted calibration process with GA by using both CORSIM and TRANSIMS, and showed the benefits of using a GA for automated calibration.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

However, most of these approaches used a few selected calibration parameters due to the complexity of the optimization surface (i.e., stochastic nature of microscopic simulation models and the total number of combinations when all the possible calibration parameters are considered). A few exceptions were Park and Schneeberger (9) and Park and Qi (10). They adopted a statistical experimental design approach to reduce the number of combinations and also considered feasibility of the initial ranges of calibration parameters. In addition, Park and Qi (10) emphasized the importance of feasibility testing. They showed if initial ranges of calibration parameters do not contain optimal solution (i.e., field condition), the simulation model cannot be calibrated even the optimization method finds optimal solution within the search region. Obviously, this is because the search region does not include field condition. Even though their approach was successfully tested at an isolated signalized intersection using VISSIM simulation model, it is desirable to test it with a sizable arterial network.

The objective of this paper is to evaluate the systematic simulation model calibration and validation procedure proposed by Park and Qi (10) using a coordinated actuated traffic signal system with both VISSIM and CORSIM simulation models.

The remainder of this paper is consisted as follows. The test bed network section presents test site descriptions, data collection efforts, and simulation models used in this study, while proposed procedure section provides details on the procedure developed by Park and Qi (10). Application of calibration and validation the procedure section provides implementations of the procedure with test bed network using VISSIM and CORSIM simulation models, followed by conclusions and recommendations.

TEST BED DEVELOPMENT

Test Site

A coordinated actuated signalized network was modeled in VISSIM and CORSIM microscopic simulation models. These networks are used for the applications of the proposed calibration and validation procedure. Test site contains an urban arterial, Lee Jackson Memorial Highway (U.S. Route 50), and twelve intermediate intersections between Sully Road and the Fairfax County Parkway in Fairfax, Virginia. The network consisted of eleven consecutive intersections on U.S. Route 50 and one adjacent intersection on Majestic Lane which was added to the network because it is located close enough to affect the traffic flow on U.S. Route 50. The test site is referred as “Route 50” throughout the remaining part of this paper.

Data Collection

The data needed for this application study included simulation input data and measures of effectiveness (MOE) data for calibration and validation purpose. It is noted that the data was obtained from previous study conducted by Park and Schneeberger (9) and a brief description on data collection process is shown below.

Geometry, traffic signal timing plan, and traffic counts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

Geometric characteristics such as distance, speed limits, detector locations, etc, for this Route 50 network were obtained from Synchro file that was provided by Virginia Department of Transportation (VDOT). The signal timing plans for intersections within test site were extracted from Management Information System for Transportation (MIST) terminal located at the Smart Travel Laboratory at the University of Virginia. Traffic counts were collected on normal weekday, Wednesday July 11, 2001, between 4:45 p.m. and 6:15 p.m. via video cameras and human counters.

Measures of performance data

Travel time data was selected as a calibration data because of its easiness to collect from the field and availability of simulation outputs. Also, it directly reflects level of service (LOS) of signalized arterial network. Two cameras were installed and recorded each vehicle's license plate on the both ends of the corridor to capture the travel time of left-most lane on Route 50. After data collection from the field, all the license plate data were extracted from the videotape and matched to measure a travel time of each vehicle. The average travel time of the vehicles on the left-most lane was 613.2 seconds.

In addition to the calibration data, validation data was also collected using the Smart Travel Van equipped with an AUTOSCOPE video detection system. The maximum queue length data was collected on a different weekday from when the travel time data was collected on August 2001, so it can be used as an untested data set for validation process. The maximum queue length for each cycle was counted at the end of each red time and field measured maximum queue length of 24 vehicles was obtained.

Simulation Models

VISSIM (Verkher In Staedten SIMulation) was selected as a first simulation model. It was developed at the University of Karlsruhe, Germany during 1970s. VISSIM is a microscopic and behavior based simulation model that uses psychophysical driver behavior model developed by Wiedemann (1991) which uses stochastic car-following models and dynamic speeds. It consisted of two different programs which are traffic simulator and signal state generator. Traffic simulator comprised of car-following and lane changing logic and it is capable of simulating up to one tenth of second. VISSIM version 4.0 was used in this paper.

CORSIM (CORridor SIMulation) was selected as a second microscopic simulation model which is also a stochastic microscopic simulation model. It was first developed by the Federal Highway Administration (FHWA) during 1970s. CORSIM is a core component of Traffic Software Integrated System (TSIS) package which is one of the most widely used microscopic simulation models in U.S. CORSIM combines two different models which are the arterial network model NETSIM and the freeway network model FRESIM. Traffic flow algorithms such as car-following, traffic control module, etc. in CORSIM simulates the interactions of individual vehicles. TSIS version 5.1 was used in this paper.

Performance measures such as travel times and maximum queue length are extracted from the simulation models. In the case of VISSIM, performance measures were provided as a standard

text format at locations specified by the user. For example, path travel times were obtained from two data collection points. However, CORSIM does not provide path travel times as a standard output. Thus, they were extracted from TSD file that stores individual vehicle information during the simulation.

