# MICROSCOPIC SIMULATION MODEL CALIBRATION AND VALIDATION HANDBOOK

Traffic Operations Laboratory
Center for Transportation Studies
University of Virginia

VTRC
Virginia Transportation
Research Council

Standard Title Page - Report on Federally Funded Project

1. Report No. FHWA/VTRC 07-CR6	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Microscopic Simulation Model Calibration and Validation Handbook		5. Report Date October 2006	
		6. Performing Organization Code	
7. Author(s) Byungkyu (Brian) Park and Jongsun Won		8. Performing Organization Report No. VTRC 07-CR6	
9. Performing Organization and Address Virginia Transportation Research Council 530 Edgemont Road Charlottesville, VA 22903		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No. 75583	
12. Sponsoring Agencies' Name and Address Virginia Department of Transportation FHWA 1401 E. Broad Street P.O. Box 10249 Richmond, VA 23219 Richmond, VA 23240		13. Type of Report and Period Covered Final: December 2004 – September 2006	
		14. Sponsoring Agency Code	
15. Supplementary Notes			

16. Abstract

Microscopic traffic simulation models are widely used in the transportation engineering field. Because of their cost-effectiveness, risk-free nature, and high-speed benefits, areas of use include transportation system design, traffic operations, and management alternatives evaluation. Despite their popularity and value, the credibility of simulation models falls short due to the use of default parameters without careful consideration. Improper model parameters prevent simulation models from accurately mimicking field conditions, limiting their ability to aid decision-making. Therefore, the user needs to pay more attention to fine-tune each model that they are using by calibrating the parameters inside the model. To summarize, we can define calibration as the adjustment of model parameters such that the model's output more closely represents field conditions.

The intention of this handbook is to outline and explain the calibration and validation procedure for the parameters controlling human and vehicle characteristics for CORSIM and VISSIM.

| 17 Key Words
Calibration and Validation, Microscopic Simulation Models, CORSIM, VISSIM, Handbook, Experimental Design, Genetic Algorithm, Optimization | | 18. Distribution Statement
No restrictions. This document is available to the public through NTIS, Springfield, VA 22161. | |
| 19. Security Classif. (of this report)
Unclassified | 20. Security Classif. (of this page)
Unclassified | 21. No. of Pages
174 | 22. Price |

Form DOT F 1700.7 (8-72)   Reproduction of completed page authorized.

FINAL CONTRACT REPORT

MICROSCOPIC SIMULATION MODEL CALIBRATION AND VALIDATION HANDBOOK

Byungkyu "Brian" Park, Ph.D.
Faculty Research Scientist
and
Assistant Professor of Civil Engineering

Jongsun Won
Graduate Research Assistant

Traffic Operations Laboratory
Center for Transportation Studies
University of Virginia

Project Manager
Michael A. Perfater, Virginia Transportation Research Council

Table of Contents

List of Figures

List of Tables

Introduction

Microscopic traffic simulation models are widely used in the transportation engineering field. Because of its cost effectiveness, risk-free nature, and high-speed benefits, areas of use include transportation system design, traffic operations, and management alternatives evaluation. Figure 1 shows applications and benefits of microscopic simulation models.
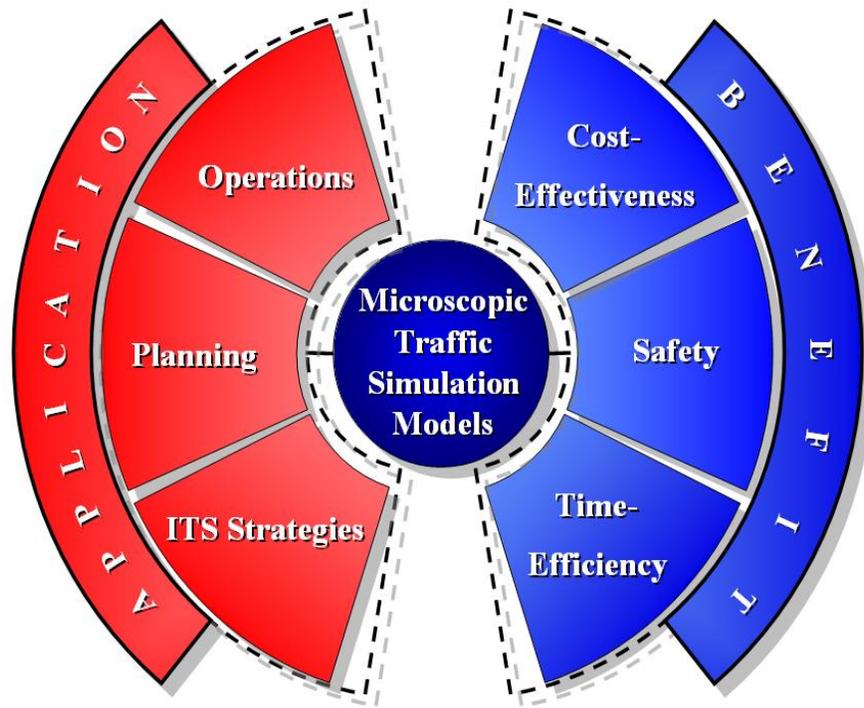
Figure 1. Application Area and Benefits of Microscopic Traffic Simulation Models

Up to this point in development, many different microscopic traffic simulation models are available to be applied to the various projects and research, the most popular existing models being:

- CORSIM (FHWA, 2003)
- VISSIM (PTV America, 2005)
- PARAMICS (Quadstone Limited, 2002)
- AIMSUN (TSS, 2006)
- WATSIM (KLD Associates, 2006)
- TRANSIMS (Barrett et al., 1995)
- MITSIM (Yang, 1997)
- SimTraffic (Trafficware, 2003)
- INTEGRATION (Van Aerde et al., 1996)

All of these models have achieved certain degrees of success in application, including traffic operations, transportation planning and intelligent transportation systems (ITS) strategies.

Despite their popularity and value, the credibility of simulation models falls short due to the use of default parameters without careful consideration. Improper model parameters prevent simulation models from accurately mimicking field conditions, limiting their ability to aid decision-making. Therefore, the user needs to pay more attention to fine-tune each model that they are using by calibrating the parameters inside the model. To summarize, we can define calibration as the adjustment of model parameters such that the model's output more closely represents field conditions.

The intention of this handbook is to outline and explain the calibration and validation procedure for the parameters controlling human and vehicle characteristics for CORSIM and VISSIM.

STOP **Calibration process**

When running traffic simulation models, distinguished calibration inputs among model inputs (e.g., Number of vehicles, Heavy vehicle percentage, etc.) for the simulation model are the most important thing. Such inputs that explain the driver's behavior and vehicle characteristics are difficult to collect from the field. Consequently, the user of simulation models needs to fine-tune all inputs that are related to the driving behavior and vehicle characteristics by comparing and adjusting some absolute measures. We call this overall process "Calibration."

Calibration Process Overview

The procedural guidelines presented in this handbook can be used to calibrate some parameters in simulation modeling. Prior to starting, it is important to understand the overall process for achieving calibration.

The procedure for calibration and validation can be divided into seven main steps:

1) The microscopic simulation model is set up in the usual manner for the particular simulation model. Details on base network coding for each simulation model are not covered in this handbook because they can easily be found in the user manual of each simulation model.
2) The feasibility of the default parameter set is evaluated. If the default parameter set yields acceptable results (results that accurately reproduce field conditions), it can be used for further process without calibration. If not, the next step needs to be considered.
3) Each calibration parameter is carefully examined and an acceptable parameter range is determined. This is done by manually defining an acceptable range for each parameter and analyzing its feasibility based on the knowledge that the engineer has for that specific site, utilizing graphical and statistical methods.
4) Optimization methodology (Genetic Algorithm) is applied to find the best calibration parameter set from pre-determined acceptable parameter ranges.
5) The performance (similarity of simulation model with the field condition) of all different parameter sets—default set and calibrated parameter set—is compared by observing the result of multiple runs of simulation models and field measured data.

6) The reality of the simulation model is checked once again with the animation that each simulation model is providing to find unrealistic conditions from the animation.
7) Validation is the final step where a final check is performed with an unused set of field data such as new field conditions or different day conditions.

A flowchart for the calibration and validation process is presented in Figure 2. Detailed techniques and explanations are given in each chapter.

① SIMULATION MODEL SETUP

- Determination of Measure of Effectiveness
- Field Data Collection
- Network Coding

See Chapter 1

② INITIAL EVALUATION

- Default Parameter set
- Multiple run

See Chapter 2

③ INITIAL CALIBRATION

- Identification of Calibration Parameters
- Experimental Design
- Multiple run

See Chapter 3

④ FEASIBILITY TEST

- X-Y Plots
- Statistical test
- Identification of Key Parameters

⑤ ADJUSTMENT

- Adjustment of key Parameters

See Chapter 4

⑥ PARAMETER CALIBRATION

- Genetic Algorithm
- Multiple run

See Chapter 5

⑦ EVALUATION OF PARAMETER SET

- Statistical test
- Visualization Check

⑧ MODEL VALIDATION

- Multiple run

See Chapter 6

FEEDBACK

Figure 2. Calibration and Validation Process Flow Chart

4

Handbook Organization

This handbook is divided into three parts and each part discusses specific aspects of the calibration and validation procedure.

Part I: Before Beginning Calibration Process

Part I discusses simulation setup that must be completed before starting the calibration and validation process, including the network building process and relevant data collection from the field.

Part II: Starting Calibration Process

Part II describes the entire calibration and validation procedure. It includes testing the default value of the calibration parameter set, verifying the acceptable calibration parameter ranges, adjusting each parameter range, and calibrating the parameter sets. Each detailed step is described in its own chapter.

Part III: Case Studies

Four case studies with real field conditions are presented in the last part of this handbook. Each case study involves the calibration and validation procedure that was explained in the previous sections. Part III shows how the details of each step are applied to a real project.

In this handbook, only two microscopic simulation models (CORSIM 5.1 and VISSIM 4.1) were used to explain the calibration and validation procedure and those simulation models were selected based on the popularity of the application from the field. However, the same calibration and validation procedure can also be applied to other types of models in the same manner.

Tips on This Handbook

Icons are used in this book to draw attention to certain points. The meaning of each icon is as follows.

Important concept that you should remember.

Flowchart called "Where are we?" is presented at the start of each chapter to give an idea of its position with respect to the complete process.

Part I
Before Beginning
Calibration Procedure

In this part...

General requirements for the simulation model setup are introduced with some examples. Also, alternative calibration data for each simulation model are introduced with recommendations. And finally, guidelines for data collection are provided.

Chapter 1
Getting Ready to Calibrate

In this chapter:
- Why the simulation model setup step is important
- How can we select calibration parameters?
- How can we collect field data that are relevant to the calibration data?

"Garbage In Garbage Out."

This familiar computer axiom also applies to the use of microscopic traffic simulation models. Each simulation model requires many different parameters that are necessary to be defined for each model's car-following and lane-changing characteristics that are embedded in the model. If inappropriate values are entered as a parameter into a simulation model, the resulting output will also be invalid. Therefore, it is essential to have valid input values, such as network and corresponding parameters.

As stated previously, the steps for setting up a base simulation network will not be discussed in this handbook because it is part of the manual for each simulation model.

In this chapter, the data required for the calibration and validation of each simulation model are presented with a brief description of field data collection.

1.1. Getting Ready to Calibrate

In the most general terms, the *"calibration procedure"* is the repeated comparison of numerical field data to the simulation model outputs. For instance, if travel time is selected as an index of comparison, travel time data both from the field and simulation output need to be collected. As a first step, travel time data from the field need to be collected to be used as a target value that simulation model has to produce as an output value. Once target data, field measured travel time data, are recorded and compared to travel time output from the simulation model, the difference
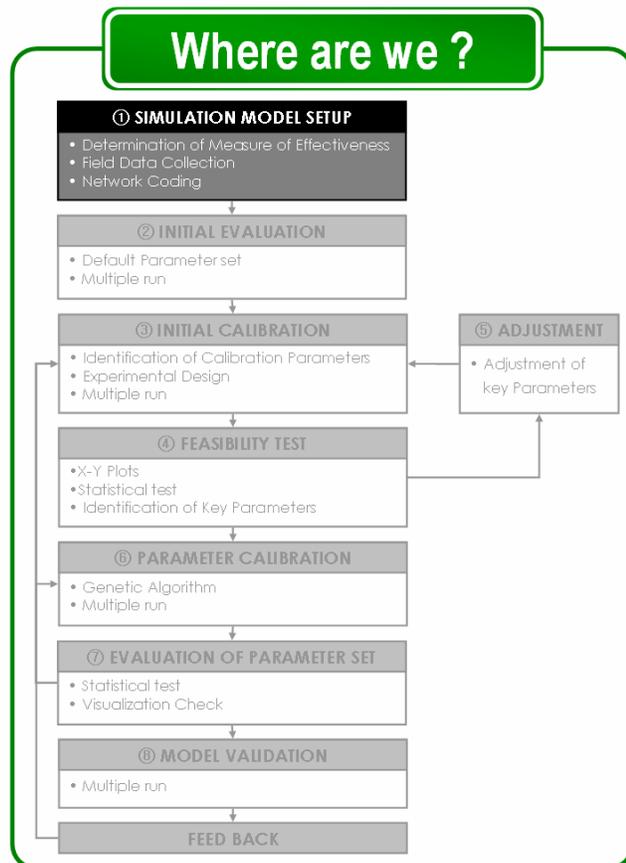
between the two is taken for each run. If that difference is not on a certifiable level, this simulation run is conducted with a fine-tuning process of calibration parameter until the difference is within the pre-determined error range and considered as a best calibration parameter set.

Therefore, it is an essential process to select an index of comparison because this calibration procedure considers a data index as a matching point that simulation model must get at.

1.2. Calibration Data Selection

Before thinking about specific calibration data selection, let's consider the criteria that calibration data should meet. The following three criteria need to be considered:

Calibration data

A measure must be selected to compare with field conditions in order to conduct the calibration process. The measure collected from the field to compare with the simulation result is called "Calibration Data".

- Ease of collection: It is very important to consider the availability of data collection.
- Sensitivity to traffic conditions: Calibration parameters should be sensitive to the field traffic condition because it becomes hard to calibrate if the measure is not sensitive to the traffic conditions.
- Consistency of calculation by different simulation models: Some simulation models have their own logic or a way of obtaining calibration data. So, it is important to make sure that that simulation model calculates the calibration data in a similar manner if you are willing to use multiple simulation models.

1.2.1. Calibration Data from Field

When determining calibration data, in addition to the criteria listed, it is necessary to consider whether or not the data reflect the Level of Service (LOS) of that study area. Therefore, it is also meaningful to consider current Measure of Effectiveness (MOE) for each section type as defined in the HCM. The MOE for each specific type of facility is presented in Table 1.

Despite the recommendations, it is reasonable to select calibration data based on field conditions and other constraints such as geometry, budget and resource limitations. Generally, the following types of calibration data are recommended for use:

- Delay
- Queue Length
- Speed
- Travel Time
- Traffic Counts or Flow Rates

Table 1. Measure of Effectiveness by Facility Type (Highway Capacity Manual, 2000)

Type of Facility	Measure of Effectiveness (MOE)
Basic Freeway Segments	Density (pc/mi/ln)
Ramps	Density (pc/mi/ln)
Ramp Terminals	Density (sec/veh)
Multi-Lane Highways	Density (pc/mi/ln)
Two-Lane Highways	Percent-Time Following (%), Average Travel Speed (mi/hr)
Signalized Intersections	Control Delay per Vehicle (sec/veh)
Unsignalized Intersections	Average Control Delay per Vehicle (sec/veh)
Urban Streets	Average Travel Speed (mi/hr)

1.2.2. Calibration Data from Simulation Runs

Together with measurability of calibration data from the field, it is also important to consider the availability of calibration data with each simulation model. So, we need to know what calibration data results simulation models can provide. Instructions for getting calibration data results in each simulation model and the list of alternate output files are presented in the following section.

① CORSIM

Even though it is comfortable and easy to change the Graphic User's Interface (GUI) tool, the fundamental CORSIM input file consists of simple text. So, the detailed list and the contents of the output values can be easily controlled by editing the CORSIM input file with a text editor. Note that the extension of CORSIM input file is ".trf" and it can be opened by both Traffic Software Integrated System (TSIS) and any kind of text editing program. Figure 3 shows an example of CORSIM input file "CORSIM City.trf" that has been provided by FHWA as an example.

```
Created by TSIS Thu Feb 06 10:11:06 2003 from TNO Version 61
This is a new demonstration test case.   The intent is to show different        0
types of interchanges and different capabilities.  Eventually it should have     0
some of all the capabilities of CORSIM.                                          0
12345678 1 2345678 2 2345678 3 2345678 4 2345678 5 2345678 6 2345678 7 234567
FHWA                                   2    62003FHWA                        1 | 1
        1   0   0  10     7981 0000 21              3 700        7781     7581 | 2
   600                                                                         | 3
                    60                                                         | 4
     0         0   0   0    0   0   0   0   0   0   0                           | 5
    17 136Lake Street                                                          | 10
    29  41Lake Street                                                          | 10
    32  31Lake Street                                                          | 10
    37  39Urban Clover                                                         | 10
    40   1Lake Street                                                          | 10
    43  42Freeway Clov                                                         | 10
    50  49SPUI                                                                 | 10
   109 176Park Way                                                             | 10
   183 184State St                                                            | 10
   183 185State St                                                            | 10
   183 186Main Street                                                         | 10
   183 187Main Street                                                         | 10
   242 243Bus Station                                                         | 10
   246 148Bus Depot                                                           | 10
   253 258Subdivision                                                         | 10
   254 256Residence Wa                                                        | 10
   255 251Subdivision                                                         | 10
   257 252Residence Wa                                                        | 10
     1  40 940        3    01 TT        35        39     20  18  30  0        | 11
```

Column →

← Row →

Figure 3. CORSIM Input Example

As can be seen from the figure, all data in a CORSIM input file are formatted in row and column base and data on each row represent the required data for the different models that are necessary components of a CORSIM model. Each row can be identified with the record type number and it is given at the end of each row—column 78~80, inside the box—with a two digit number. Among many different record types, record type 5 manages the contents and duration of the output file in CORSIM. The specifications on report can be modified by editing the value on that row. It can be used to generate cumulative output and intermediate output.

Record Type

STOP

The input parameters of CORSIM are divided with code called "Record Type". CORSIM consists of 84 record types and each record type controls a specific parameter.

When building a network, record type 5 is automatically inserted into the input file with a default value – all zeros as in the figure. Therefore, it is required to modify record type 5 with the specific need for the calibration and further analysis. Table 2 shows types and details of required input for record type 5.

Once record type 5 is included and edited, it will automatically output a file that contains the data that the user requested including calibration data for each simulation run. Note that the extension of a CORSIM output file is ".out" and it is also possible to open it with any of the text editing programs. A CORSIM output file contains some of the different calibration data in one file and Table 3 shows the list of available calibration data provided by CORSIM.