PROPOSED PROCEDURE

Previously proposed microscopic simulation model calibration and validation procedure by Park and Qi (10) was applied to evaluate an arterial network by using VISSIM and CORSIM. In order to provide readers with the completed procedure, a brief description of the procedure is provided in this section. Readers who wanted to access to the entire procedure should refer to Park and Qi (10). The procedure is also depicted in a flow chart (Figure 1).

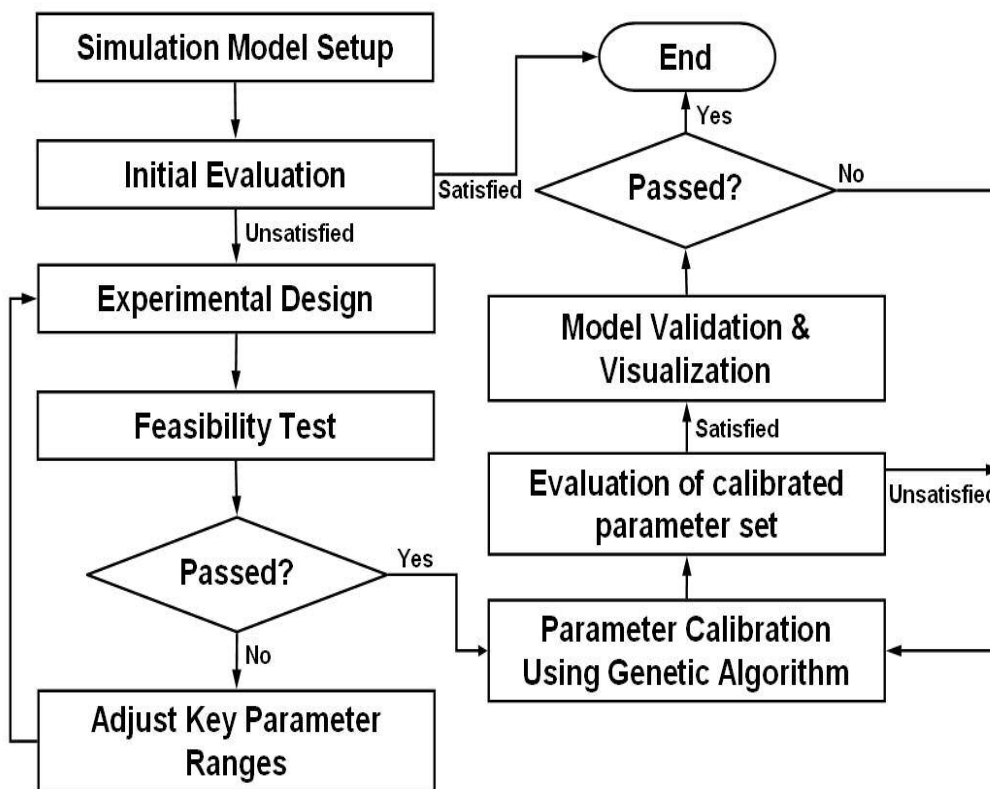


Figure 1. Calibration and Validation Procedure Flow Chart [Modified from Park and Qi (10)]

Simulation Model Setup

Simulation model setup includes defining study scope and purpose, site selection, measure of effectiveness (MOE) determination, data collection and network coding.

Initial Evaluation

This step is to test whether default calibration parameters in the simulation model is sufficient to represent field conditions or not.

Initial Calibration

Initial calibration step consisted of three steps: (1) identify calibration parameters and their acceptable ranges, (2) conduct statistical experimental design and generate reasonable number of parameter sets, and (3) implement multiple runs with each parameter set.

Feasibility Test

This step is to verify whether the distribution of simulation results from previous multiple runs includes field data. If the distribution includes the field data, then it is determined as the acceptable range. Otherwise, modifications on the existing parameter sets need to be conducted to adjust either the ranges or the list of parameters.

Parameter Calibration Using Genetic Algorithm

This step is to find an optimal calibration parameter set using a genetic algorithm (GA) optimizer. During the GA optimization, multiple runs were conducted for each feasible parameter set to consider the variability of simulation results. Fitness value for each individual was calculated by comparing field data and simulation results using Equation 1.

$$FV = \frac{|TT_{Field} - TT_{Sim}|}{TT_{Field}} \quad (\text{Eq. 1})$$

where

FV = Fitness value

TT_{Field} = Average field-measured travel time

TT_{Sim} = Average travel time output from multiple simulation runs with same parameter set

Evaluation of the Parameter Set

Once the GA finds an optimal calibration parameter set, multiple runs (e.g., 100 runs) with the calibrated parameter set are conducted to compare the performance of the simulation model with field data. Thus, a distribution of calibrated simulation model outputs is compared with field data. A visualization testing is also conducted.