Table 2. CORSIM Output Record Type (FHWA, 2003)

No.	Start Col.	End Col.	Description	Type	Range	Units	Default
1	1	4	Number of time intervals between cumulative simulation statistics reports	Integer	0~9999	Time Interval	
2	9	12	Time to begin first set of reports	Integer	0~9999	Seconds	0
3	13	16	Duration of first reporting period	Integer	0~9999	Seconds	0
4	17	20	Time between each intermediate set of reports in the first reporting period	Integer	0~9999	Seconds	0
5	21	24	Time to begin the second reporting period	Integer	0~9999	Seconds	0
6	25	28	Duration of the second reporting period	Integer	0~9999	Seconds	0
7	29	32	Time between each intermediate set of reports in the second reporting period	Integer	0~9999	Seconds	0
8	33	36	Time to begin the third set of reports	Integer	0~9999	Seconds	0
9	37	40	Duration of the third reporting period	Integer	0~9999	Seconds	0
10	41	44	Time between each intermediate set of reports in the third reporting period	Integer	0~9999	Seconds	0
11	48	48	Turn-movement-specific output request	Integer	0~1	N/A	0
12	53	58	Name of an output file	Text		N/A	
13	79	80	Record type	Integer	05	N/A	

Table 3. Calibration Data Provided by CORSIM

Delay	▪ Control Delay ▪ Total Delay	▪ Delay Time
Queue Length	▪ Average Queue Length ▪ Queue Length	▪ Maximum Queue Length
Speed	▪ Average Speed	
Travel Time	▪ Travel Time	
Others	▪ Average Occupancy (by detector) ▪ Phase Failure ▪ Vehicle Miles ▪ Densities	▪ Average Volume ▪ Stopped Time and Percentage ▪ Vehicle Trips

② VISSIM

Unlike CORSIM, VISSIM requires the data collection points and the list of data to be determined prior to the actual simulation running. VISSIM has three different types of data collection points that can be installed on the network that are "Data collection points (▯)", "Travel time sections (⏱)" and "Queue counters (⚠)". Each data collection point has its own numbers and can be used to select the point or section to be recorded. Once the previous step is completed, it is required to select the type and other specifications for data collection. The window called "Offline Analysis" —refers to Figure 4 on the right side—manages output and can be activated by selecting "Evaluation—Files" or by typing the shortcut <Alt> + <a> + <f> in order. To obtain outputs from VISSIM, the user needs to check the checkboxes that correspond to the desired output and configure the specific conditions for each output. As mentioned previously, detailed instructions for VISSIM simulation outputs will not be discussed in this handbook. The calibration data given by VISSIM are listed in Table 4.
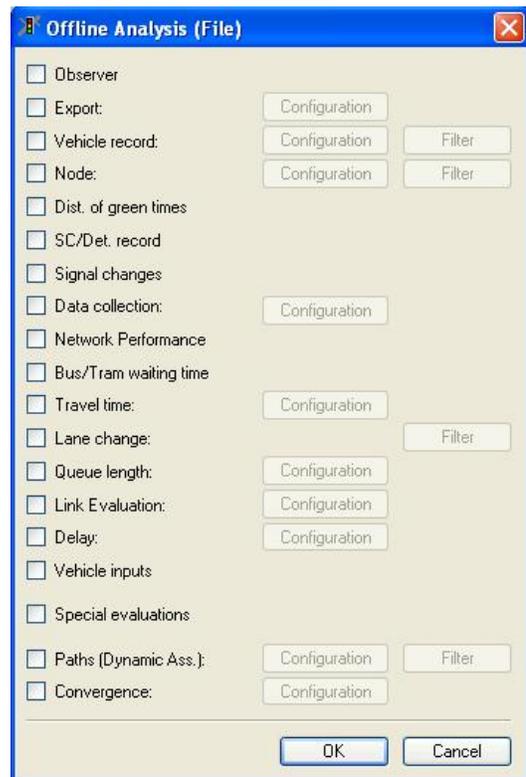
Figure 4. VISSIM Output Control Box

Table 4. Calibration Data Provided by VISSIM

Delay	▪ Stopped Delay		▪ Total Delay	
Queue Length	▪ Average Queue Length		▪ Maximum Queue Length	
	▪ Number of Vehicles stopped in the queue			
Speed	▪ Average Speed		▪ Speed	
Travel Time	▪ Average Travel Times		▪ Travel Time	
Other	▪ Acceleration		▪ Average Volume	
	▪ Bus/Tram Waiting Time		▪ Emissions	
	▪ Fuel Consumption		▪ Number of Stops	
	▪ Occupancy		▪ Vehicle Density	

VISSIM provides two different types of output: raw output and the other is aggregated output. Aggregated outputs are used most of the time due to their convenience. Some representative output file types are listed in Table 5.

Table 5. Representative Output File Types for VISSIM

File type	Calibration Data	File type	Calibration Data
*.vlz	Delay data	*.stz	Queue data
*.fzp	Speed, Occupancy data, etc.	*.rsz	Travel time data

1.3. Field Data Collection

Many different types of field data need to be collected, either from the field or from an available database in order to build a simulation network and calibrate simulation models. In general, necessary data can be categorized into two parts: fundamental data" and "calibration data". First, fundamental data are the data required to build a base simulation network. In order to build a network for the simulation model, the following types of data are required to be collected from the field:

- Geometry data
  — Number of lanes
  — Lane width
  — Distance
- Traffic data

— Volume
- Control data
  — Traffic signal setting
  — Priority rule
  — Stop sign

Second, calibration and validation data are necessary to be used as a measure of the calibration and validation procedure. For more accurate calibration, multiple data sets (either different data collection date or types) are desired to be collected and used to consider day-to-day variability and differences in different data types. Also, validation data are required to be obtained from the field, and have to be different from the data used for the calibration procedure.

Likewise, the accuracy of the data collected from the field is very important because if fundamental data and calibration data are not accurate, then the calibration procedure is nothing but a waste of resources. Thus, it is essential to check whether or not the traffic flow is affected by any of the following conditions:

- Bad weather
- Construction
- Incidents or crashes
- Special events

If the model is to be used for a specific condition, such as a work-zone, bad weather conditions, or special events, it is desirable to collect all the data on a day when similar traffic conditions exist in the field.

Part II
Starting the
Calibration Process

In this part
The main calibration steps are introduced and explained with examples.
This part consists of five chapters, and each chapter talks about the following subjects:
- Chapter 2: Feasibility of default parameter set
- Chapter 3: Range check of each calibration parameter
- Chapter 4: Adjustment of parameter range
- Chapter 5: Finding optimized parameter set with selected ranges
- Chapter 6: Evaluating and validating calibrated parameter set

Chapter 2
Trial with Default Parameter Set

In this chapter:
- Conducting multiple runs with default parameter set
- Using histograms for decision-making
- Determining feasibility of a default parameter set

Every simulation model has an uncalibrated parameter set inherent to the simulation model that we call the default parameter set. It may work for some different cases but it is impossible to guarantee. However, if the default parameter is acceptable to be used for that specific network, it is fine to skip the calibration and validation procedure and use the simulation model with the default parameter set for further analysis. This is beneficial because a huge amount of time and effort can be saved by skipping the main calibration procedure. To ensure that kind of possibility, the applicability of the default parameter set needs to be tested after the simulation model setup process. If the default parameter set is not acceptable, calibration and validation procedures—refer to chapters 3 through 7—need to be conducted.

The feasibility test consists of two steps: multiple runs of the simulation model with the default parameter set and subsequent comparison to calibration data collected from the field. Figure 5 shows the inputs and outputs of the process.
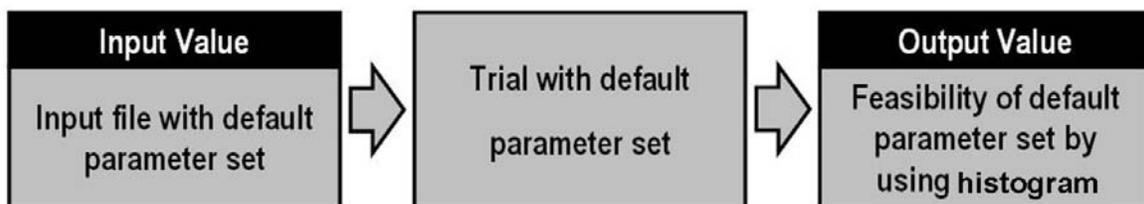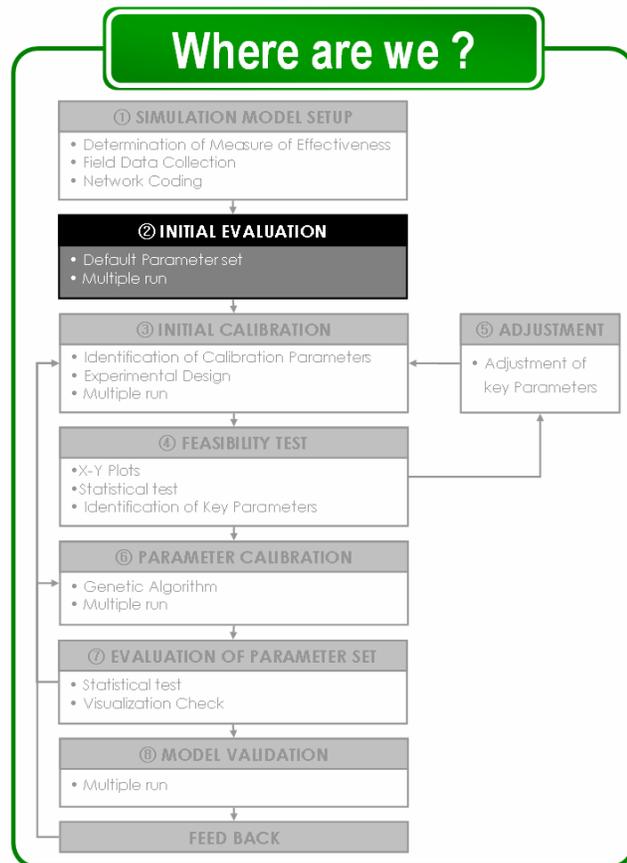
Figure 5. Inputs and Outputs of Default Parameter Set Evaluation

2.1. Multiple Runs with Default Parameter Set

2.1.1. Why Do We Have to Run Simulation Models Multiple Times?

Generally, simulation models give slightly different output values for repeated simulation runs. This is due to a randomly generated seed number for each simulation run. The randomly generated seed number is used to make decisions for the simulation, such as the timing of vehicle loading, the type of vehicle that will be loaded, and the path for each vehicle. Thus, if we aggregate the output data from multiple runs with a given network, they will have a certain distribution with minimum and maximum values. Figure 6 shows an example of the distribution of travel time output from 100 runs and clearly shows a distribution rather than specific value.
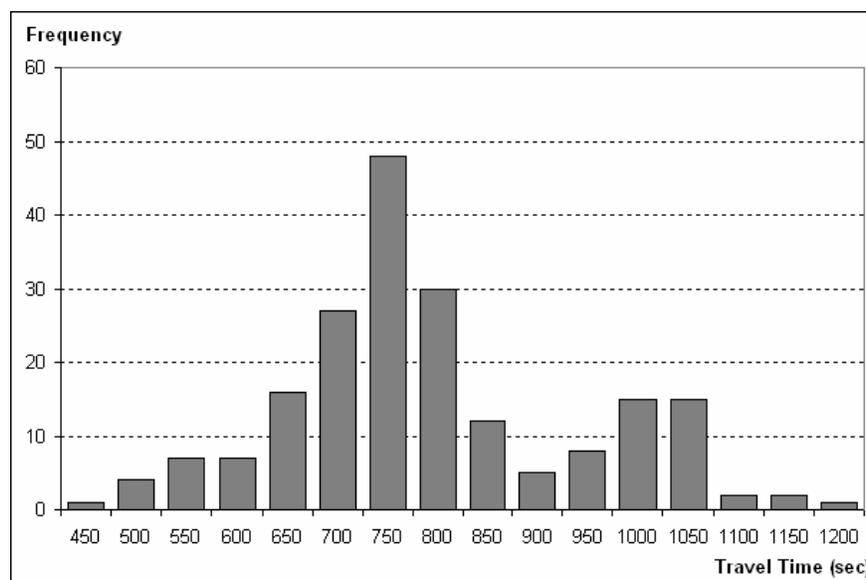
Figure 6. Example of Travel Time Output of Multiple Runs

The data from numerous repetitions of the simulation model can be classified within a certain interval. In the case shown on the previous page, the intervals are 50 seconds apart. The corresponding histogram is used to find the distribution of simulation outputs.

The X-axis represents the intervals of calibration data that you selected in the simulation model setup part (refer to Chapter 1) and the Y-axis represents the frequency of occurrence of that specific interval. Therefore, the X-axis on the graph on the previous page represents travel time and the Y-axis represents the frequency of simulation run results for each travel time class. As the graph shows, the result of simulation model repetition is a distribution of values, not one specific value. The more times the simulation model is run—the larger the sample size—the more representative the distribution will be.

2.1.2. How Many Times Do We Have to Repeat?

Since it is clear that the repetitions of the simulation model are important, the next question would be "How many times do I have to run?" A statistical process needs to be conducted to get a theoretical idea of the reasonable number of repetitions. Without this confirmation process, you may not be able to conclude that you covered all possible cases. The process for getting the minimum number of repetitions is taken from the Federal Highway Administration's (FHWA) "Guidelines for Applying Micro-simulation Modeling Software (2004)." The complete process has four main steps, as shown in Figure 7.

① **Executing few simulation model repetitions**

In order to conduct the following steps, sample data—calibration data from simulation model repetitions—must be collected. In other words, we might need to run more than 100 times in some cases to cover the whole distribution. However, we can predict the possible variation of output results by analyzing the calibration data from several preceding repetitions and try to find an acceptable number of runs in advance. Normally, four runs are considered to be the minimum number of preceding repetitions.

Figure 7. Sample Size Selection Process

② **Estimating sample standard deviation**

With the sample calibration data from ①, estimate the sample standard deviation. Sample standard deviation is calculated with the following equation:

$$S^2 = \frac{\sum (x - \bar{x})^2}{N - 1}$$

In this equation, x represents output value for each repetition and $\bar{x}$ represents an average value of all repetitions. The difference $(x - \bar{x})$ tells how far away from the mean each output value is, which is a measure of variance. This term is squared to make it positive. Essentially, the variance of each individual data point is summed and divided by $(N - 1)$, which is the number of

repetitions minus 1, to give an overall "average" variance. Why do we need to calculate sample variation and not standard deviation? Standard deviation, S, is just the square root of the variance, S^2.

Standard Deviation

STOP

Standard deviation is a measure of the variability of a series of observations. It answers the following question: "By how much is a set of data different from an average or fitted value?"

For instance, if we ran 4 times and we got 50, 53, 48 and 52 for each run, we can apply these values to the previous equation and find,

$$S^2 = \frac{(50-50.8)^2 + (53-50.8)^2 + (48-50.8)^2 + (52-50.8)^2}{4-1} = \frac{14.76}{3} = 4.92$$

Therefore, $S = \sqrt{4.92} = 2.218$

③ Selecting Confidence Level

The next step is to decide how accurate you want your estimation to be. This is called a confidence level. A 95% confidence interval is widely used and the concept is shown in Figure 8.

2.5%   97.5%
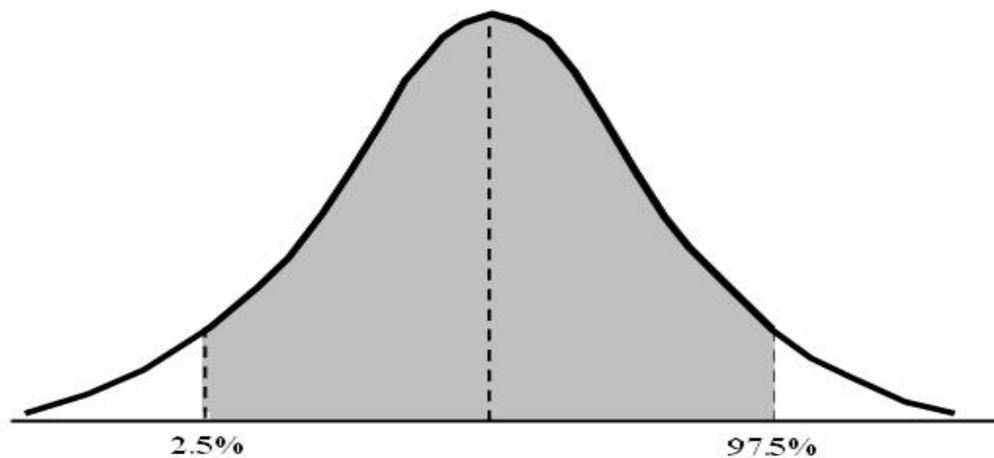
Figure 8. Confidence Level Graph

The shaded area contains 95% of the area under the curve, the curve representing a normal distribution of possible outcomes. You can be 95% confident that for any given repetition the measured value will be within this range. As the confidence level increases, more area under the curve becomes shaded. Similarly, as the confidence level increases, either the coverage rate or the number of replications will increase.

④ Calculating Minimum Repetitions

At this point, you are ready to calculate the number of minimum repetitions. The required minimum can be calculated by using the following equation:

$$C = 2 \times t_{(1-\alpha/2),N-1} \frac{s}{\sqrt{N}} \qquad \text{where,} \ \ C = 1 - Confidence \ Level$$

For example, if your confidence level is 95%, C equals 0.05. $t_{(1-\alpha/2),N-1}$ is a t-statistic value for the probability of a two-sided error summing to alpha with N-1 degrees of freedom. This value can be found in the Appendix and α is the confidence level that you selected in the previous step. Also, the standard deviation that you got from step ② needs to be used, noted as s.

Frequently used sets are presented in Table 6. In order to use this table, select a confidence level and calculate C/S (ratio of confidence level to standard deviation).

Table 6. Minimum Number of Replications with Desired Confidence Interval

$C_{1-\alpha}/S$	Selected confidence level	Minimum number of repetitions
0.5	99%	130
	95%	83
	90%	64
1.0	99%	36
	95%	23
	90%	18
1.5	99%	18
	95%	12
	90%	9
2.0	99%	12
	95%	8
	90%	6

It is clear that statistical knowledge should be applied to the number of repetitions decision step; however, an engineer's judgment is also frequently applied to determine the number of repetitions.

2.2. Is the Default Parameter Set Valid?

Validity of the default parameter set can also be verified by using graphical methods rather than analyzing in a statistical manner, and one could find it more straightforward. A histogram analysis and X-Y plot analysis are used to check the validity of the calibration result by using single and multiple performance measures.

2.2.1. Histogram Analysis

A histogram is a graphic presentation of the frequency distribution of a discrete variable. Each axis represents the frequency and the interval respectively and the vertical length of each

rectangle shows the frequency of that specific interval. In order to draw the histogram, all the data need to be arranged in a frequency table form. Table 7 shows a simple example of a speed data frequency table.

Table 7. Example of Frequency Table

Interval	Up to 30	30~35	35~40	40~45	45~50	50~55	55~60	Above 65	Total
Number of Vehicles	130	167	209	278	328	516	374	193	2,195

The frequency table indicates the number of data points that fit in that specific class as we discussed earlier. For example, we can find out that 2,195 speeds were collected and each value in the second row corresponds to the class that was defined in the first row. Since data have been converted to category (discrete) format, it is possible to draw a histogram that is a graphical expression of the frequency data. By drawing a histogram, it is possible to draw the distribution of collected data frequency. Figure 9 shows an example of a histogram that has been drawn with the speed data provided.
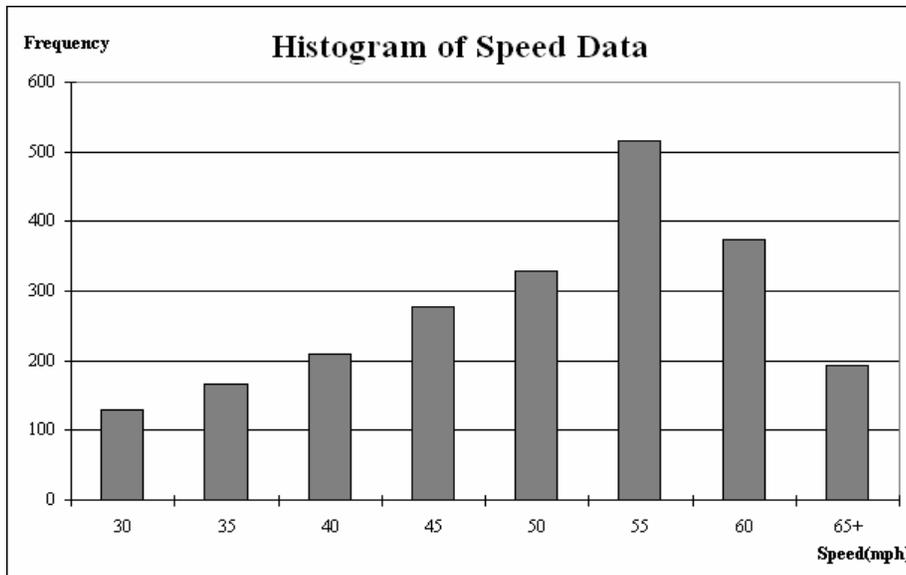
Figure 9. Example of Histogram

Each bar shows each class and the length of each bar varies with the frequency of each specific class. For more information on the histogram, refer to available statistics books. Example of a travel time output histogram is shown once again in Figure 10.
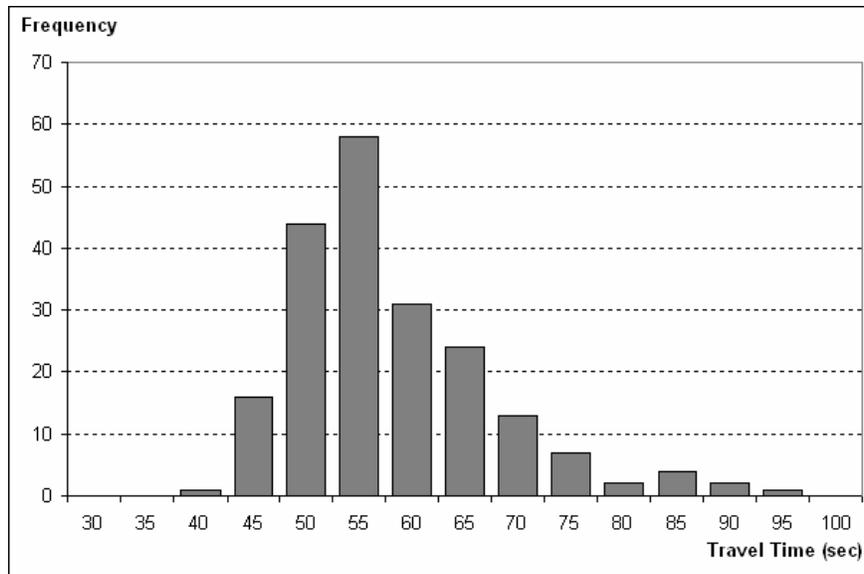
Figure 10. Travel Time Output Example

When is it feasible?