Validation and Visualization

In order to validate the calibrated simulation model, it is desirable to collect a new set of data under different condition (e.g., new traffic signal timing plan, new traffic volume, etc.). Thus, in this case study, the data collected on different day was used as a new data set. Again, multiple runs (e.g., 100 runs) are to be made and a distribution of simulation model outputs is compared with validation field data. If field data falls within the 95 percentile range of the model output distribution, the calibrated model is considered to be valid. In addition, visualization testing is also conducted for a few selected runs (e.g., 25th, 50th and 75th percentile runs).

APPLICATION OF CALIBRATION AND VALIDATION PROCEDURE

VISSIM Case Study

The proposed procedure was applied to the Route 50 network using VISSIM. This part describes the details on each step within the calibration and validation procedure for VISSIM.

Simulation Model Setup

The Route 50 network was coded into the VISSIM simulation model and data collection points for calibration and validation measures were installed.

Initial Evaluation

A total of 100 simulation runs with the default parameters were conducted and the distribution of simulation outputs (i.e., travel time) is compared with the field travel time. As shown in Figure 2, the simulated travel time (average of 724.3 seconds) was significantly higher than field travel time of 613.2 seconds. Therefore, it was decided to conduct the calibration and validation procedure.

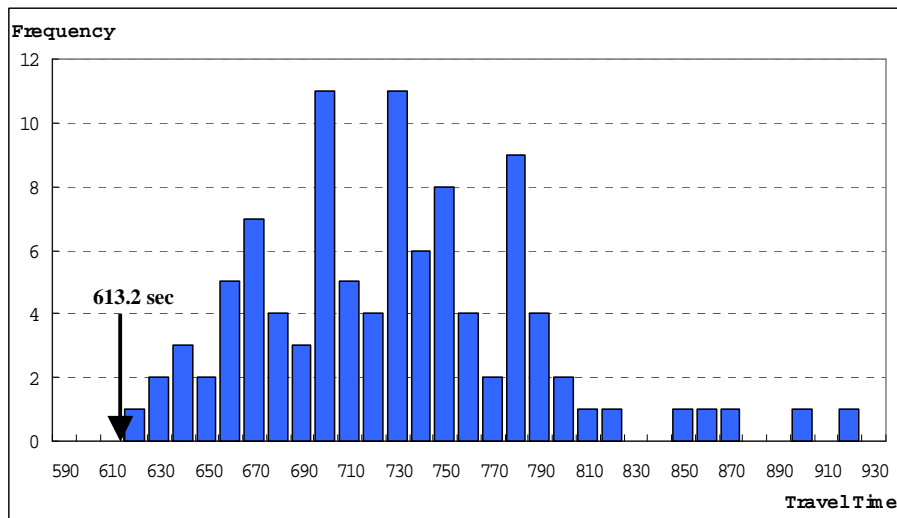


Figure 2. Initial Evaluation Results for Route 50 - VISSIM

Initial Calibration

Identification of Calibration Parameters

As an initial phase, 14 calibration parameters in the VISSIM model were selected. These include car following model, lane change behavior and priority rules. The following is the initially determined list of parameters and their ranges. The detailed descriptions of each parameter could be found in VISSIM user manual (11). The simulation resolution is included as it impacts on the response to traffic controls such as priority rules or traffic signals.

1. Simulation resolution (Time steps/Simulation second): 1 – 9
2. Number of observed preceding vehicles: 1 – 4
3. Maximum look ahead distance (meter): 200 – 300
4. Average standstill distance (meter): 1 – 5
5. Saturation flow rate

- Additive part of desired safety distance: 1.0 – 5.0
- Multiple part of desired safety distance: 1.0 – 6.0
6. Priority rules
 - Minimum gap time (second): 3 – 6
 - Minimum headway (meter): 3 – 20
7. Desired speed distribution (mph): 30~60, 35~55, 40~50
8. Waiting time before diffusion (second): 20 – 40
9. Minimum headway (meter): 0.5 – 7.0
10. Maximum deceleration (m/s^2): -5.00 – -1.00
11. Reduction rate (meter): 20 – 80
12. Accepted deceleration (m/s^2): -1.50 – -0.10

Experimental Design for Calibration

When 14 calibration parameters with 5 possible values for each parameter are considered, the total number of possible combination is 5^{14} (i.e., 6,103,515,625). This is almost impossible to evaluate within a reasonable amount of time. Thus, a total of 200 evaluation parameter sets were generated by using Latin Hypercube sampling method (13). The method was coded in a Visual Basic Application in MS Excel. The number of evaluation parameter sets should be determined on the basis of the number of parameters considered, the size of network, computation run time, etc. This study used 200 sets based on the previous experience that have been accumulated by the authors and it seemed quite effective. Of course, more is better even though it is costly.

Multiple Runs

For 200 calibration parameter sets, five random seeded runs were conducted for each set to consider variability of the microscopic simulation outputs. As a result, 1,000 runs were made and average travel time for each five runs was recorded. A distribution of average travel times for 200 cases was prepared for feasibility testing.

Feasibility Test

The distribution of 200 average travel times developed in the previous step is compared with field measured travel time. As shown in Figure 3, the field travel time lies within the distribution indicating that an optimal parameter set can be found from the calibration parameters and their ranges. Thus, the parameters and their ranges were considered as feasible and next step is implemented.