To examine the feasibility of the default model by using histogram analysis, it is necessary to have both field collected data value or ranges and the distribution of simulation output data that correspond to the selected performance measure drawn in histogram format. For easy understanding, explanations are presented with examples here.

Let's assume that we drew a histogram with the travel time data from repetitions of a certain simulation model. At the same time, we collected four travel time data sets from the field. The field collected travel time data values are 55, 67, 72 and 59 seconds. As shown in Figure 11, if we combine the field travel time data with the histogram, it is easy to sense that all four field data values fit inside the distribution shown by the histogram. Therefore, in this case we conclude that the default parameter set is feasible because all collected travel time data from the field could be expressed by the current parameter setting.
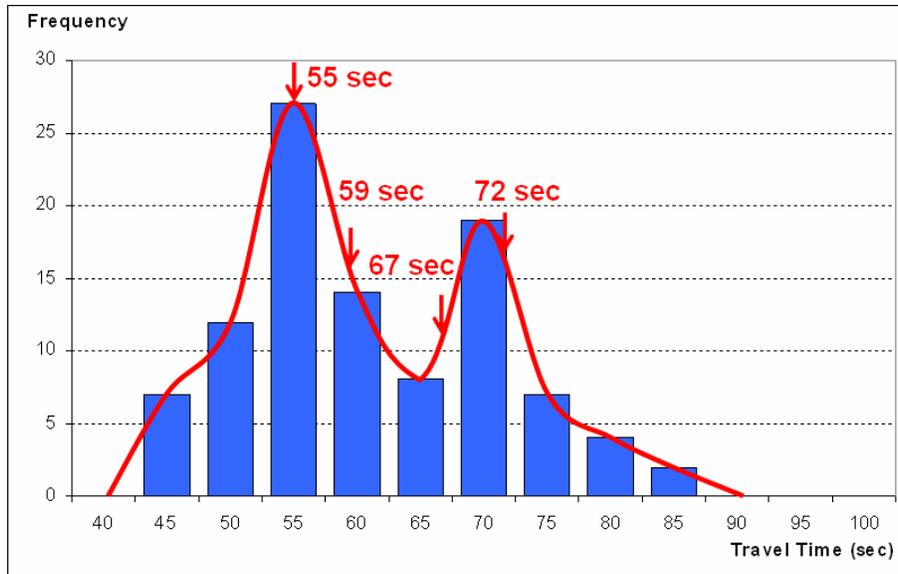
Figure 11. Example of Acceptable Case—Histogram

On the other hand, if the collected travel time data is 79, 83, 94, and 102 seconds, we conclude that the default parameter set is not feasible with the same criteria we used previously. The first two travel time data points fit within the distribution, so they can be expressed by this parameter set. However, the last two travel times do not fit inside the distribution, so we can say that the last travel times cannot be expressed by the distribution. Therefore, we conclude that the default parameter set is not feasible. Figure 12 shows an example of an unacceptable case.
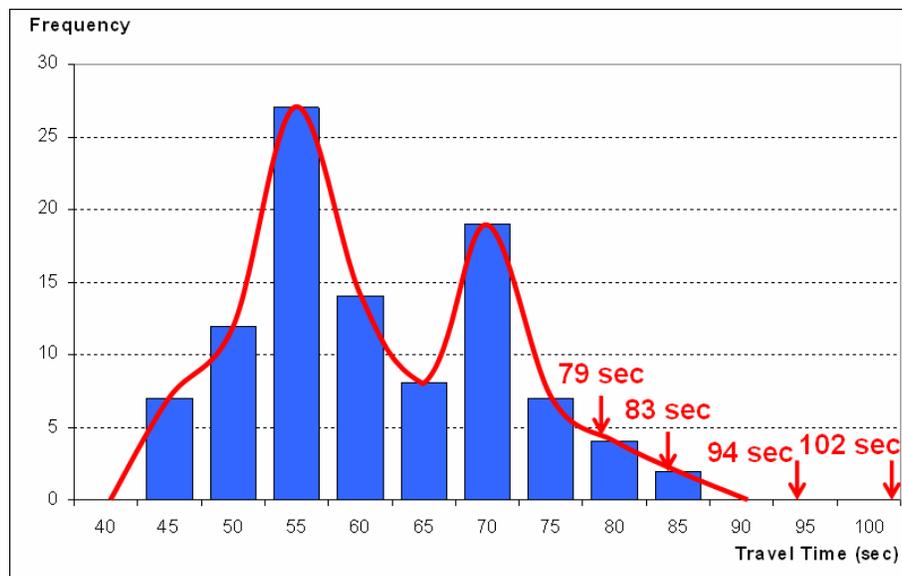
Figure 12. Example of Unacceptable Case—Histogram

23

2.2.2. X-Y Plot Analysis

An X-Y plot is a two-dimensional graph representing a set of bivariate data. In order words, unlike a histogram, the X-Y plot uses a level or location of each data rather than the frequency of each interval's representative value. As a result, each data point (dot) is located at the intersecting point of two values that correspond to two different performance measures. Figure 13 and Table 8 show an example of the X-Y plot and performance measure data combination.
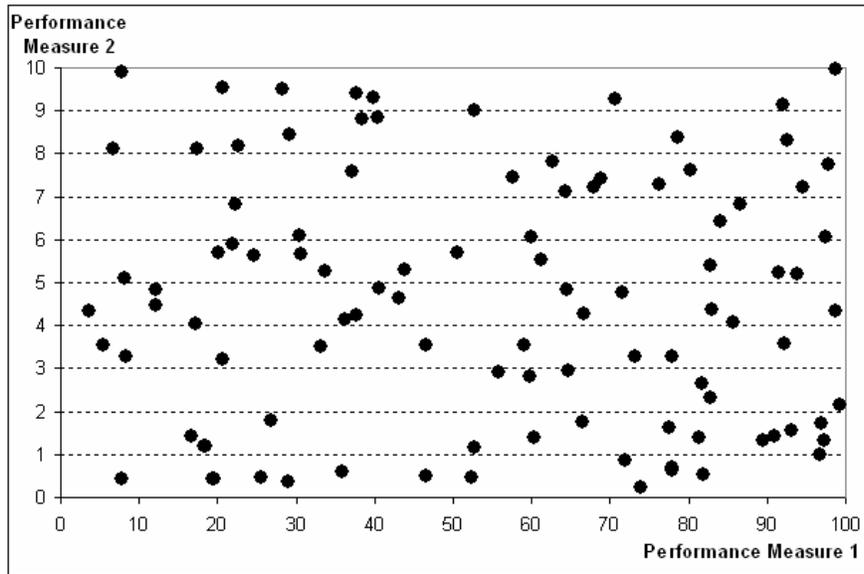
Figure 13. Example of X-Y Plot

Table 8. Example of Performance Measure Data Combination

P1*	P2*	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2
34	5	78	1	97	1	38	9	63	8	5	4	64	7	37	8	71	9	65	3
92	5	12	5	8	0	53	1	8	3	60	1	73	3	47	0	97	1	8	5
82	1	90	1	95	7	29	0	56	3	12	4	80	8	40	9	83	5	98	8
33	4	25	6	8	10	36	4	65	5	21	3	81	1	79	8	87	7	67	4
44	5	99	2	91	1	27	2	67	2	22	7	28	10	61	6	18	1	22	6
92	4	82	3	78	2	78	1	93	2	46	4	60	6	38	9	60	3	36	1
83	4	99	4	72	1	97	2	17	4	93	8	92	9	18	8	23	8	74	0
21	10	40	9	78	3	20	6	58	7	31	6	59	4	68	7	29	8	4	4
30	6	26	0	84	6	86	4	38	4	69	7	7	8	97	6	72	5	94	5
52	0	99	10	19	0	83	2	43	5	17	1	76	7	51	6	53	9	41	5

*Note: Letter "P" from "P1" and "P2" represents "Performance Measure".

As shown in Figure 11, each point is located at the intersecting point of two performance measures' intersecting point. After drawing graphs such as histograms or X-Y plots, it is

necessary to determine the feasibility of the model with a default parameter set. As a first step of the feasibility check, the cases that determined to be feasible were described in this part with examples. The criteria that need to be used for the X-Y plot analysis are presented in the following part.

When is it feasible?

As was done for the histogram analysis, it is necessary to have both field-collected data values or ranges and the distribution of simulation output data that correspond to the selected performance measure drawn in an X-Y plot format. A quick example by using the data set in Table 8 is presented here.

Let's say that we have two performance measures that have been obtained from the simulation model. And identical types of data sets have been collected from the field for multiple days to be used as comparison with the simulation outputs. Assume that the ranges of field-collected data for performance measures 1 and 2 are 62 to 86 and 8.3 to 15.2 as presented in Figure 14 within the dark-shaded box. That range should be overlapped with the X-Y plot of simulation outputs and the 90% confidence interval region of the total data point cluster, shown as a light-shaded box, should be drawn as well. If the 90% confidence interval region falls on the field-collected performance measures region, it can be considered to be feasible. However, if those two regions are not overlapping at all, it cannot be considered to be feasible and a calibration procedure needs to be conducted.
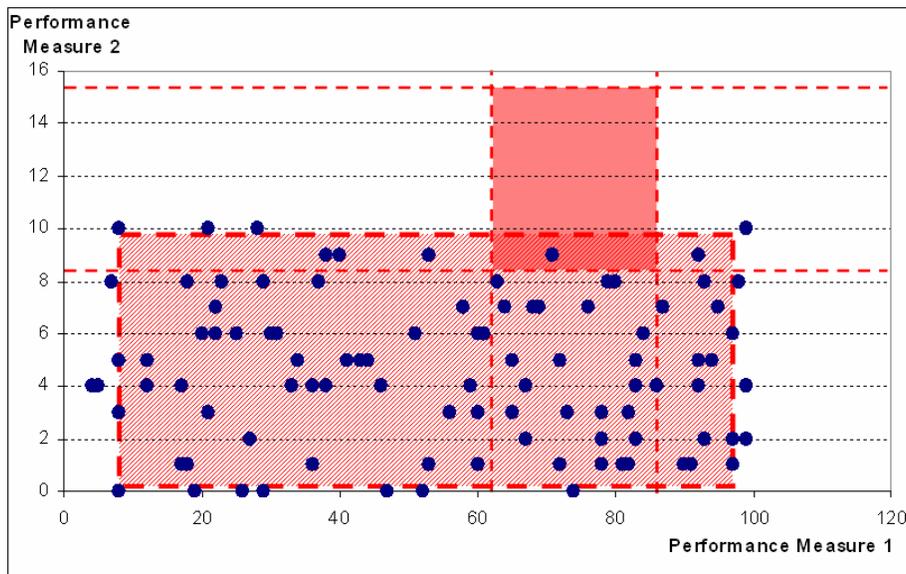
Figure 14. Example of Acceptable Case—X-Y plot

Chapter 3
Initial Calibration

In this chapter:
- Selection of parameters to be calibrated
- Sampling different cases within determined range
- Verification of the result

Let's start to talk about the main calibration procedure; our main interest in this investigation. The main idea of parameter calibration can be summarized as follows.

①. Among many different calibration parameters for each simulation model, it is necessary to select parameters that will affect the result.

②. Set up a range for each selected parameter expected to be appropriate based on your experience or available traffic data.

③. Check whether the ranges for each parameter are reasonable based on multiple runs of the simulation model.

④. If the range turns out to be unreasonable, try to figure out which parameter affected this biased result more by using plotting, statistical testing, and/or graphical methods.

And then conduct step ③ again.

⑤. Once ranges for each parameter satisfy the conditions in step ③, continue on with the calibration procedure. An optimization tool can help find an optimized parameter set that satisfies certain criteria within the range constraints.

⑥. If you obtained the result from step ⑤, this is the end of the calibration process. Once the calibrated parameter set has been found, the validation procedure needs to be conducted.

In this chapter, we will talk about the details that were briefly explained in steps ① to ③, which mainly consider settings and verifying parameter ranges. So this chapter includes the following sections.
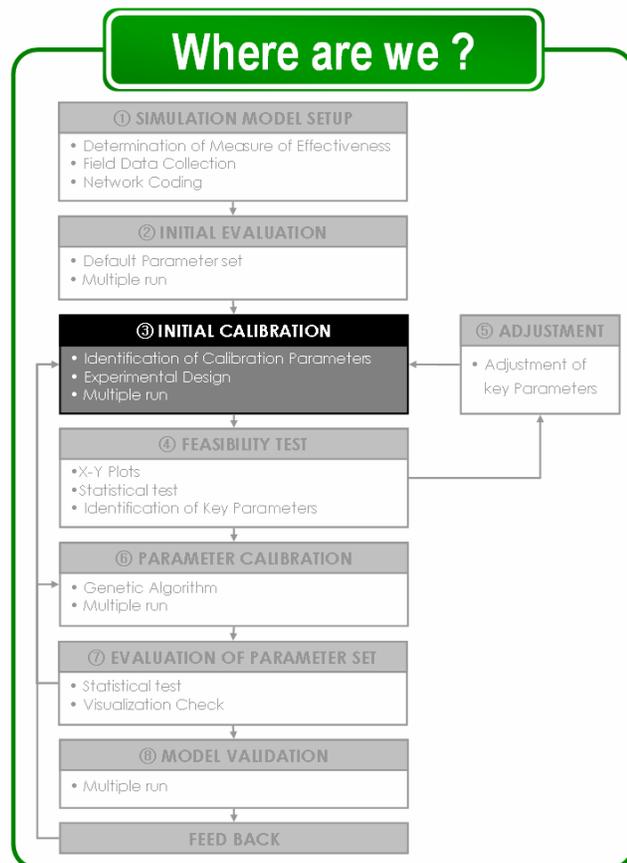
- Identification of calibration parameters
- Sampling process of a parameter set
- Multiple runs with sampled parameter sets

3.1. Identification of Calibration Parameters

There are several inputs that are known as calibration parameters for each microscopic simulation model. These inputs allow users to fine-tune simulation models to match real traffic conditions, generally consisting of a driver's behavior parameters and vehicle performance parameters. In this section, the following contents for each simulation model are introduced and explained.
- The list of calibration parameters of each simulation model
- Acceptable ranges of each parameter in applicable case

Further information, such as explanations on each parameter, will not be discussed in this handbook. Refer to the user manual of each model for that information.

3.1.1. CORSIM

In this section, helpful information on different calibration parameters of CORSIM is discussed. CORSIM provides two types of calibration parameters that deal with driver's behavior and vehicle performance. So, let's first look at the list of calibration parameters that CORSIM provides. After reviewing this section, you will have an idea about the calibration parameter selection for your project.

Tables 9, 10, and 11 show the list of calibration parameters for CORSIM. They also provide the record types (see chapter 1 for details) and acceptable ranges and default values for each parameter.

Table 9. Calibration Parameter List 1—CORSIM

Parameter			Range	Default
Description	**Record Type**	**Columns**		
Mean Value of Start-up lost time	11	57~60	0~99(0.1sec)	20
Mean queue discharge headway		61~64	14~99(0.1sec)	18
Desired free-flow speed		65~68	0, 10~65(mi/h)	30
Mean Start-up delay	20	19~20	0~60(0.1sec)	10
Desired free-flow speed		21~22	0~70(mi/h)	65
Car following sensitivity multiplier		69~72	1~1000(%)	100
New car-following sensitivity factor (each driver type)	68	1~40	0~9999(0.01s)	N/A
New value for Pitt car following constant		43~44	3~10(ft)	10
Time to complete a lane change maneuver	70	1~4	0~9999(0.1sec)	20
Mandatory lane change gap acceptance parameter		12	1~6	3
% of drivers desiring to yield right-of-way to lane changing vehicles attempting to merge		13~16	0~100(%)	20
Multiplier for desire to make a discretionary lane change		19~20	1~10(0.1unit)	5
Advantage threshold for discretionary lane change		23~24	1~10(0.1unit)	4
Maximum non-emergency freeway deceleration (each vehicle type)		25~60	$0 \sim 9999$ (0.1 ft/sec^2)	80
Leader's maximum deceleration perceived by follower		65~68	$5\sim20 \text{ (ft/sec}^2)$	15
Minimum Deceleration for lane change	81	13~16	$1\sim10(\text{ft/sec}^2)$	5
Difference in max. and min. acceptable deceleration for mandatory lane change		17~20	$5\sim15(\text{ft/sec}^2)$	10
Difference in max. and min. acceptable deceleration for discretionary lane change		21~24	$5\sim15(\text{ft/sec}^2)$	5
Deceleration rate of leading vehicle		25~28	$10\sim15(\text{ft/sec}^2)$	12
Deceleration rate of following vehicle		29~32	$10\sim15(\text{ft/sec}^2)$	12
Driver factor used to compute driver aggressiveness		33~36	15~50	25
% of drivers who cooperative with lane changer		45~48	10~100(%)	50
Headway below which all drivers will attempt to change lanes		49~52	1~30(0.1sec)	20
Headway above which no drivers will attempt to change lanes		53~56	30~100(0.1sec)	50
Mean longitudinal distance over which drivers decide to perform one lane change		57~60	50~2500(ft)	300

Table 10. Calibration Parameter List 2—CORSIM

Parameter			Range	Default
Description	**Record Type**	**Columns**		
Probability of Jumping the number of lanes (each entry)	140	5~56	0~100(%)	38
Left-Turn speed		57~60	0~44(ft/sec)	22
Right-Turn speed		61~64	0~26(ft/sec)	13
Spillback probability (each vehicle type)	141	1~16	0~100(%)	N/A
Left-Turn lagging for 2 seconds		17~20	0~100(%)	N/A
Left-Turn lagging for 2~4 seconds		21~24	0~100(%)	N/A
Left-Turn lagging for 4~5 seconds		25~28	0~100(%)	N/A
Acceptable Gap (each driver type)	142	1~40	15~75(0.1sec)	N/A
Additional Gap Time for crossing 1~10 lanes	143	1~40	10~75(0.1sec)	N/A
Acceptable Deceleration (for each driver type)	144	1~40	$2{\sim}30(\text{ft/sec}^2)$	N/A
Acceptable Gap (for each driver type)	145	5~44	10~100(0.1sec)	N/A
Free-Flow Speed Adjustment (for each driver type)	147	1~40	0~1000(%)	N/A
Start-up Lost-time distribution percentage (for each driver type)	149	4~48	0~1000(%)	N/A
Vehicle Queue Discharge distribution percentage (for each driver type)			0~1000(%)	N/A

The brief explanations on each record type can be found in the user manual. Thus, refer to CORSIM user's guide for more detailed information.