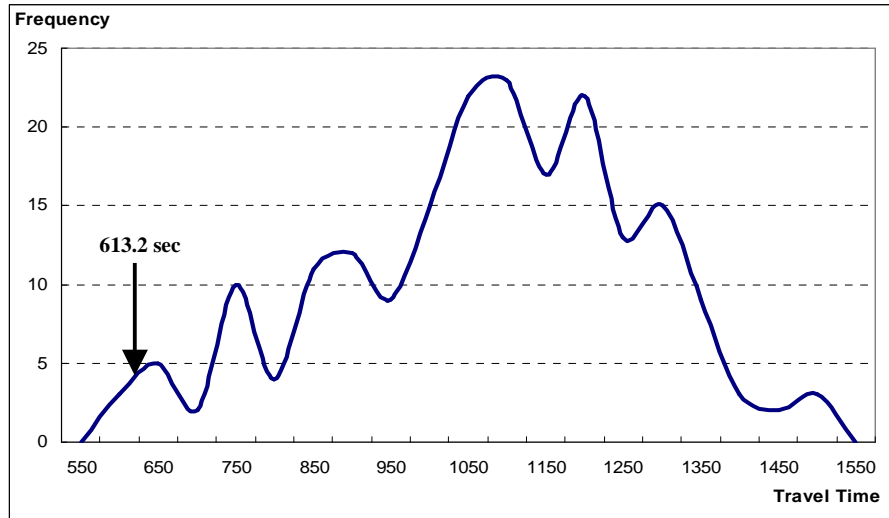


Figure 3. Feasibility Test Result for U.S. Route 50 – VISSIM

Parameter Calibration Using Genetic Algorithm

A GA program, integrated with the VISSIM model, was used to calibrate parameters within the feasible ranges. The GA parameters used in this study included maximum number of generations of 10, population size of 10, crossover rate of 0.8, and mutation rate of 0.05. The GA program converged within 10 generations. A few multiple GA runs showed similar convergences.

Evaluation of the Parameter Sets

An optimal calibration parameter set (named as 1st trial) was selected from the best solution at the last generation. A 100 multiple simulation runs were made with the optimal parameter set and the distribution of the travel time from the simulation model outputs was compared with the field travel time. Even though, as shown in Figure 4, the optimal calibration parameters show a good match with field data, the validation results were not acceptable. Another GA-based calibration parameter optimization attempt (named as 2nd trial) was made and the distribution of travel times were developed using 100 multiple runs. Figure 6 shows travel time distributions of both trials and field travel time. Visualization testing conducted for both calibrated models indicated that both models were acceptable.

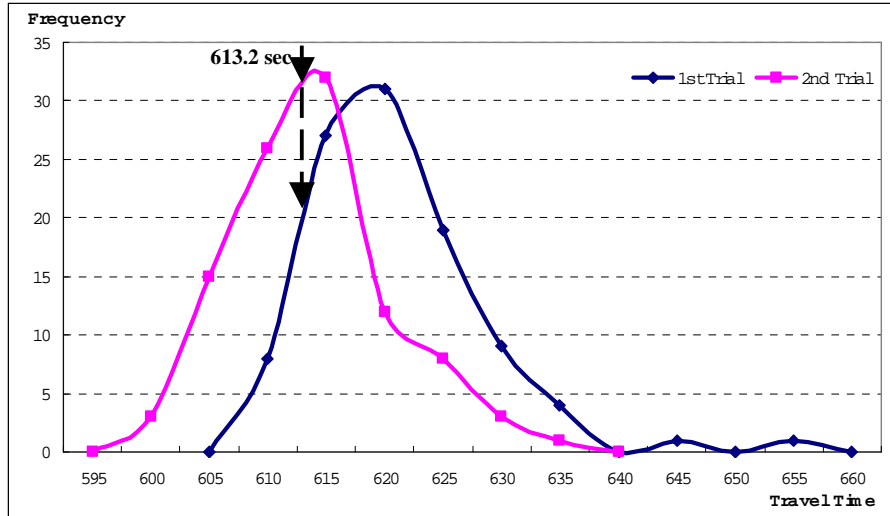


Figure 4. Travel Time Distribution of GA-based Parameter Set – VISSIM

Validation with a new data and visualization testing

In order to validate the calibrated parameter set identified from the previous section, a new field data in which was not exposed during calibration was used. A distribution of maximum queue lengths obtained from VISSIM simulation runs was compared with field measured maximum queue length. As briefly mentioned, the results from the firstly calibrated model (1st trial) were not acceptable (see Figure 5). This triggered the second optimization and its results (2nd trial) were acceptable. Furthermore, the visualization testing was conducted by carefully watching simulation animations during validation procedure. No abnormal vehicle movements were observed.

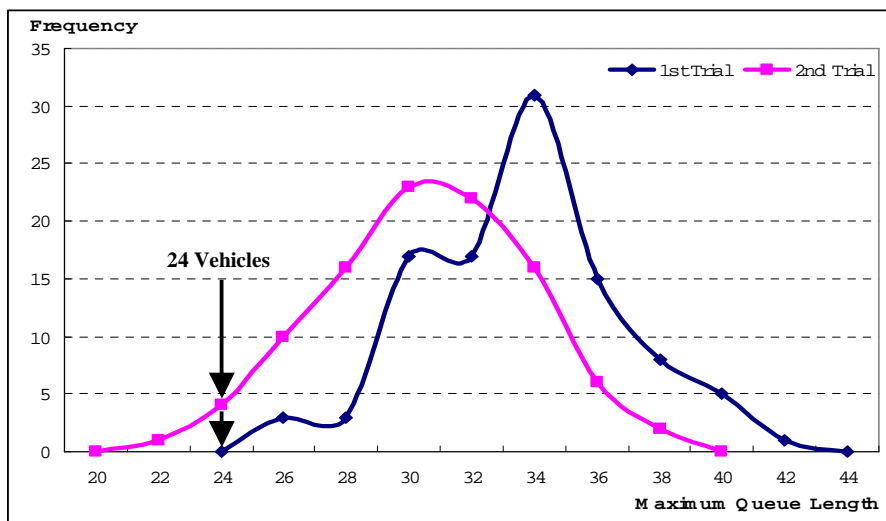


Figure 5. Maximum Queue Length Distribution of Calibrated Parameter Set - VISSIM

For comparison purpose, three calibration parameter sets (i.e., default, 1st trial and 2nd trial) and their average travel times are summarized in Table 1. Based on the ANOVA analysis, the most significant key parameters were additive part of desired safety distance and desired Speed Distribution.

Table 1. Comparison of three parameter sets for the Route 50 - VISSIM

Route 50 - VISSIM	Default	Calibrated (1st Trial)	Calibrated (2nd Trial)
Simulation Resolution	5	2	5
Number of observed preceding vehicles	2	1	1
Maximum look ahead distance (meter)	250.00	244.44	273.74
Average standstill distance (meter)	2.00	1.93	1.73
Additive part of desired safety distance	3.00	1.53	1.08
Multiple part of desired safety distance	3.00	2.67	2.57
Priority rules - Minimum gap time (second)	5.0	4.58	5.82
Priority rules - Minimum headway (meter)	8.0	11.06	9.09
Desired speed distribution (mph)	40-50	30-60	30-60
Waiting time before diffusion (second)	60.00	38.59	32.73
Minimum headway (meter)	0.50	1.68	1.75
Maximum deceleration (m/s^2)	-4.00	-4.47	-3.26
Reduction rate (meter)	100.00	50.30	52.12
Accepted deceleration (m/s^2)	-1.00	-0.41	-0.28
Average travel time (second)	724.3	617.6	611.6

CORSIM Case Study

CORSIM was selected as another microscopic simulation model tested for the microscopic simulation model calibration and validation procedure. The application procedure to the CORSIM simulation model and the results are presented in this section.

Simulation Model Setup

The network geometry, signal timing plan, traffic volume, and other input variables were coded into CORSIM.

Initial Evaluation

Multiple simulation runs with the default parameter set were conducted and average travel time for each run was extracted. Then, a distribution of CORSIM generated travel times was compared with field travel time. As shown in Figure 6, the travel time distribution based on the default parameter set was not even close to the field average travel time. Thus, it is deemed that the calibration and validation procedure should be implemented.

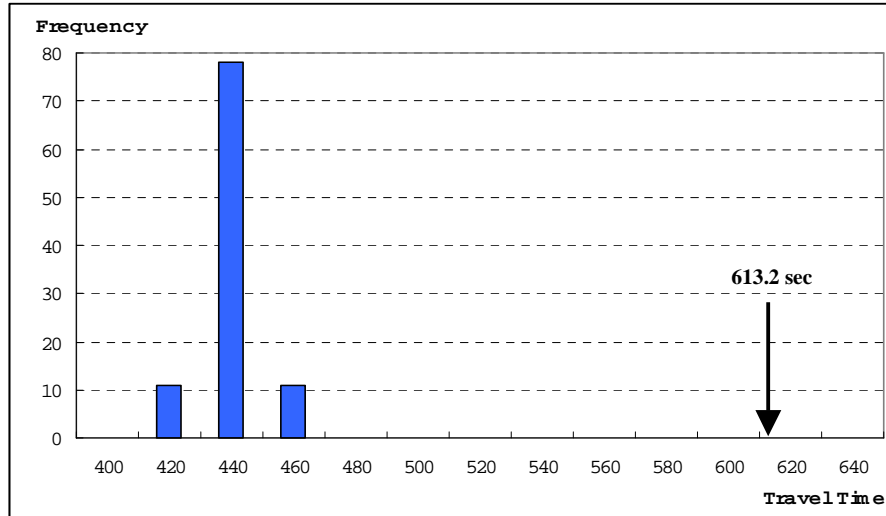


Figure 6. Initial Evaluation Results for Route 50 - CORSIM

Initial Calibration

Identification of Calibration Parameters

CORSIM provides numerous calibration parameters which user can fine-tune them to replicate field conditions. CORSIM contains several distinctive parameters that are represented by a discrete distribution or 10 percentile values indexed by driver types. In order to change behaviors of driver characteristics, parameters that control proportions of driver types are included. The following is the initially determined list of parameters and their ranges. The detailed description of each parameter could be found in CORSIM user's guide (4).