3.1.2. VISSIM

In this section, calibration data information on VISSIM is provided. VISSIM provides a few more calibration parameter categories, which are car following, lane changing, priority rule, and desired speed distribution parameters. Table 10 shows the list of calibration parameters. Note that VISSIM parameters do not have specific restrictions on range.

Table 11. Calibration Available Parameter List 1—VISSIM

No.	Parameter	
1	*Simulation Resolution*	
2	Car Following Parameter	Maximum Look Ahead Distance
3		Average Standstill Distance
4		Additive Part of Safety Distance
5		Multiple Part of Safety Distance
6		CC0, Average Standstill Distance
7		CC1, Headway Time
8		CC2, "Following" Variation
9		CC3, Threshold for Entering "Following"
10		CC4, Negative "Following" Threshold
11		CC5, Positive "Following" Threshold
12		CC6, Speed Dependency of Oscillation
13		CC7, Oscillation Acceleration
14		CC8, Standstill Acceleration
15		CC9, Acceleration at 80km/h
16	Lane Change Parameter	Maximum Deceleration
17		Reduction Rate (-1m/s^2 per Distance)
18		Accepted Deceleration
19		Waiting Time Before Diffusion
20		Minimum Headway
21	Desired Speed Distribution	
22	Priority Rule	Minimum Gap Time
23		Minimum Headway

* Note: The range of each parameter was not presented because no specific ranges vary with different cases.

3.2. Sampling Possible Parameter Sets

The number of combinations for many parameters are enormous; moreover, many of the parameters are continuous values rather than discrete. So, in fact, it is unattainable to examine all possible combinations of a parameter set. For instance, if 10 parameters need to be tested and each discrete parameter has 5 levels, $5^{10} = 9,765,625$ combinations need to be tried (this could take years). It would be impossible to analyze all these combinations; however sampling from these combinations must be conducted and equivalently from the entire range of possibilities.

3.2.1. Latin Hypercube Design

A Latin Hypercube Design (LHD) algorithm is used to reduce the number of combinations into a reasonable level, while still reasonably covering the entire parameter surface. It is part of a space-filling method that tries to maximally cover space. If we divide the sample space with 20% intervals, the same number of samples will be selected from each interval. Figure 15 represents this case with a distribution of samples following a normal distribution and shows uniform distribution.
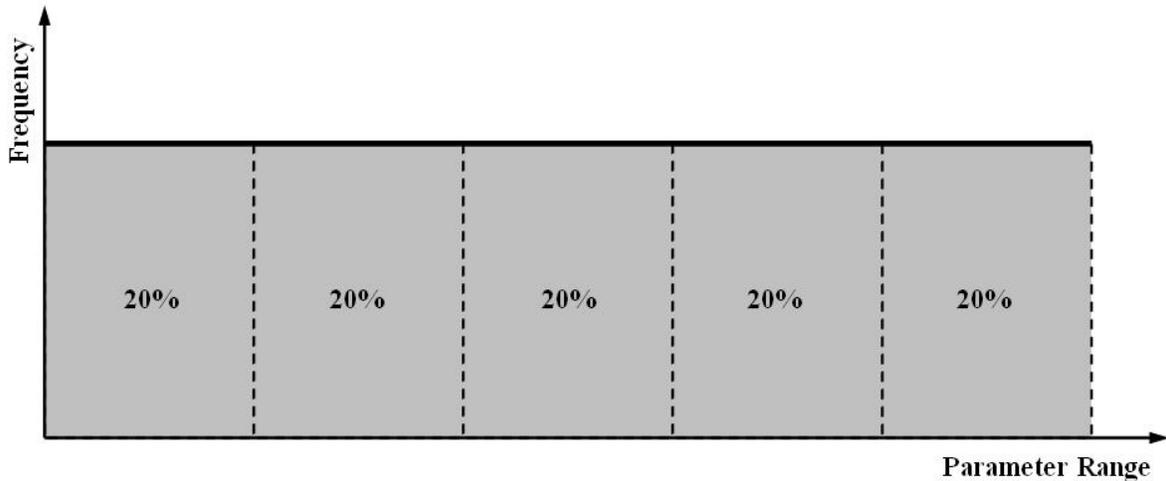
Figure 15. Sampling Region Concept

For the purpose of calibration, 200 different parameter sets seem to be a minimum number that should be conducted at least. More sets will guarantee more accurate results; however when it is impossible to conduct sufficient parameter sets, it is required to keep that minimum criteria.

3.2.2. Multiple Runs with Designed Parameter Set

As mentioned in a previous step, simulation models with specific parameter sets need to be run multiple times to consider the effect of randomly generated seeds (i.e. variability of real world day-to-day conditions). In this case, 5 or perhaps fewer runs are sufficient because of the large number of runs generated, (around 10,000 runs) and this requires a large amount of time and effort.

3.2.3. What Is the Acceptable Range?

After the sampling and multiple runs steps, it is necessary to conduct a step that is verifying the feasibility of determined calibration parameter ranges. Identical steps that have been used for the test of a default parameter set are used for this part for verification purposes. Details on the feasibility test step of both single and multiple performance measures are presented in the following part.

Single Performance Measure Case

The result of multiple simulation runs can be presented in a histogram format. Once the distribution has been obtained, it is required to check whether the distribution includes the field measured values or not. The concept of parameter range acceptance is shown in Figure 17. If the field-measured value falls within the middle 90% of the distribution, it can be considered as acceptable. As shown in the figure, the region with blue circles is an acceptable region and the region with X marks is an unacceptable region.

31
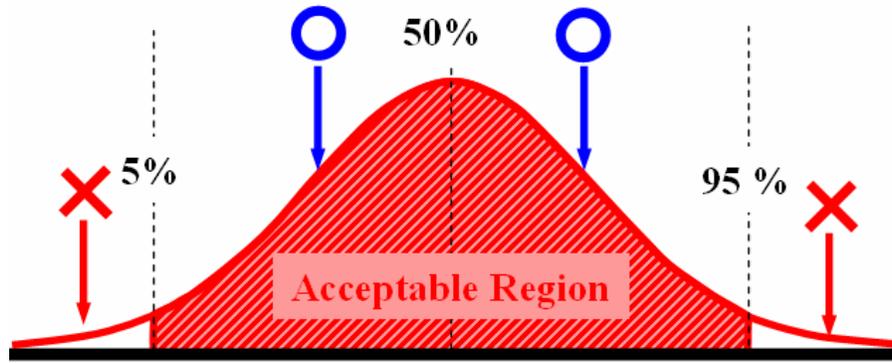
Figure 16. Concept of Parameter Range Acceptance—Single Performance Measure

Multiple Performance Measures Case

When multiple performance measures are used, it is necessary to consider the combinations of those multiple performance measures. Therefore, the X-Y plot analysis was used in the feasibility test for determined calibration parameter ranges. As shown in Figure 17, the field-collected data ranges for two performance measures are presented on the X-Y plot as a region (shaded box). If the 90% confidence interval region overlaps this light-shaded box with mark O (region), it can be considered as a valid calibration parameter range. If this is the case, then the parameter ranges adjustment step described in the next chapter can be skipped and move straight to the GA optimization step. However, if the 90% confidence interval region of simulation output does not fall within the light-shaded box but in the dark-shaded area with mark X, then the parameter ranges adjustment described in the next chapter should be conducted to determine a new set of parameter ranges.
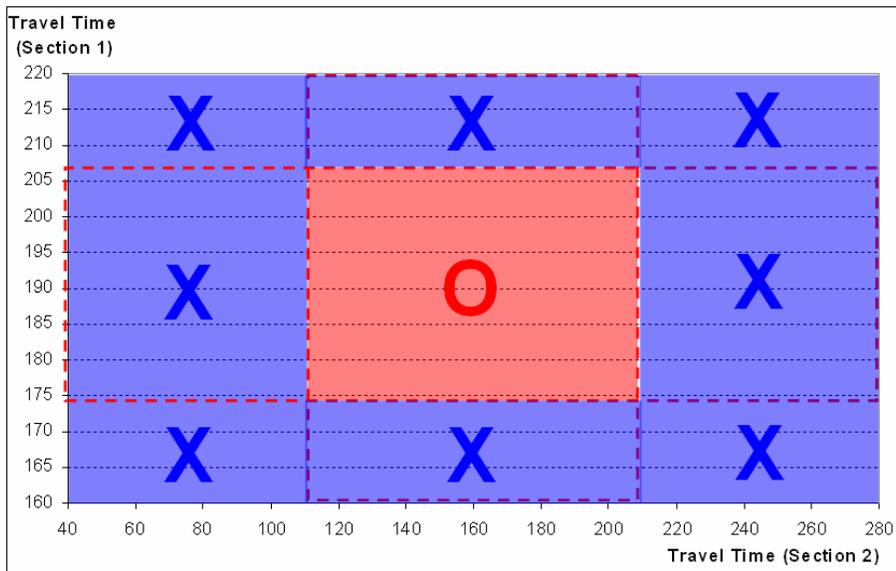
Figure 17. Concept of Parameter Range Acceptance—Multiple Performance Measures

3.2.4. Example

At this point, let's refresh ourselves on the processes that have been explained in this chapter by going through an example. Table 12 shows 8 parameters in VISSIM and each parameter's range. In this example, 8 key parameters in VISSIM are selected as calibration parameters and the ranges of each parameter are defined as shown in Table 12. The ranges of each parameter are defined based on the default parameter of a value described in the VISSIM manual and ranges to have a symmetric range for both sides.

Once the ranges are determined, 20 parameter sets are generated by using LHD. The 20 parameter sets within the parameter ranges provided are shown in Table 13.

Table 12. Parameter Selection Example

	Description	Range		Description	Range
P1	Simulation Resolution	1~3	P5	Additive Part of Desired Safety Distance	1.00~5.00
P2	Number of Observed Preceding Vehicles	1~4	P6	Multiple Part of Desired Safety Distance	1.00~6.00
P3	Maximum Look Ahead Distance	200~300	P7	Minimum Gap	3~6
P4	Average Standstill Distance	1.00~5.00	P8	Minimum Headway	5~20

Table 13. Example of Generated Sample Space by Latin Hypercube Design Method

Scenario Number	P1	P2	P3	P4	P5	P6	P7	P8
1	2	1	208.35	3.28	1.77	4.81	4.67	11.30
2	1	4	253.59	2.34	3.20	4.95	5.80	6.78
3	2	2	224.33	1.47	3.58	1.66	3.61	10.87
4	2	2	217.78	2.86	4.13	3.68	4.82	17.66
5	2	3	242.67	1.02	1.56	3.52	3.32	8.20
6	3	2	263.84	1.19	3.44	1.37	3.72	5.18
7	1	1	294.83	2.47	2.43	5.02	5.34	12.84
8	1	4	231.24	2.17	4.73	3.50	5.15	11.99
9	1	4	222.85	3.75	4.05	5.19	5.25	12.24
10	2	4	270.65	4.40	1.97	1.74	5.10	16.74
11	3	3	282.48	4.96	4.97	3.30	3.01	7.29
12	2	2	244.79	2.31	4.83	2.48	5.46	14.62
13	2	3	287.41	1.37	4.45	1.06	4.06	17.45
14	2	3	261.13	3.63	1.05	1.94	4.29	19.43
15	2	1	258.77	3.49	2.88	2.01	4.18	19.90
16	3	3	235.00	4.66	3.74	3.91	3.92	9.04
17	1	2	291.44	1.85	1.34	4.58	3.85	14.02
18	3	2	299.19	1.97	2.26	3.07	5.55	10.50
19	1	3	213.15	4.09	2.85	5.70	3.22	7.70
20	3	4	237.91	3.13	2.54	1.20	4.76	13.24

Correlation of each parameter pair is minimized to test the effect of individual parameters. As shown in Table 14, correlation values of each pair are low. Thus, the surface of the parameter is considered to be adequately covered.

Table 14. Example of Correlation values of Sample Space Generated by Latin Hypercube Design Method

	P2	P3	P4	P5	P6	P7	P8
P1	-0.089	0.072	0.110	-0.025	-0.091	-0.079	-0.123
P2		-0.054	0.079	0.069	-0.087	0.025	-0.067
P3			0.000	-0.070	-0.071	-0.095	0.022
P4				0.000	0.062	-0.082	-0.047
P5					-0.026	-0.024	-0.091
P6						0.084	-0.097
P7							0.099

Chapter 4
Working with Calibration Parameters

In this Chapter:
- Identify key calibration parameters
- Adjust key parameter ranges

4.1. Feasibility Test

In a previous step, it was possible to confirm whether the range of a parameter set was acceptable or not. If the parameter set was acceptable, it is adequate to skip the process in this chapter and move on to the next part, which is Chapter 5. However, if the ranges were not acceptable, further analysis on the parameter range alteration needs to be conducted to find an acceptable range that will be used for the main calibration process. Parameter range alteration should be conducted based on the following two aspects.
- Range for each parameter
- Adding or removing parameters

A feasibility test can be achieved by using two different methods to figure out key parameters: X-Y plots and statistical analysis.
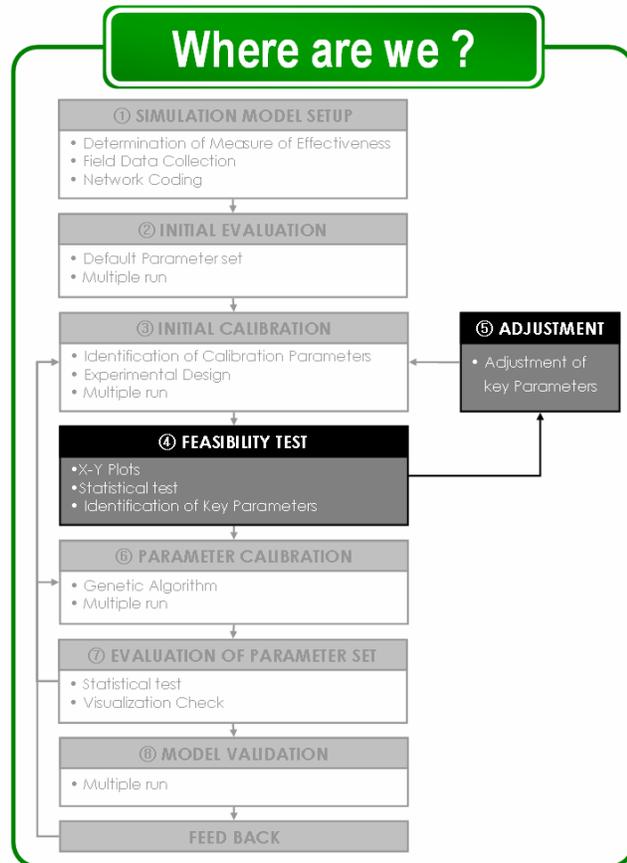
4.1.1. X-Y Plot

An X-Y plot produces a two-dimensional histogram with the variable in the analytical variable box plotted on the X-axis, and the variable in the reference variable box plotted on the Y-axis. The main purpose of using an X-Y plot is to discover the relationship between two different variables (analytical and response) regarding each variable's different level. Figure 18 shows an example of an X-Y plot. In this case, it is analyzing the relationship of a student's final examination score and the amount of hours the student spent studying for the exam.
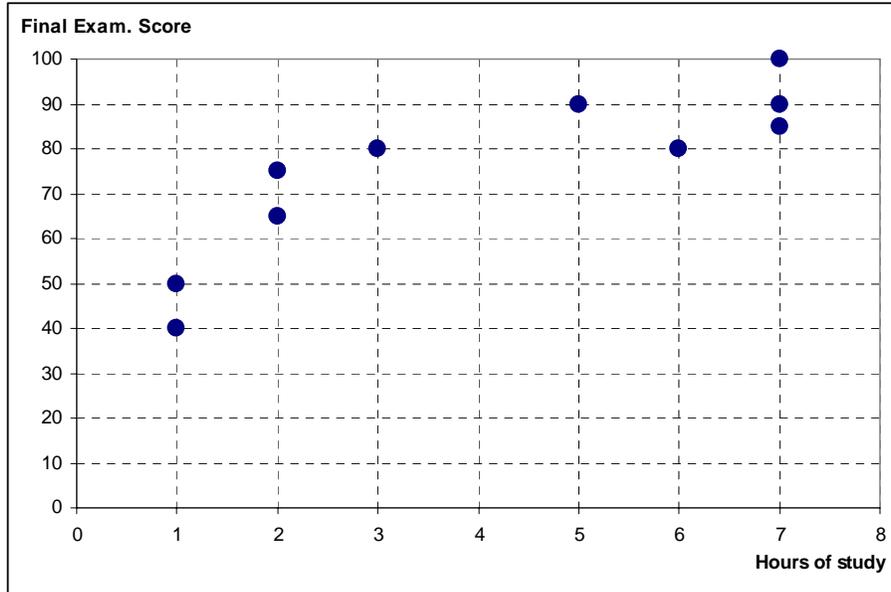
Figure 18. Example of X-Y Plot

With the X-Y plot in Figure 18, it is possible to conclude that the amount of time each student spends on preparing for the final examination is proportional to the final examination score. This is possible because the plotted result of each student's data showed approximate tendency.