1. Link mean free flow speed (mph): -10, 0, +10 (from each free flow speed)
2. Mean queue discharge headway (second): 1.5 – 2.5
3. Mean start-up lost time (second): 1.5 – 3.0
4. Left turn jumper probability (%): 10 – 40
5. Left turn speed (mph): 13 – 31
6. Right turn speed (mph): 13 – 25
7. Left-turn lagging within 2 seconds (%): 20 – 50
8. Left-turn lagging for 2-4 seconds (%): 5 – 15
9. Amber interval response (fps²)
 - Shift to left: 19, 16, 13, 10, 7, 5, 4, 3, 2, 2
 - Shift to right1: 23, 20, 17, 14, 11, 9, 8, 7, 6, 6
 - Shift to right2: 25, 22, 19, 16, 13, 11, 10, 9, 8, 8
10. Gap distribution for left turns (second)
 - Shift to left: 6.8, 5.6, 5.0, 4.4, 3.8, 3.5, 3.2, 2.9, 2.6, 1.7
 - Shift to right: 8.8, 7.6, 7.0, 6.4, 5.8, 5.5, 5.2, 4.9, 4.6, 3.7
11. Gap distribution for right turns (second)
 - Shift to left1: 9.0, 7.8, 7.0, 6.2, 5.4, 5.0, 4.6, 4.2, 3.8, 2.6
 - Shift to left2: 8.0, 6.8, 6.0, 5.2, 4.4, 4.0, 3.6, 3.2, 2.8, 1.6

12. Distribution of free flow speed by driver type (%)

Narrow (0.8): 82, 86, 94, 96, 98, 100, 105, 108, 112, 119

Wider (1.2): 73, 80, 91, 94, 97, 100, 107, 112, 118, 128

13. Start-up lost time distribution (%)

Narrow (0.8): 195, 132, 120, 115, 102, 89, 82, 70, 57, 38

Wider (1.2): 240, 147, 130, 121, 102, 83, 74, 56, 37, 10

14. Discharge headway distribution (%)

Narrow (0.8): 156, 116, 116, 108, 100, 100, 92, 76, 76, 60

Wider (1.2): 184, 124, 124, 112, 100, 100, 88, 64, 64, 40

Experimental Design

The 200 candidate parameter sets were generated using a Latin Hypercube sampling method. It is again noted that the number of evaluation parameter sets should be determined on the basis of the number of parameters, the size of network, computation run time, etc. The 200 sets seemed working well for this study.

Multiple Runs

Five random seeded simulation runs were conducted in CORSIM for each of 200 parameter sets, for a total of 1,000 runs. An average travel time for each five runs was recorded.

Feasibility Test

A distribution of average travel times for 200 parameter sets was prepared for feasibility testing. The distribution was compared with field travel time. As shown in Figure 7, the travel time distribution obtained from the initial parameters and their ranges definitely contain field travel time. Thus, the next step is conducted.

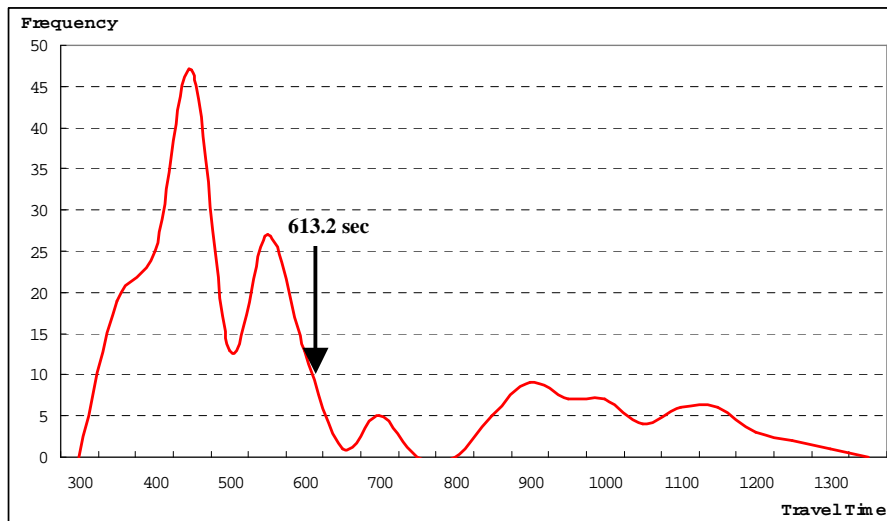


Figure 7. Feasibility Test Result for U.S. Route 50 - CORSIM

Parameter Calibration Using Genetic Algorithm

The GA-based optimization run was conducted and the optimal calibration parameter set was obtained. The GA parameters used in the CORSIM calibration were the same as those used in the VISSIM calibration.