In the same manner, it is possible to determine key parameters by observing the X-Y plot of each calibration parameter and the corresponding measures. If some relationship can be observed from the plot, it can be identified as a key parameter. On the other hand, if the dots on the plot are scattered without a specific pattern, it can be considered as a parameter that did not affect the corresponding measure significantly. Figure 19 shows an example of an X-Y plot of a calibration parameter that can be considered as a key parameter. In this case, "Mean Desired Speed" can be considered as a key parameter. If the initial parameter range set was not acceptable, this parameter can play an important role in the next step, which is parameter range adjustment.
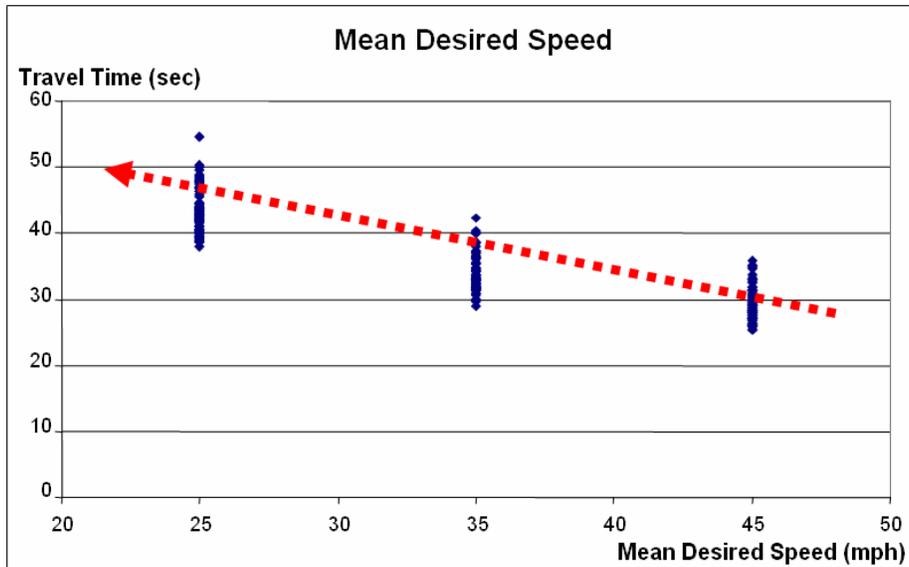
Figure 19. Example of X-Y Plot of Key Parameter

Figure 20 is showing a non-key parameter case. As can be seen from the figure, it is not showing any specific pattern. In this case, "Simulation resolution" can either be ignored during parameter adjustment procedure or omitted from the calibration parameter list because this parameter is not significantly affecting the measure.
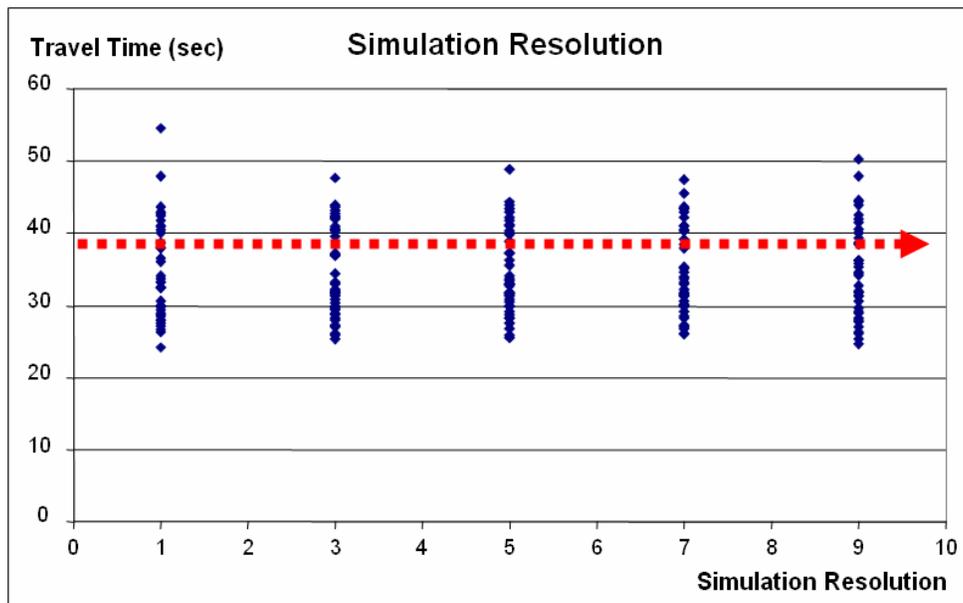
Figure 20. Example of X-Y Plot of Non-Key Parameter

4.1.2. Statistical Analysis

Key calibration parameters can also be determined by using a statistical method known as the Analysis of Variance (ANOVA). This test is widely used among engineers to test hypotheses

37

about the differences between two or more means. In other words, ANOVA tries to figure out whether two different variables, each calibration parameter and measure in this case, have a specific relationship with a select level of certainty. In this handbook, details on ANOVA will not be covered because it is not a main concern of this document and can be easily found from many different references on statistics. However, the way the user can obtain and analyze a result from the ANOVA is briefly mentioned.

ANOVA usually presents a result with an ANOVA table as shown in Table 14. Most of the statistical programs such as Microsoft Excel, Minitab, SPSS, etc. have an ANOVA function under the data analysis menu and all of the programs give a result in a similar format.

Table 15. Example of ANOVA Table

Sources of Variance	Degrees of Freedom	Sums of Squares	Mean Squares	F-Statistic	P-Value
Minimum Gap (ft)	3	457.3	152.4	5.68	0.003
Error	36	966.6	26.9	-	-
Total	39	1423.9	-	-	-

Among the many values on the ANOVA table, the only value the user should pay attention to is the "p-value". The p-value should be compared with a predetermined confidence level to select a key calibration parameter. For instance, if the user determined the confidence level to be 95% (95% is the most commonly applied confidence level in the engineering application), the predetermined confidence level is 0.05 (found by subtracting 0.95 from 1.00). So, by applying that confidence level to Table 15, minimum gap can be considered as a key parameter because the given p-value (0.003) is smaller than confidence level (0.05). If the p-value was greater than the confidence level, it should not be included in a key calibration parameter list.

4.1.3. 3D Contour Plot

Usually, the microscopic simulation model is known as a black box technology that tries to reproduce the field condition based not only on the parameter values that the users typed in but also the interaction of multiple parameters. Therefore, it is desirable to check whether there is a significant interaction between two or more specific calibration parameters or not to ensure the effect of range adjustment. For that purpose, a 3D contour plot was used due to its easy observation of three different measures. 3D contour plot is, as well known, frequently applied when there is a necessity to present the relationship between three measures and presents a third measure value that corresponds to the other two measure values with different color tones.   An example of 3D contour is shown in Figure 21.
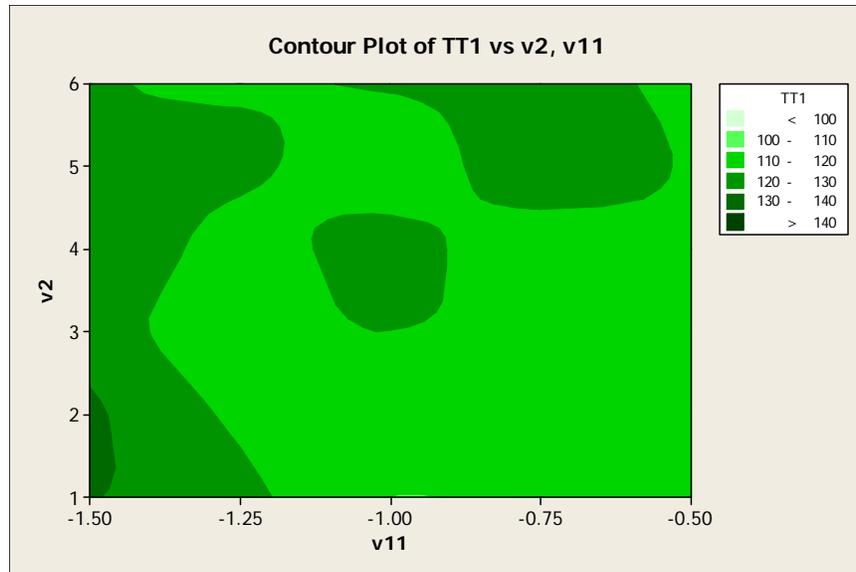
Figure 21. Example of 3D Contour Plot

As shown in the figure, the different output value can be observed with a different level of two calibration parameter values. Therefore, it is able to observe the trend of output value change with respect to the different level of two calibration parameter values and guess the interaction of two calibration parameters.

4.2. Parameter Range Adjustment

4.2.1. Independent Key Parameter

Once the key parameters are identified through ranges either on the X-Y plot or by statistical analysis, the next step is the adjustment of key parameter ranges and the omission or addition of calibration parameters. The parameter range can be modified by two different methods. First, simply shifting the range based on the X-Y plot that was obtained from the key calibration parameter determination step. In the case of Figure 22, it is clearly shown as a key calibration parameter. And if the measure (travel time in this case) needs to be greater than the current value, it will have a greater minimum gap value. Therefore, the maximum value of the minimum gap parameter range should be extended. However, a critical gap value larger than 6 seconds is not usually accepted. As such, a parameter may not be extended beyond 6 seconds.
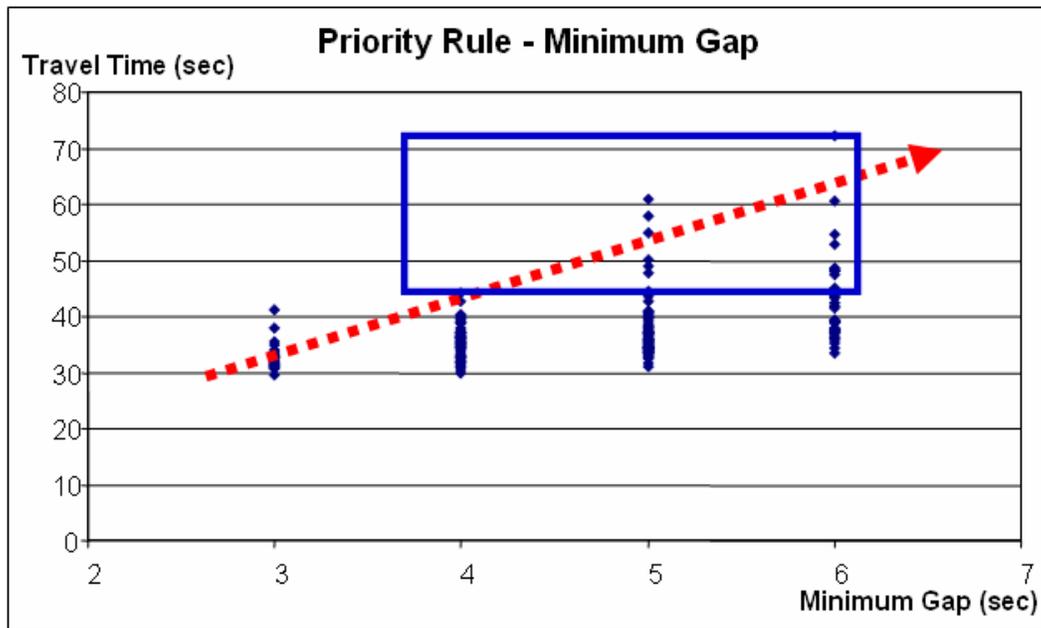
Figure 22. Example of Parameter Range Modification with X-Y Plot

Even though it is more appropriate to conduct the following procedure prior to determining the range of each parameter, it can also be applied to adjust the parameter ranges.

Another available method is comparing the field value or calculated value with a different method, such as the HCM procedure with the parameter range. For instance, speed data can be measured from the field by using many different methods such as radar. Then the field speed data can be compared with initial speed distribution. Saturation flow rate is another key aspect that affects the traffic flow model and other corresponding measures. In that case, saturation flow rate from the field measured data, by means of the HCM procedure and simulation, can be compared and the range can be modified by minimizing the differences of various results.

4.2.2. Non-Key Parameter

Even though some parameters are defined as non-key parameters, it is not appropriate to ignore all those non-key parameters because every single parameter is known to affect the simulation results despite its magnitude. Therefore, it is desirable to consider the combined effect of the non-key parameters by using 3D contour plot. Example contour plot is shown in Figure 23.
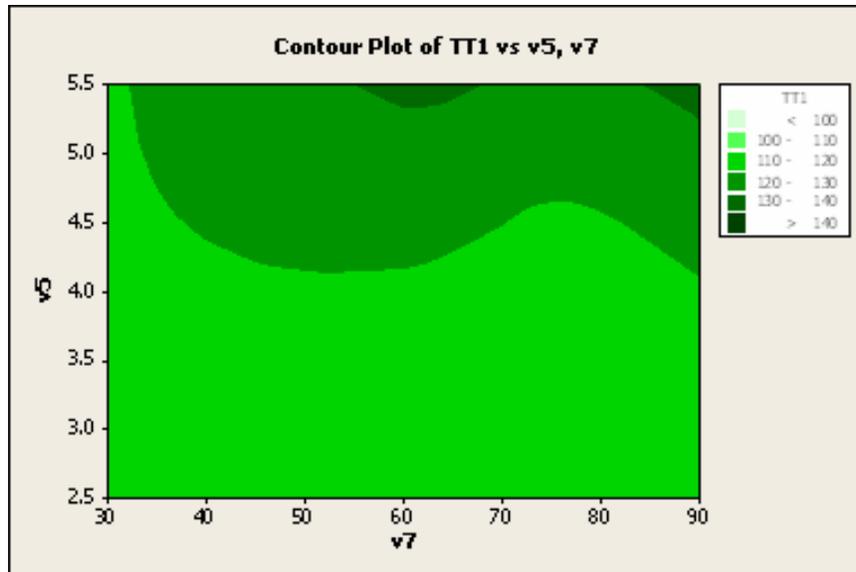
Figure 23. Example of Parameter Range Modification with 3D Contour Plot

Parameter 'v5' and parameter 'v7' in Figure 23 represent key parameter and non-key parameter. As shown in Figure 23, the combined effect of two parameters is presented. If only key parameter has an effect, then the pattern should be monotonic. In other words, because parameter 'v5' is a key parameter, the saturation of the green color should be increased consistently as the key parameter value changes. However, it is impossible to conclude that the saturation is changing consistently with the key parameter. Rather, it can be concluded that non-key parameter has a combined effect and non-key parameter also should be adjusted.

4.3. Parameter Range Acceptance

As stated in 3.2.2 through 3.2.4, once the parameter ranges are adjusted, the feasibility of adjusted parameter ranges need to be checked. Again, the result of multiple runs of simulation runs with a different calibration parameter combination needs to be presented in either histogram or X-Y plot format. In the case of single performance measure, the distribution of simulation output needs to include the field measured value within the 90% range of the distribution. In the multiple performance measures case, any data point should fall within the field-collected performance measure ranges. The concept of the parameter range acceptance is shown in Figures 24 and 25 again.
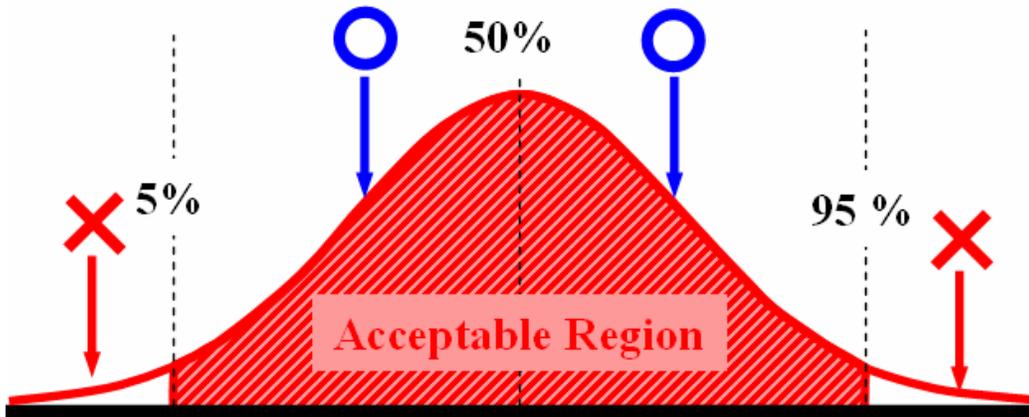
Figure 24. Concept of Parameter Range Acceptance—Single Performance Measure
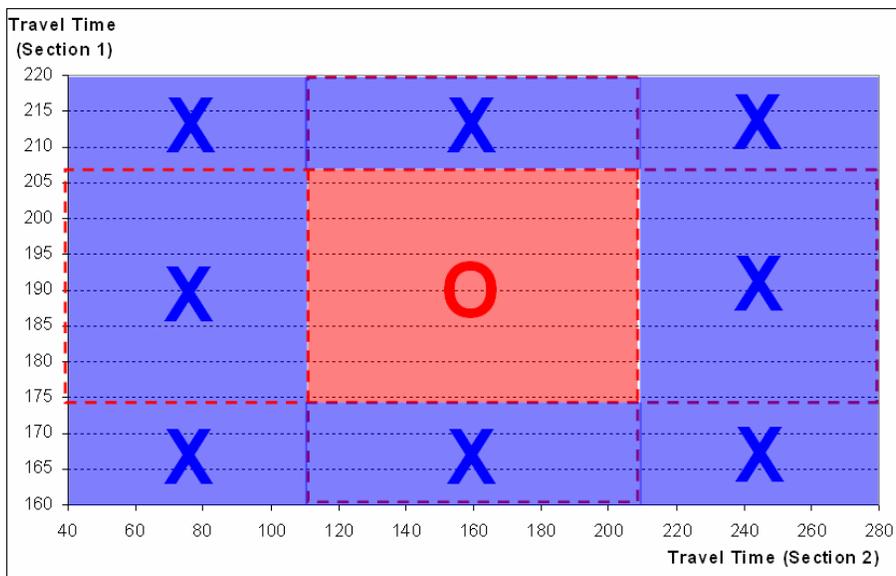
Figure 25. Concept of Parameter Range Acceptance—Multiple Performance Measures

Chapter 5
Searching for Best Parameter Set

In this chapter:
- Introducing the concept of a genetic algorithm
- Choosing parameter values using a genetic algorithm approach
- Evaluating the effectiveness of parameter sets
- Further parameter calibration using these results

5.1. Overview

Once the range of calibration parameters is determined, a specific set of parameters that reproduces the field condition needs to be found, which is what we call calibration. In this calibration procedure, we use an optimization method called a genetic algorithm (GA). The GA uses a specific number of digits, called a chromosome, that is generated at random. The chromosome's specific digits correspond to parameter values. By generating a single chromosome, values for each parameter are generated and with it a completely randomized simulation run can be completed.

5.2. Genetic Algorithm (GA)

A GA has several different steps that have to be conducted to complete the optimization (calibration) procedure. Figure 26 shows a concept of GA.

Where are we ?

① SIMULATION MODEL SETUP
- Determination of Measure of Effectiveness
- Field Data Collection
- Network Coding

② INITIAL EVALUATION
- Default Parameter set
- Multiple run

③ INITIAL CALIBRATION
- Identification of Calibration Parameters
- Experimental Design
- Multiple run

⑤ ADJUSTMENT
- Adjustment of key Parameters

④ FEASIBILITY TEST
- X-Y Plots
- Statistical test
- Identification of Key Parameters

⑥ PARAMETER CALIBRATION
- Genetic Algorithm
- Multiple run

⑦ EVALUATION OF PARAMETER SET
- Statistical test
- Visualization Check

⑧ MODEL VALIDATION
- Multiple run

FEED BACK

Figure 26. Conceptual Diagram of Genetic Algorithm

In this handbook, details on GA are not attempted to be explained because GA itself is a very complicated methodology that requires a significant amount of background knowledge that cannot be covered in a few pages. For more detailed explanations on GA, please refer to available references (Mitchell, 1996).

5.2.1. Evaluation

The bottom line of GA that needs to be understood by the user is the way it evaluates the candidate parameter combinations.

The way GA evaluates the performance of a new chromosome is the utilization of a fitness function. Once the fitness function is formulated, it calculates the fitness value. The result with each new chromosome can be used as an input value to the fitness function and generates the fitness value for that specific chromosome. Then, GA memorizes the fitness of each chromosome setting. For the purpose of calibration, the following fitness functions (equations) are used for the single performance measure case and multiple performance measures case.

Single Performance Measure Case

When only one performance measure is used for the fitness value calculation, relative error can be used as a fitness value calculation method. The fitness function that can be used for the single performance measure case is shown here.

$$\text{Fitness Value} = \frac{\left| \text{APM}_{Field} - \text{APM}_{Simulation} \right|}{\text{APM}_{Field}}$$

Where, APM_{Field}: The average value of performance measure from the field
$\text{APM}_{Simulation}$: The average performance measure value from multiple simulation runs

Multiple Performance Measures Case

When multiple performance measures are used for the fitness function, unlike the single performance measure case, the scale issue rises due to a different magnitude and unit. In order to avoid that problem, two different methods that can minimize the difference or separate the value with different scales are presented here.

- Log Transformation Method
- Constraint Insertion Method

Log Transformation Method

Log transformation is widely applied in practice where the data are skewed, or contain a significant number of outliers, or have unequal variation. For instance, when travel time and traffic volume data are considered as two performance measures, the effect of traffic volume alteration would be greater than the effect of travel time alteration. If field-measured travel time data are 150 and traffic volume is 2,000, 10% alterations for each measure represent 15 and 200. However, by transforming with log, this inequality can be minimized to 1.2 (log 15) and 2.3 (log 200). Therefore, all performance measure values both from the field and the simulation output were transformed into a log format and the fitness value calculated. The fitness function could be reformed as the following equation after inserting the log transformation concept.

$$\text{Fitness Value} = \sum_{i=1}^{M} \left(\frac{\sum_{j=1}^{N} \alpha_{ij} \cdot (\ln (| PM_{field\,(i,k)} - PM_{sim.\,(i,j)} |))}{\ln (APM_{field\,(i)})} \right)$$

Where, α_{ij}: 0, if j^{th} output with performance measure i falls within an acceptable range
         1, others
   $PM_{Field(i,k)}$: i^{th} Performance measure value from the field that collected on k^{th} day
   $PM_{Sim(i,j)}$: i^{th} Performance measure value from the simulation output from j^{th} replication
   $APM_{Field(i)}$: Average value of ith Performance measure value from the field

Constraint Insertion Method

The concept of the constraint insertion method is very simple and straightforward. The existing objective function only calculates the fitness value of each trial when the constraint with additional performance measure has been satisfied; otherwise, it imposes a certain fitness value, which is significantly higher than the usual fitness value that can be obtained. Figure 27 shows a concept of constraint insertion of additional performance measure. It is recommended to set a performance measure that considers being the primary criteria and has less variability as a value included in the main objective function and other performance measures can be inserted in the constraint format.
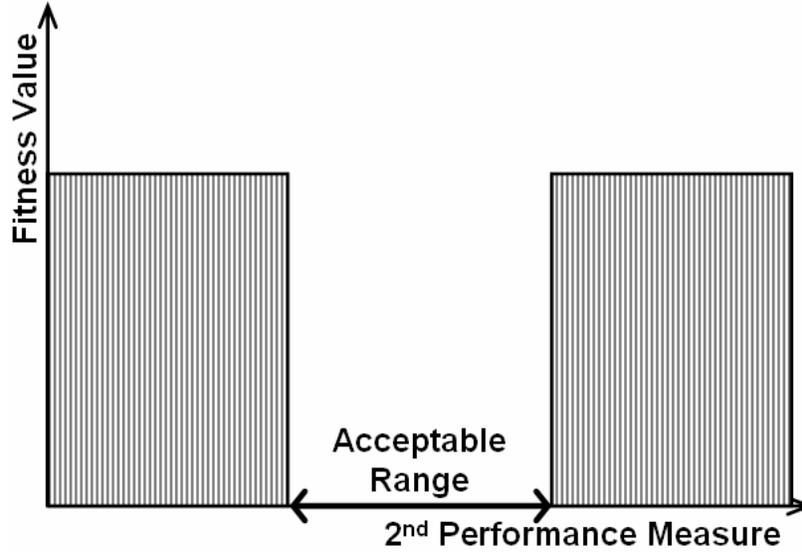
Figure 27. Concept of Additional Performance Measure Insertion—Constraint Insertion Method

The following equation shows the fitness function after inserting constraints.

$$\text{Fitness Value} = \sum_{i=1}^{M} \left(\frac{\sum_{j=1}^{N} \alpha_{ij} \cdot (\mid PM_{Field\,(i)} - PM_{Simulation\,(i,\,j)} \mid)}{APM_{Field\,(i)}} \right) \quad \dots\dots\dots\dots\text{(Eq. 3)}$$

$$\text{s.t. } PM_{min\,(k)} \leq PM_{simulation\,(k)} \leq PM_{max\,(k)}$$

Where, $PM_{min(j)}$: Minimum value out of jth Field measured performance measure value
$PM_{max(j)}$: Maximum value out of jth Field measured performance measure value

The equation is simply calculating the ratio of the difference between the performance measure value obtained from the field and that from the simulation. If the evaluation result is acceptable (i.e., fitness value is less than predetermined criteria), then the current chromosome can be considered as a calibrated result. However, if the result is not acceptable, the whole process needs to be conducted again with the current chromosome as an initial value. A simplified GA procedure is shown in Figure 28.

46
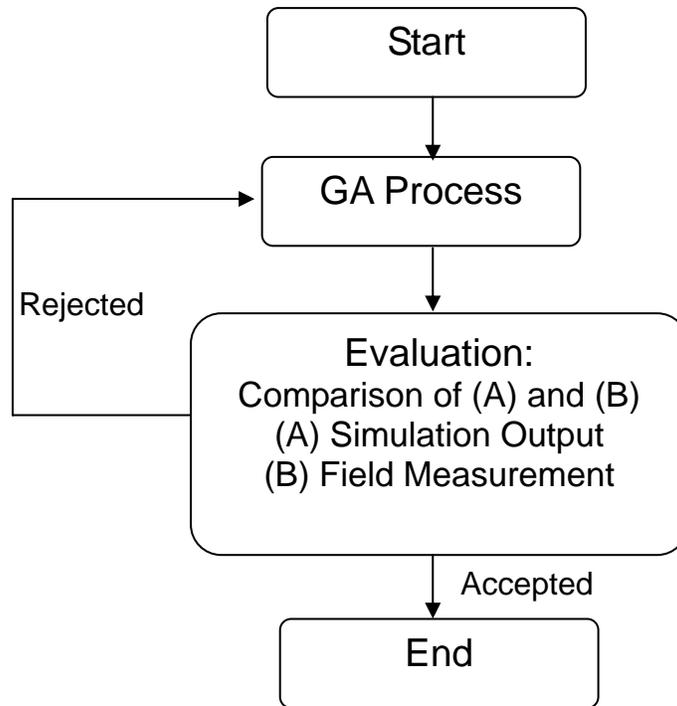
Figure 28. Conceptual Diagram of Calibration Procedure

5.3. Fitness Confirmation

As stated previously, the performance of every calibration parameter set that was generated from a previous procedure was investigated by using a certain fitness function. As the number of generations increases, the fitness value (outputs of fitness function) becomes smaller because GA keeps trying to find a calibration parameter set that reproduces the field condition more closely. Figure 29 shows an example of a fitness value convergence as the number of generation increases. In the figure, the following equation was used as a fitness function. The blue and pink colored lines in the figure represent the best and average fitness value within each generation, respectively.

$$\frac{\left| \text{Travel Time}_{Field} - \text{Travel Time}_{Simulation} \right|}{\text{Travel Time}_{Field}}$$
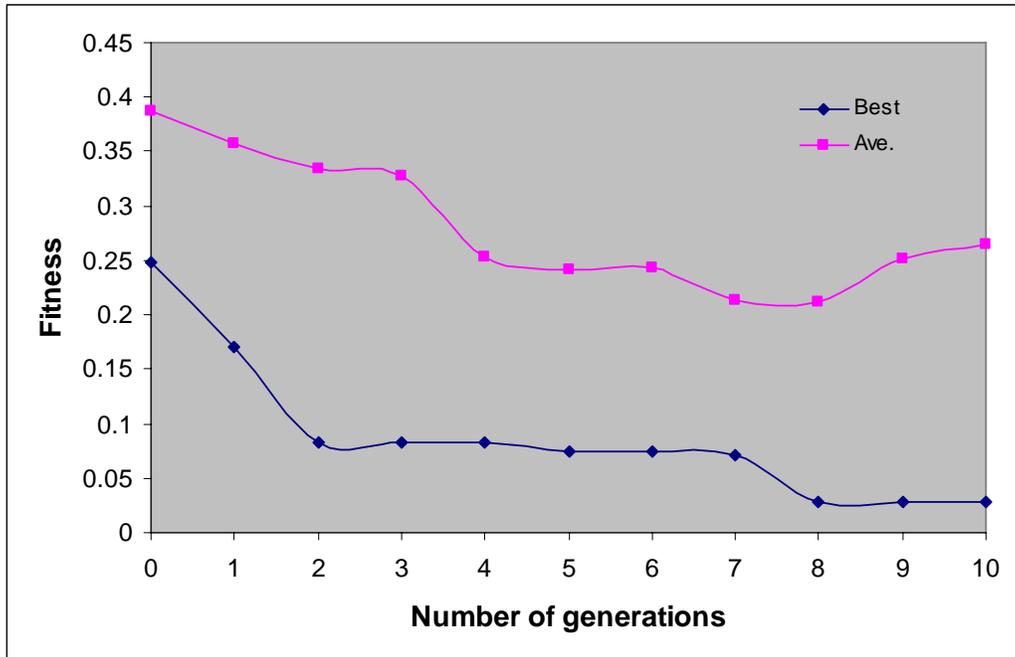
Figure 29. Example of Convergence

As the number of generations increase, up to a certain point, a better calibration parameter set is found that reproduces the field condition more closely.

5.4. Verification of Parameter set

Once the calibration procedure is completed and a calibrated parameter set is obtained, the performance of a calibration parameter set with a different random seeded number needs to be conducted. As was done in Chapters 3 and 4, the distribution of simulation outputs that correspond to the field-measured data should be presented in a histogram format and compared with the performance measured from the field.

Chapter 6
Evaluation and Validation of Calibrated Parameter Set

In this Chapter:
- Comparison of calibrated parameter set with default parameter set
- Visualization check with animation
- Validation of the model with calibrated parameter set

This chapter consists of two main components:
- Evaluation of parameter sets
- Validation of calibrated parameter set

These comprise the last step of the model calibration and validation procedure that mainly concerns the verification of a calibrated parameter set. Detailed instructions on each aspect is described in the rest of this chapter.

6.1. Evaluation of Parameter Sets

Since the calibration parameter set has been obtained from a previous procedure, it is necessary to investigate the relative performance of the model with one calibrated parameter set to another parameter set, which confirms the better performance of a calibrated parameter set. Moreover, it is important to check whether the calibrated parameter set can reproduce the field condition or not.

6.1.1. Default vs. Calibrated Parameter Set

As the first step of evaluation, the performance of a default and calibrated parameter set needs to be compared. As stated in the earlier part (Chapter 2) of this handbook, this calibration procedure is necessary when a simulation model with a default parameter set does not reproduce the field condition. Thus, it is required to check whether the calibrated parameter set reproduces the field condition or not, as well as the relative performance of a calibrated parameter set with a default value.
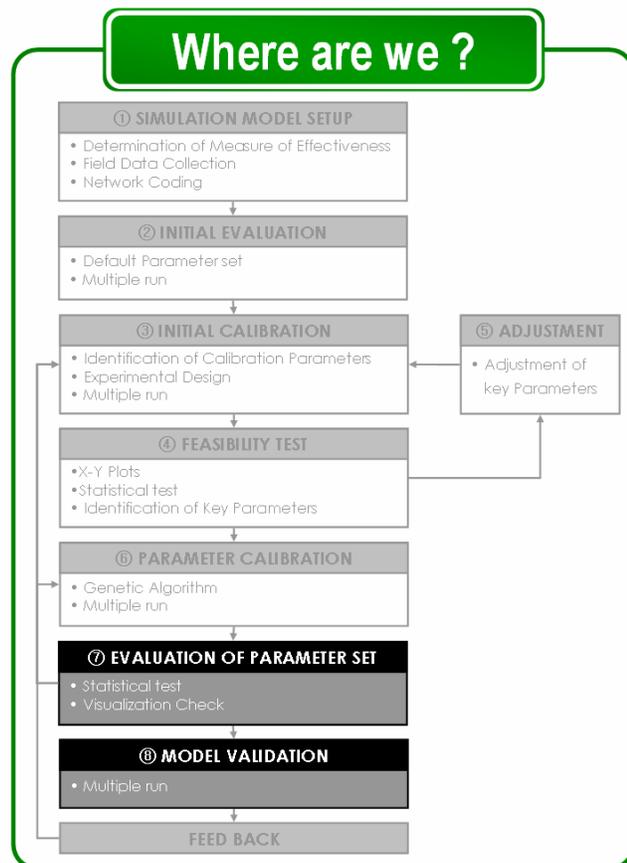
The performance of each model can be compared by simply using a line chart (or histogram) that has been used previously to check the acceptability of each model (see Chapter 3). Similar to what has been done from a previous step, multiple runs (say 100 runs) for each model with a different parameter set need to be conducted to represent the variability of performance measured as the random seed number changes.

As shown in Figure 20, three different aspects can be presented on the same chart, which are as follows:
   - Distribution of 100 results with a default parameter set (Blue line)
   - Distribution of 100 results with a calibrated parameter set (Red line)
   - Performance measure collected from the field (Black arrows and numbers)

Since the precondition of conducting a calibration procedure was that the model with a default parameter set does not reproduce the field condition, the distribution of a performance measure with a default parameter set may not include some of, or the entire, field measured values. In the same manner, the distribution of a performance measure with a calibrated parameter set will include the entire field-measured values. Therefore, usually, the evaluation result seems definite in most cases; however, this step mainly works as a last confirmation step to prevent an unreasonable result. For instance, as can be seen from Figure 30, the distribution of the performance measure needs to be compared with the performance measure from the field.
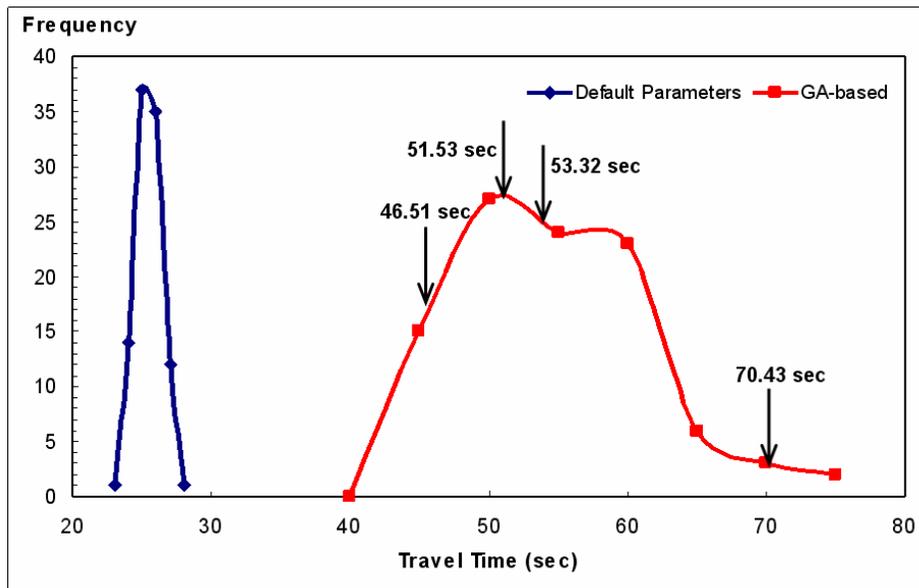
Figure 30. Example of Evaluation of Parameter Sets with Line Chart

If the result is similar to that presented in Figure 30, it is possible to conclude that the model with a calibrated parameter set performs better than the model with a default parameter set and reproduces the field condition adequately. If the result is opposite or the distribution of a simulation output does not include the performance measure from the field, it is necessary to conduct a calibration procedure again.

When multiple performance measures are used for the calibration, it is more appropriate to use an X-Y plot analysis as was done from the previous parts (see 2.2.2). A different simulation output set should be presented with different colors and 90% confidence interval region should be compared with the field-collected performance measure ranges. Refer to Chapter 2 for details.

Other parameter sets that could be included in the comparison are the best-guessed parameter set and the optimal parameter set identified in the LHD approach, where best-guessed means were used to assign the most accurate value to each parameter based on engineering judgment and knowledge of local traffic conditions.

6.1.2. Visualization Check

The final step that needs to take place in the calibration procedure is a visualization check. Since this calibration work is done by computer, it is important to ensure the reliability of the calibrated parameter set. As shown in Figure 31, sometimes GA produces a parameter set that gives an acceptable output distribution, but shows unrealistic vehicle movements. So, it is necessary to check an animation of each simulation model with the random seed that produced an output at a certain percentile (say 50th, 75th, 85th, etc). It is very important and conducting this procedure is strongly encouraged.
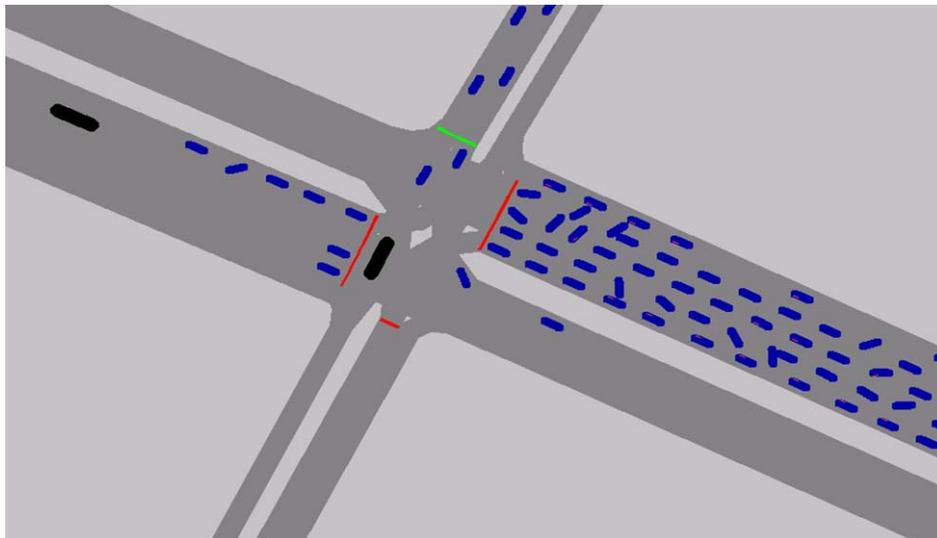
Figure 31. Example of Unacceptable Animation

6.2. Validation Process

The very last step of simulation model calibration and validation procedure is the validation with untried data. It is meaningful because the data that have not been used for the calibration procedure are used, which indicates, if successful, that the model with a calibrated parameter set works for different cases.

First, a new simulation model with untried data needs to be built. Then, multiple runs with different random seeded numbers with the calibrated parameter set needs to be conducted. The

simulation outputs that correspond to the performance measure collected from the field need to be extracted and presented in a histogram format as shown in Figure 32. If the performance measure from the field falls within the 'Acceptable Region,' it can be considered a well-calibrated and validated parameter set. However, if the field-measured value does not fall within an 'Acceptable Region,' the calibration procedure needs to be conducted again.

In the multiple performance measures case, the X-Y plot again needs to be used to conduct the validation step. As shown in Figure 33, 90% confidence interval region of simulation output data needs to overlap the combination of field-collected performance measure data. If this condition could not be satisfied, the GA optimization step needs to be conducted again.
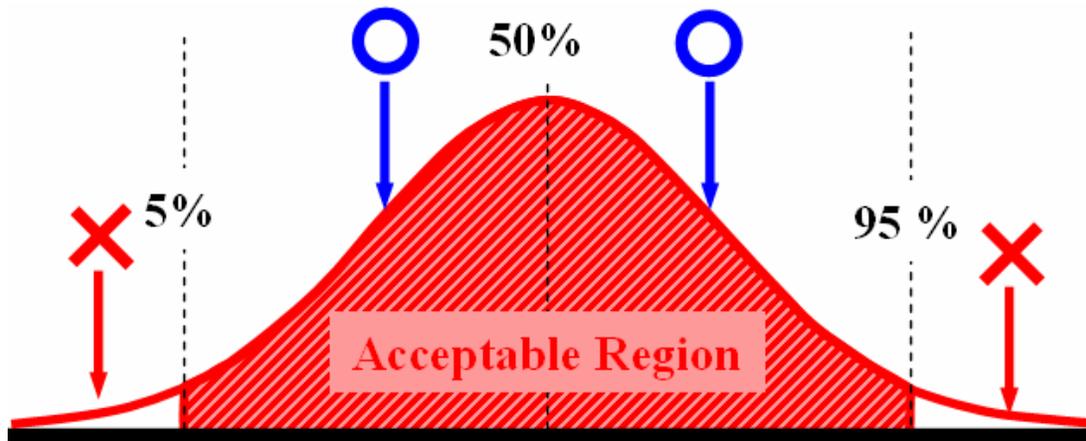
Figure 32. Concept of Acceptable Region—Single Performance Measures
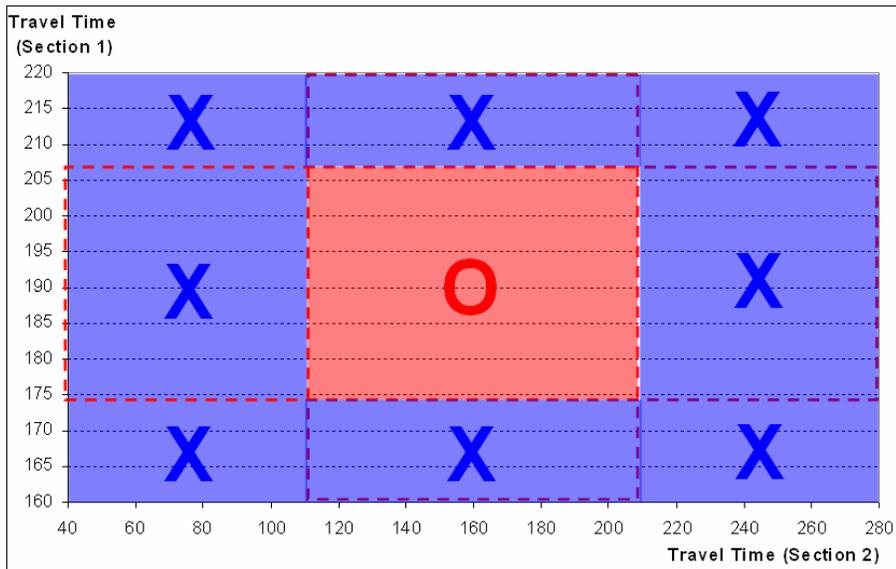
Figure 33. Concept of Acceptable Region—Multiple Performance Measures

Part III
Case Studies

In this part:

Actual application case studies are introduced in this part.