Evaluation of the Parameter Sets

Again, 100 multiple CORSIM runs were made using the optimal parameter set found in the GA optimization. The distribution of CORSIM produced travel times was compared with field travel time. As shown in figure 8, the field travel time reasonably lies within the distribution of travel times. Thus, it was determined that the CORSIM model is well calibrated.

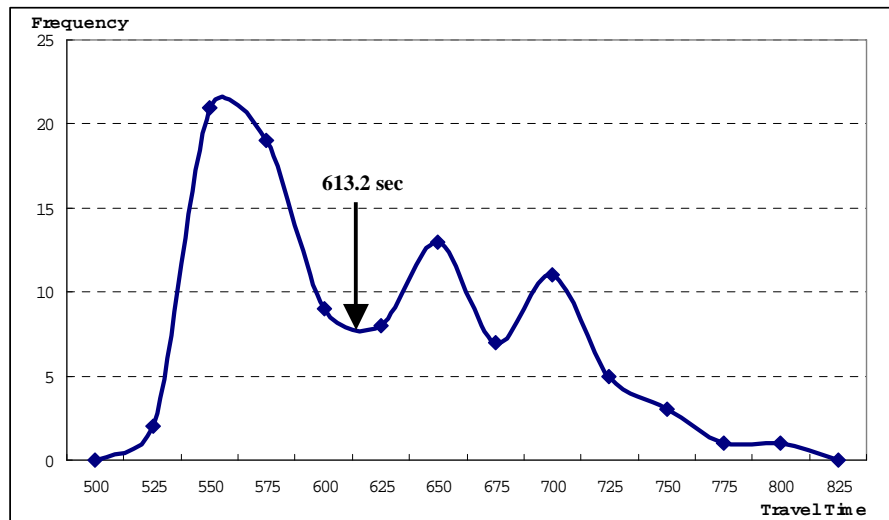


Figure 8. Travel Time Distribution of GA-based Parameter Set - CORSIM

Validation with a New Data

In order to validate the calibrated CORSIM simulation model, a maximum queue length data set in which was not exposed during the calibration optimization was used. The maximum queue length data from a predetermined approach was extracted from each of 100 simulation runs. Figure 9 shows the distribution of maximum queue lengths obtained from CORSIM. It is clear that the field data lies within the distribution. In addition, the animations of the CORSIM models were carefully checked for visual verification and no abnormal situations were observed. As such, the validation of the calibrated model is considered to be satisfactory.

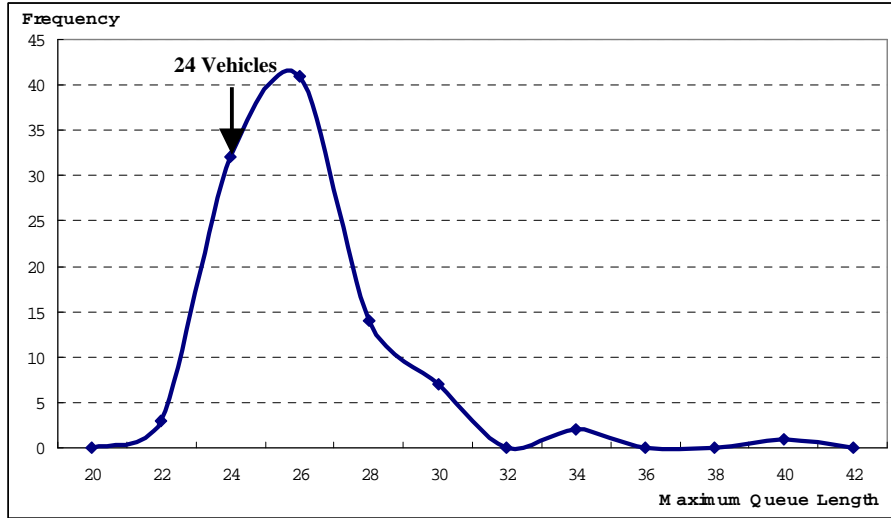


Figure 9. Maximum Queue Length Distribution of Calibrated Parameter Set - CORSIM

For comparison purpose, default and calibrated simulation parameter sets as well as their average travel times are summarized in Table 2. Apparently, some significant changes were made between these two parameters sets. In general, the calibrated parameters made the CORSIM model to be more conservative than the default parameters to match with field conditions (i.e., both travel time and maximum queue length).

Table 2. Comparison of two parameter sets for the Route 50 - CORSIM

Route 50 - CORSIM	Default	Calibrated
Link mean free flow speed (mph)	45 – 55	35 – 45
Mean queue discharge headway (second)	1.8	2.5
Mean start-up lost time (second)	2.0	4.0
Left turn jumper probability (%)	38	13
Left turn speed (ft/sec)	22	17
Right turn speed (ft/sec)	13	40
Left turn lagging within 2 seconds (%)	50	10
Left turn lagging within 2-4 seconds (%)	15	25
Amber interval response (fps ²) Index	1	2
Gap distribution for left turns (sec) Index	1	2
Gap distribution for right turns (sec) Index	1	1
Distribution of free flow speed by driver type (%) Index	1	2
Start-up lost time distribution (%) Index	1	1
Discharge headway distribution (%) Index	1	2
Average travel time (second)	429.7	609.6

CONCLUSIONS AND RECOMMENDATIONS

This paper applied previously developed microscopic simulation model calibration and validation procedure, which was only tested with an isolated signalized intersection, for the calibration and validation of a coordinated actuated arterial network with two microscopic simulation models: VISSIM and CORSIM. The performance of calibration and validation procedure was evaluated by comparing the distribution of simulation outputs and an average value of field data. The travel time measure was used for calibration, while the maximum queue length measure was used for validation.