This part consists of four chapters and each chapter talks about the following subjects:

- Chapter 7: Isolated signalized intersection
- Chapter 8: Urban Network
- Chapter 9: Highway Basic Segment
- Chapter 10: Highway Merging Segment

Chapter 7
Isolated Signalized Intersection

In this chapter:
- Case study of calibration and validation with an isolated intersection

---

7.1. Site Description

This site is an isolated intersection with an actuated signal system. It is the junction of U.S. Route 15 and U.S. Route 250 at Zion Crossroads, Virginia, which has four legs with single lanes for all four approaches. This site is referred to as "Site 15" throughout the remaining part of this handbook for convenience. The location and alignment of the test site is shown in Figures 34 and 35.

Figure 34. Schematic of Site 15

Figure 35. Geometry of Site 15

The geometry for each approach is identical with 110 feet long right-turn bay and shared left-turns and through lane. The southbound approach carries relatively heavy traffic volume because this intersection is a little south of the I-64 exit and both routes are major routes in this area. Also, large proportions of vehicles make left-turns (toward Richmond, Virginia) during peak hour.

7.2. Data Collection

7.2.1. Required Types of Data

The data needed in this case study included simulation input data and performance measure data for calibration and validation purposes. To build the network in each simulation model, input data such as traffic counts, intersection geometric characteristics, detector locations (for signal), and posted speed limits were required.

The signal timing plan was obtained from VDOT personnel. The performance measure used in this network was travel time on the southbound approach because it carries the heaviest volume among four approaches. Furthermore, it was easy to collect both from the field and the simulations. Comparison of other measures such as delay and queue length data are considered to be not so meaningful because each simulation model calculated the measure in their own way. For example, CORSIM provides queue length data in the number of vehicles format; however, VISSIM only provides the length of the queue. While converting the queue length data to the number of vehicles format, another error might occur that diminishes the quality of the calibration result. The details on this required data are described in the following part.

7.2.2. Data Collection

Based on the preliminary site visit, the data collection plan was decided to include the following: one smart travel van (STV), two video cameras, and five persons, as shown in Figure 36.

Person 2 recorded actual signal timing changes from the actuated signal system during the data collection period. Persons 1 and 3 were responsible for the two video cameras, which were located in the southbound approach to obtain the license plates and times of each vehicle at the upstream and downstream points. The distance between these two cameras was 960 feet. The

other two cameras attached to the STV collected detailed data in two areas corresponding to the dashed circle and rectangle in Figure 36. The camera for circular area provided traffic counts and vehicle classification for each approach, whereas the camera for the rectangular area provided detailed traffic movements on the southbound approach.

To account for day-to-day variability, data were collected during the evening peak hour between 5:00 and 6:00 pm on multiple weekdays, April 15, 22, May 13, 20, and June 5 in 2003. Synchronization between clocks of all equipment such as video cameras, surveyor's watch and video recorder in the STV were performed before data collection.
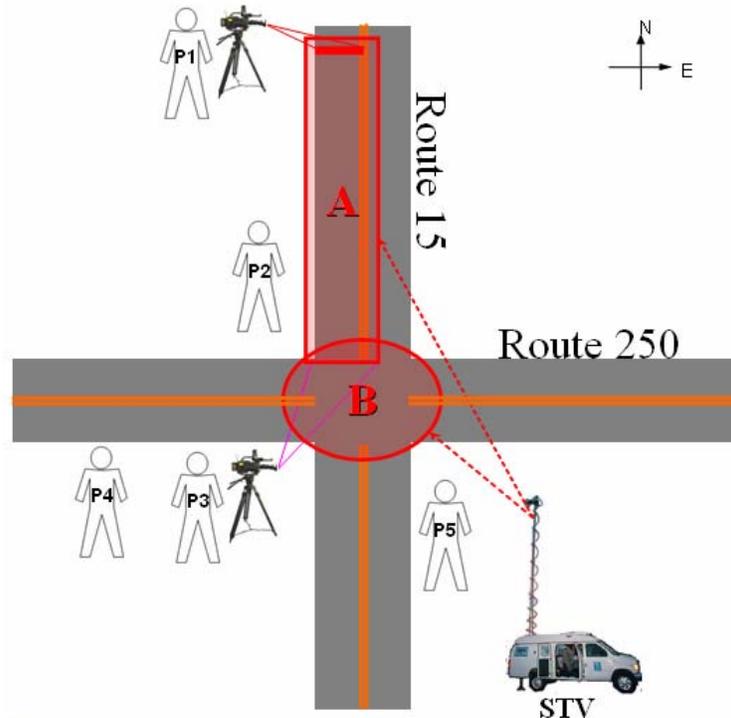
Figure 36. Site 15 Data Collection Plan

7.2.3. Data Reduction

Videotapes were reduced to obtain traffic counts, individual travel times on the southbound approach segment, and real signal timing. Traffic counts and heavy vehicle counts for each approach were obtained through watching the video attached to the STV. The signal statistics such as maximum/minimum green, red and yellow times were extracted from the recorded signal data. The posted speed limit of 45 mph was observed at the site. Travel time was collected by matching license plates through watching videotapes from two video cameras that were recording from each end of the travel time measuring segment. The procedure was taken to note the license plate numbers and times of each vehicle passing the observation points, then match the license plate numbers one by one, and finally calculate the time differences and record travel times.

7.2.4. Data Summary

Signal Timing

Actuated signal setting implemented in the field actuated signal system in terms of phases, splits, minimum and maximum green times, and gap out times are shown in Table 16. Since there were only two phases in this signal system, real time signal data were recorded only at major directions: southbound and northbound, as listed in Table 16. The signal timing for minor approaches could be inferred from Table 17.

Table 16. Actuated Signal Setting—Site 15

Lane Group	SBTL and NBTL	EBTL and WBTL	Lane Group	SBTL and NBTL	EBTL and WBTL
Phase	2	4	**Red Clearance**	1.2	1.0
Minimum Initial	12	6	**Added Initial**	1.5	0
Passage Time	2.5	2.5	**Maximum Initial**	20	0
Maximum Green	30	30	**Minimum Gap**	2.5	2.5
Yellow	4.5	4.5	**Red Clearance**	1.2	1.0

Table 17. Actual Signal Timing—Site 15

Major Signal	Green (sec)			Red (sec)		
	Mean	Min	Max	Mean	Min	Max
5/13/03	30.9	12.6	58.2	24.1	12.6	36.9
5/20/03	29.6	12.0	48.1	23.0	12.1	37.2
6/5/03	30.6	13.5	70.6	25.0	6.1	37.3
Avg.	30.4	12.7	59.0	24.0	10.3	37.1

Traffic Counts

Table 18 lists traffic counts on four individual days and average traffic counts on three days: April 22, May 13, and May 20 in 2003. Data from these three days were used in the calibration process while data on June 5 were reserved for validation. Trucks and other multi-axles large vehicles were identified as heavy vehicles (HV), and converted into a percentage of the total number of vehicles.

Table 18. Field Traffic Counts—Site 15

Site 15 Traffic Counts (vph)			Left Turn	Through	Right Turn	HV%
Calibration Data	4/22/2003	Southbound	217	312	122	3
		Northbound	36	188	40	2
		Eastbound	80	86	33	1
		Westbound	40	64	80	3
	5/13/2003	Southbound	195	270	111	2
		Northbound	22	193	40	5
		Eastbound	75	111	42	3
		Westbound	34	82	104	4
	5/20/2003	Southbound	208	289	99	1
		Northbound	32	171	33	4
		Eastbound	83	94	53	2
		Westbound	33	74	90	2
Validation Data	6/5/2003	Southbound	192	300	93	5
		Northbound	32	172	29	6
		Eastbound	88	98	44	4
		Westbound	45	77	84	2

Travel Time

Table 19 shows statistics of travel time including mean, median, standard deviation, and minimum and maximum values. The data indicate a huge variation of mean travel time among different days ranging from 46.51 to 70.43 seconds. It is noted that June 5, 2003 data were reserved for validation and the other three days were used in the calibration. Even on the same day, travel time of individual vehicles varied dramatically as indicated by standard deviation, minimum and maximum values. The distributions of travel time on individual days are illustrated in Figure 37.

Table 19. Field Travel Time—Site 15

| Travel Time (sec)	Mean	Median	St. dev	Min.	Max.
4/22/03	70.43	68.00	27.98	15.00	180.00
5/13/03	53.32	46.00	25.64	16.00	121.00
5/20/03	46.51	44.00	18.91	17.00	165.00
6/05/03	51.53	42.00	30.89	17.00	157.00
Ave.	55.45	50.00	25.86	16.25	155.75
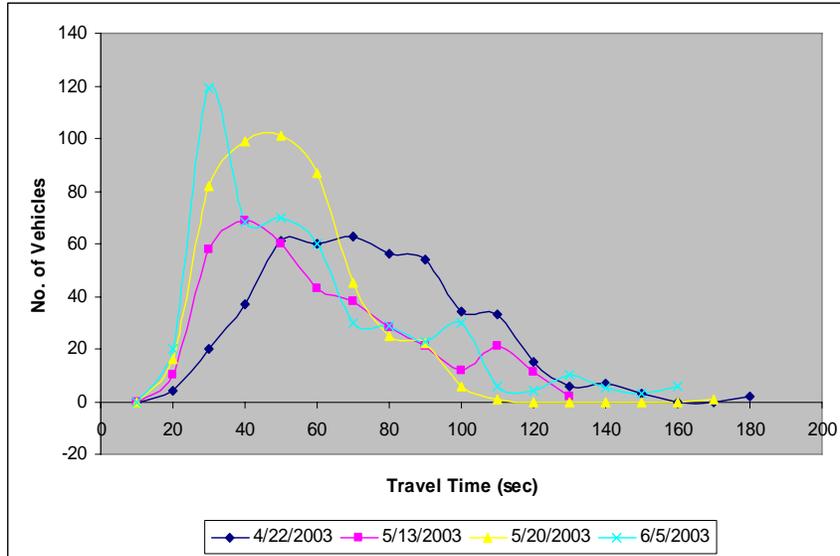
Figure 37. Field Travel Time—Site 15

7.3. CORSIM Calibration

7.3.1. Identification of Calibration Parameters

CORSIM includes numerous calibration parameters that allow users to fine-tune the model to replicate observed conditions and can be found in the user's manual. In addition to the parameters with acceptable ranges, some parameters in CORSIM are represented by a discrete distribution (e.g., amber interval response, see Table 20) or by 10 percentile values (e.g., distribution of free flow speed by driver type, see Table 21) indexed by 10 driver types from conservative to aggressive.

Table 20. Default Value Set for Amber Interval Response

Driver type	1	2	3	4	5	6	7	8	9	10
Acceptable deceleration	56	50	46	42	39	37	34	30	26	20

Table 21. Default Distribution of Free-Flow Speed Percentages

Driver type		1	2	3	4	5	6	7	8	9	10	Sum
Percentage multiplier of free-flow speed	NETSIM	75	81	91	94	97	100	107	111	117	127	1,000
	FRESIM	88	91	94	97	99	101	103	106	109	112	1,000

As can be seen from Table 21, the sum of the 10 percentile values must equal 1,000. Since driver aggressive data are not available, it was necessary to try a wide range of driving behavior

59

parameters. In order to increase aggressiveness or conservativeness for a population of drivers, the following parameters were included as calibration parameters:
- The distribution alternatives of amber interval response
- Gap distributions for left turns
- Gap distributions for right turns
- Distribution of free flow speed by driver type
- Start-up lost time distribution
- Discharge headway distribution

The designed range for each parameter is shown here.
1) Link mean free flow speed (mph): 35 ~ 45
2) Mean queue discharge headway (1/10 sec): 15 ~ 30
3) Mean start-up lost time (1/10 sec): 15 ~ 30
4) Left turn jumper probability (%): 10 ~ 40
5) Left turn speed (mph): 9 ~ 21
6) Right turn speed (mph): 9 ~ 21
7) Left-turn lagging within 2 seconds (%): 20 ~ 50
8) Left-turn lagging for 2-4 seconds (%): 5 ~ 15
9) Amber interval response (fpss)
   Default: 21, 18, 15, 12, 9, 7, 6, 5, 4, 4
   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 (sec)
    Default: 7.8, 6.6, 6.0, 5.4, 4.8, 4.5, 4.2, 3.9, 3.6, 2.7
    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 (sec)
    Default: 10, 8.8, 8.0, 7.2, 6.4, 6.0, 5.6, 5.2, 4.8, 3.6
    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 (%)
    Default: 75, 81, 91, 94, 97, 100, 107, 111, 117, 127
    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 (%)
    Default: 218, 140, 125, 118, 102, 86, 78, 63, 47, 23
    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 (%)
    Default: 170, 120, 120, 110, 100, 100, 90, 70, 70, 50
    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

7.3.2. Experimental Design for Calibration

Two hundred parameter combinations were generated within the initial parameter ranges determined by using the Latin Hypercube Sampling method.

Ten random seeded runs were conducted in CORSIM for each of the 200 combinations, for a total of 2,000 runs. Usually, 5 random seeded runs are conducted to check variability of each parameter set; however, CORSIM is known as a simulation model that generates a set of outputs with wide variability. So, 10 random seeded runs were selected to be conducted to take into account this wide variability. The average travel time of the southbound approach of each random seeded run was recorded for 2,000 runs and the results from the 10 multiple runs were then averaged to represent each of the 200 parameter sets.

7.3.3. Feasibility Test

As mentioned from the previous part of this handbook, a feasibility test was conducted to check whether the simulated outputs based on the current parameter ranges could capture the field data and to identify the key calibration parameters in CORSIM.

The travel time histograms of 200 cases are shown in Figure 38. The field travel time data (56.75 seconds) roughly fall within the top 5% of the distribution. It indicates that the selected parameters and their ranges were at the boundary so that sometimes these parameter ranges cannot reproduce the field condition. In order to shift the simulation output distribution to the right side to capture the field condition, key parameters that most affected the results needed to be identified. Then, the initial range of key parameters was expanded to an appropriate level.
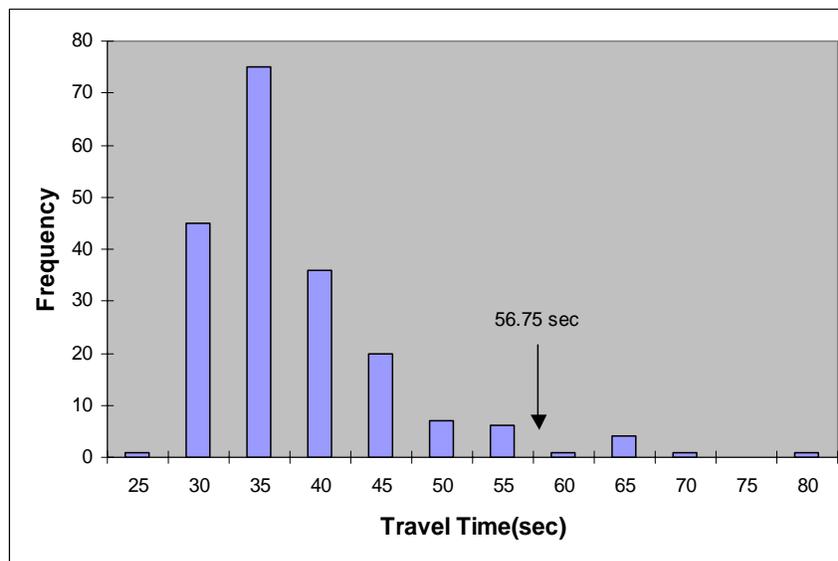

Figure 38. Feasibility Test Results for Site 15 with CORSIM

According to the X-Y plots of each parameter versus CORSIM travel time, as seen in Figure 39, link mean free flow speed (a), mean queue discharge headway (b), and gap distributions for left turns (j) were identified as key calibration parameters.
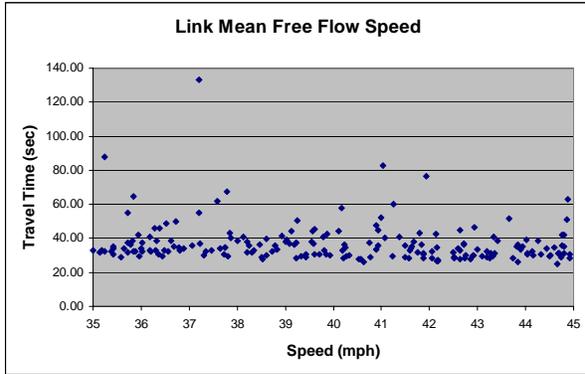
(a) Link Mean Free Flow Speed
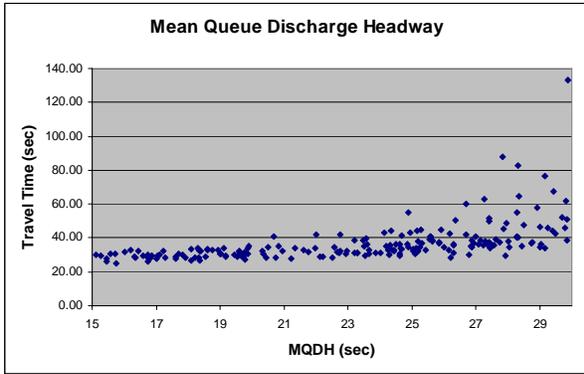
(b) Mean Queue Discharge Headway
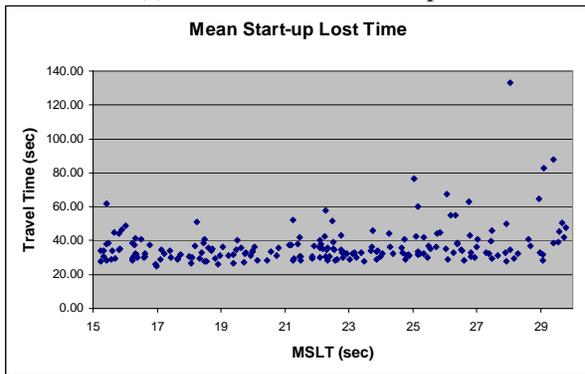
(c) Mean Start-up Lost Time
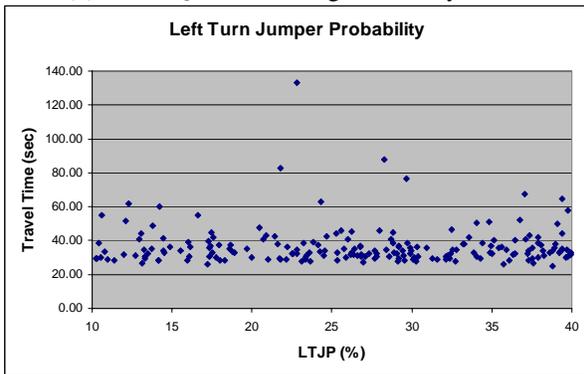
(d) Left Turn Jumper Probability
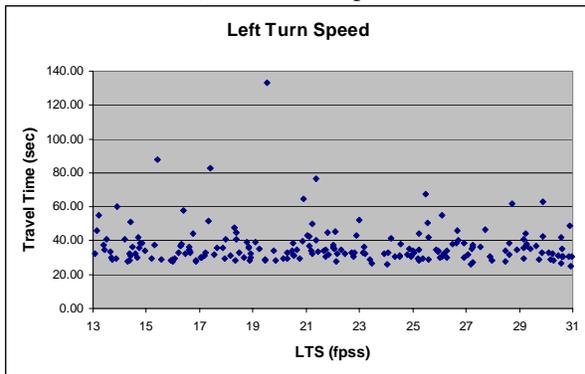
(e) Left Turn Speed
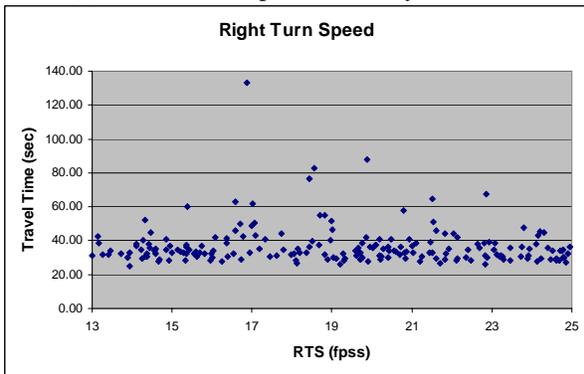
(f) Right Turn Speed
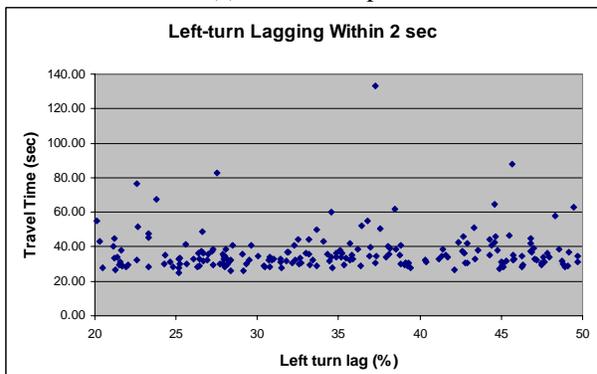
(g) Left Turn Lagging Within 2 sec