The results based on Route 50 network indicate that the default parameters of both VISSIM and CORSIM simulation models were not able to replicate field travel times. The calibrated parameters obtained from the procedure for both models showed that the field travel time lies within the distributions of travel times – indicating that the optimal parameters were well calibrated. In addition, maximum queue length (a validation data set) was used for verifying that the calibrated models can replicate additional field data that were not used for calibration. In VISSIM, the calibrated parameters did not pass the validation test at the 1st trial, and thus 2nd trial was made and it was successful. This indicates that the validation test should be conducted to make sure that the calibrated model can be used for untried conditions. In the case of CORSIM simulation model, the simulation outputs of the calibrated parameters were able to replicate field measured validation condition. Thus, this study validated that the previously developed microscopic simulation model calibration and validation procedure can be used for a sizeable signalized network and other microscopic simulation models.

In this paper, only one field collected data point was used during each of calibration and validation tests. It is more desirable to collect multiple data points, collected over time, of field data to account for the day-to-day variability. In addition, more data can be collected to reduce the number of calibration parameters such as standstill distance or saturation flow rate. Although a sizable signalized arterial network was used in this study, more complex network that includes multiple arterials should be considered to verify the performance of this procedure. Furthermore, it is more desirable to consider many different calibration parameters such as parameters related to the dynamic origin-destination (OD) demand and multiple MOEs to understand inherent variability associated with the simulation models.

REFERENCES

1. Sacks, J., N. Rouphail, B. Park, and P. Thakuriah, Statistically Based Validation of Computer Simulation Models in Traffic Operations and Management. *Journal of Transportation and Statistics*, 2002, Vol. 5, No. 1, 2002, pp. 1-24.
2. Milam, R. T. Recommended Guidelines for the Calibration and Validation of Traffic Simulation Models. *Fehr & Peers Associates, Inc.*
http://www.fehrandpeers.com/Papers/Guidlines_traff_simulation.pdf. Accessed Jul. 2005.
3. Cohen, S. L. "An Approach to Calibration and Validation of Traffic Simulation Models." 83rd Annual Meeting Preprint CD-ROM, *Transportation Research Board*, Washington, DC, 2004.

4. Dowling, R., A. Skabardonis, J. Halkias, G. M. Hale, and G. Zammit. "Guidelines for Calibration of Microsimulation Models: Framework and Applications." 83rd Annual Meeting Preprint CD-ROM, *Transportation Research Board*, Washington, DC, 2004.
5. Kim, S., Kim, W., and Rilett, L. R. "Calibration of Micro-Simulation Models using Non-Parametric Statistical Techniques." 84th Annual Meeting Preprint CD-ROM, *Transportation Research Board*, Washington, DC, 2005.
6. Cheu, R. L., X., Jin, K. C., Ng, Y. L. Ng, and D. Srinivasan. "Calibration of FRESIM for a Singapore expressway using genetic algorithm." *Journal of Transportation Engineering*, Vol. 124, No. 6, 1998, pp. 526-535.
7. Lee, D. H., Y. Xu, and P. Chandrasekar. "Parameter Calibration for PARAMICS Using Genetic Algorithm." 80th Annual Meeting Preprint CD-ROM, *Transportation Research Board*, Washington, DC, 2001.
8. Kim, K-O and L. R. Rilett. "A Genetic Algorithm Based Approach to Traffic Microsimulation Calibration Using ITS Data." 83rd Annual Meeting Preprint CD-ROM, *Transportation Research Board*, Washington, DC, 2004.
9. Park, B., and J. D. Schneeberger. "Microscopic Simulation Model Calibration and Validation: A Case Study of VISSIM for a Coordinated Actuated Signal System." *Transportation Research Record 1856*, TRB, National Research Council, Washington, D.C., 2003, pp. 185-192.
10. Park, B., and Qi, H. "Development and Evaluation of Simulation Model Calibration Procedure." 84th Annual Meeting Preprint CD-ROM, *Transportation Research Board*, Washington, D.C., 2005.
11. *VISSIM Version 4.0 User Manual*. PTV Planug Transport Verkehr AG. Innovative Transportation Concepts, Inc., June 2004.
12. *CORSIM User's Manual*. Federal Highway Administration, U.S. Department of Transportation, Office of Safety and Traffic Operations, McLean, VA, 1997.
13. McKay, M. D., R. J. Beckman, and W. J. Conover. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, Vol. 21, No. 2, 1979, pp. 239-245.