(h) Left Turn Lagging Within 2~4 sec

62

(i) Amber Interval Response

(j) Gap Distribution for Left Turns

(k) Gap Distribution for Right Turns

(l) Distribution of Free Flow Speed by Driver Types

(m) Start-up Lost Time Distribution

(n) Discharge Headway Distribution
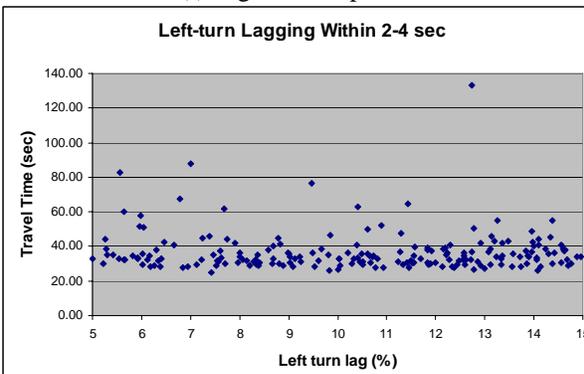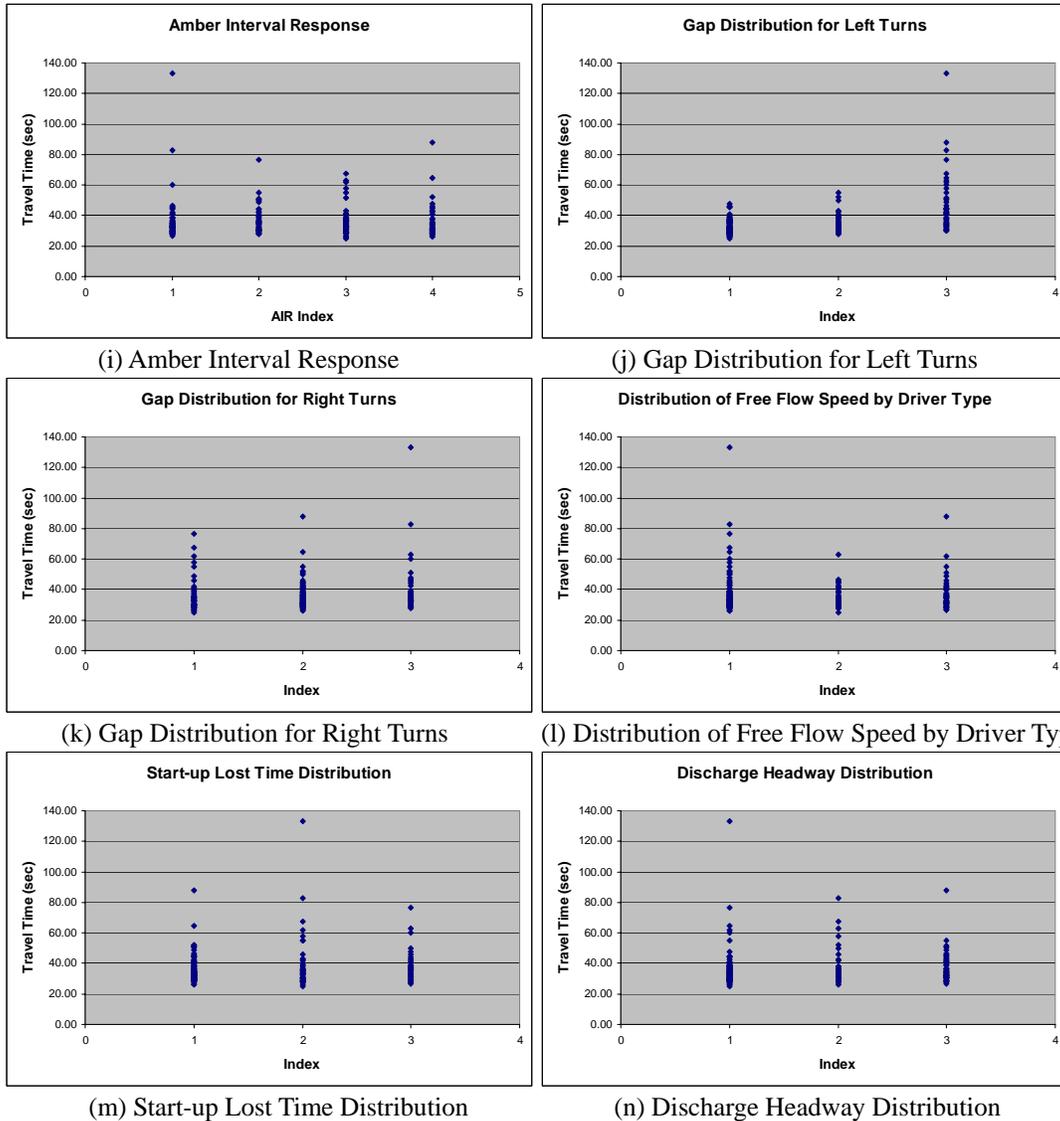
Figure 39. X-Y Plots of Travel Time vs. Calibration Parameters—CORSIM

A statistical test (ANOVA) was conducted to identify the key parameters and the results were summarized in Table 22. Prior to analysis of the statistical result (ANOVA table), a significance level needs to be selected. Here, 0.05 was selected in this case, which is a usual value in most cases. Following is the list of calibration parameters that showed their p-values to be less than 0.05 and selected as key calibration parameters.

- Gap distribution for left turn
- Mean queue discharge headway
- Mean start-up lost time
- Link mean free flow speed

63

Table 22. ANOVA Results for Site 15 with CORSIM

Site 15 – CORSIM	p-value
Link mean free flow speed	**0.045**
Mean queue discharge headway	**0.000**
Mean start-up lost time	**0.000**
Left turn jumper probability	0.440
Left turn speed	0.806
Right turn speed	0.292
Left-turn lagging within 2 seconds	0.747
Left-turn lagging within 2-4 seconds	0.159
Amber interval response	0.913
Gap distribution for left turns	**0.000**
Gap distribution for right turns	0.303
Distribution of free flow speed by driver type	0.375
Start-up lost time distribution	0.347
Discharge headway distribution	0.935

* Significant value is less than 0.05.

Note that the calibration parameters other than "Mean start-up lost time" were identified as a key parameter from both cases (X-Y plot and statistical analysis). The reason that the Mean start-up lost time was missed from the analysis of the X-Y plot as a key parameter was that it may have a joint effect with another parameter(s), so it was not able to be captured by the X-Y plot analysis that only deals with a single parameter. Continuous comparison of simulation animation and field observations was conducted to find a difference between simulation animation and field traffic flow. As a result, left-turn vehicles on the south approach sometimes waited for a long time to find an acceptable gap. Since it is a one-lane approach, these left-turn vehicles blocked an approach and caused long queue and delay for the trailing vehicles. On some occasions, vehicles waited for two cycles before being discharged. Such a phenomenon was not observed in the simulation animation. Accordingly, the following calibration parameter was added to consider this kind of longer gap problem in the field.

- Gap distribution for left-turn vehicles

Another 200 scenarios were generated and multiple CORSIM simulations were conducted based on the new parameters and their ranges. Figure 40 shows the result with the new parameter ranges, of which the field travel time falls within the distribution of the simulation outputs. Therefore, the current parameter ranges could be used for further calibration procedure.
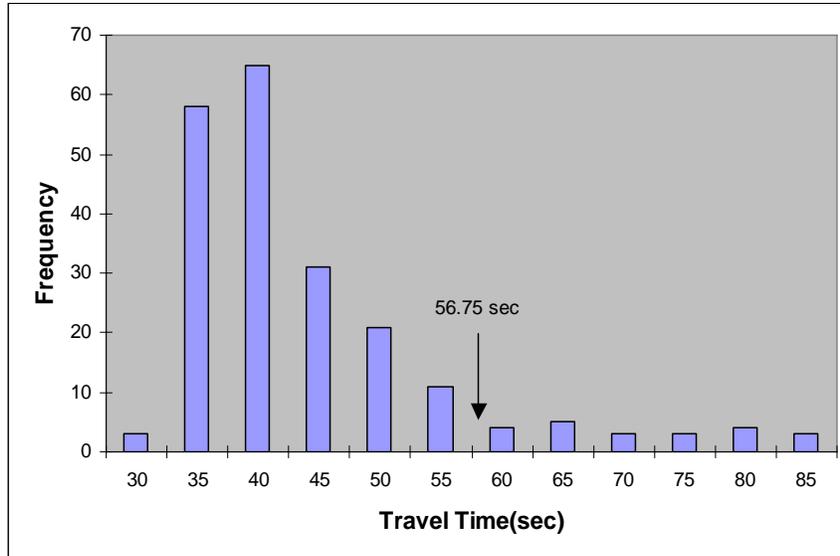
Figure 40. Feasibility Test Results for Site 15 with CORSIM

Parameter combination that showed the best performance is saved for the final evaluation part.

7.3.4. Parameter Calibration Using Genetic Algorithm

A genetic algorithm was integrated with the CORSIM model to calibrate each parameter value. This algorithm was designed to minimize the result of the following equation that minimizes the difference between the field-measured travel time and simulation travel time outputs.

$$\frac{\left| \text{Travel Time}_{Field} - \text{Travel Time}_{Simulation} \right|}{\text{Travel Time}_{Field}}$$

The simulation travel time output was the value averaged from the 10 CORSIM runs of each parameter set. Ten generations and 10 populations were adopted in the GA process. The convergence of fitness value with generations is shown in Figure 41. As can be seen from the figure, the fitness value of the best parameter set is continuously changing because the GA tests various combinations of parameters and tries to eliminate the parameter values as well as the combinations as they grow higher. The parameter set with the best fitness value was selected to represent the local traffic conditions and selected in the final evaluation.
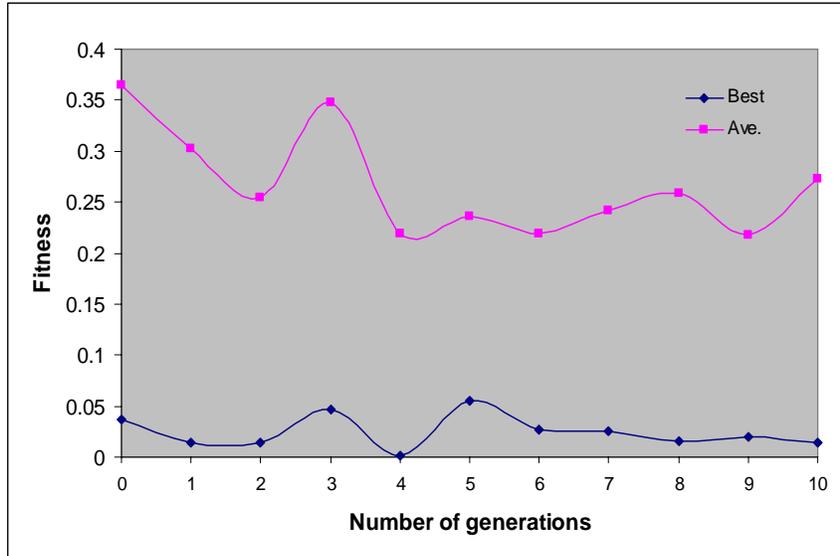
Figure 41. Convergence of Fitness Value with Generation

7.3.5. Evaluation of the Parameter Sets

For verification purposes, the comparison of the performance of each model with default parameters and GA-based parameters was conducted. These parameter values are summarized in Table 23. Figure 42 shows the distribution comparison of simulation travel time based on the three parameter sets. The result shows that the simulation model with calibrated parameter sets outperforms the model with default parameters and all three field data sets fall into their distributions. All the parameter values look reasonable, too. Therefore, this calibrated parameter set can be used for further analysis.

A visualization check was conducted with the animation tool that CORSIM provides (Trafvu) and no abnormal traffic movements were observed. The animation was similar to the traffic flow that was observed from the field as well.
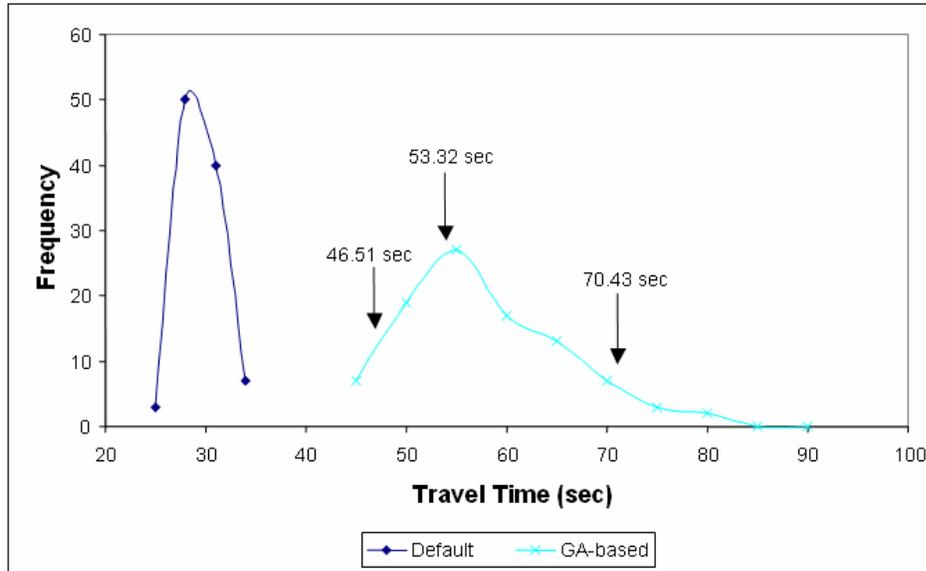
Figure 42. Comparison of Site 15 Travel Time with CORSIM

Table 23. Two Parameter Sets for Site 15 with CORSIM

Site15-CORSIM	Default	Calibrated
Link mean free flow speed (mph)	45	30
Mean queue discharge headway (sec)	1.8	2.8
Mean start-up lost time (sec)	2.0	2.8
Left turn jumper probability (%)	38	21
Left turn speed (ft/sec)	22	19
Right turn speed (ft/sec)	13	23
Left turn lagging within 2 seconds (%)	50	47
Left turn lagging within 2-4 seconds (%)	15	12
Amber interval response (fpss) Index	1	2
Gap distribution for left turns (sec) Index	1	1
Gap distribution for right turns (sec) Index	1	2
Distribution of free flow speed by driver type (%) Index	1	2
Start-up lost time distribution (%) Index	1	2
Discharge headway distribution (%) Index	1	1
Avg. Travel Time (sec)	28.2	58.4

7.3.6. Validation

Three calibrated CORSIM parameter sets were evaluated using field data of the fourth day not used in the calibration. 100 runs were made to be compared with validation data. Distributions of three parameter sets were compared in Figure 43. Similar to the calibration result, LHD-based and GA-based parameters could achieve simulated distributions with mean values close to the field data while the default parameters produced shorter travel times. Therefore, the calibrated parameter set can be used for further transportation analysis work without any further fine-tuning process.
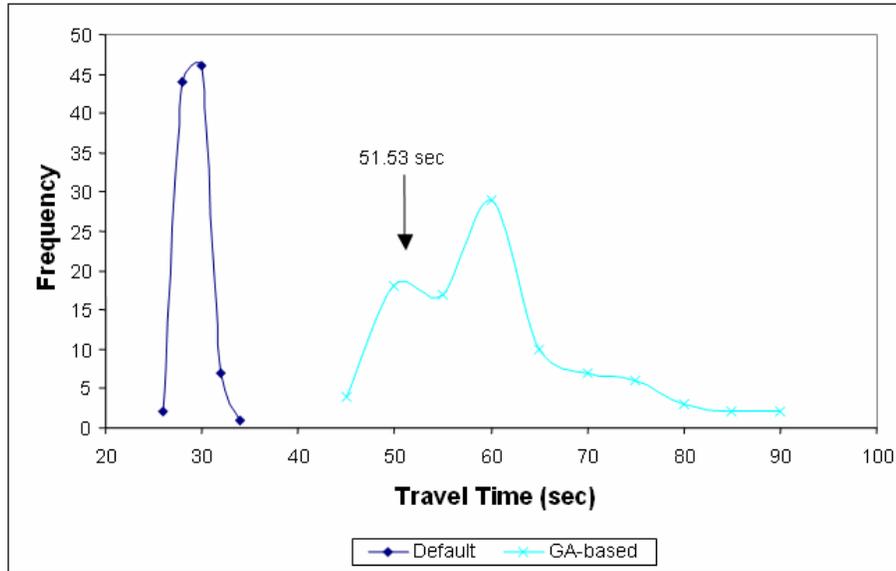
Figure 43. Validation of CORSIM Using Site 15

7.4. VISSIM Calibration

7.4.1. Identification of Calibration Parameters

VISSIM provides numerous calibration parameters that can be modified and those parameters are categorized with its own characteristics. In the Site 15 case, the following types of calibration parameters are selected:
- Basic calibration parameters: These are global parameters that affect the whole network, which is fundamental to this model.
- Wiedemann 74 model: According to the VISSIM user manual, the Wiedemann 74 model is appropriate for an urban network. So, all three parameters in the Wiedemann 74 model have been included.
- Priority Rule: As mentioned previously, left-turn at this intersection is permitted but not protected. Therefore, it is essential to consider priority rule parameters.
- Desired Speed Distributions: Of course, the speed limit for this intersection is given; however, not all the vehicles move with the speed limit and the desired speed varies with the driver's characteristics.

The following is the list of parameters and acceptable ranges based on the user manual and engineer's judgment.
1) Simulation resolution (Time steps/Sim. sec.): 1 ~ 9
2) Number of observed preceding vehicles: 1 ~ 4
3) Average standstill distance (meter): 1 ~ 5
4) Saturation flow rate (1756, 1800, 1846, 1895, and 1946 veh/h)
   - Additive part of desired safety distance*: 2.0 ~ 2.5
   - Multiple part of desired safety distance*: 3.0 ~ 3.7
5) Priority rules – minimum headway (meter): 5 ~ 20

68

6) Priority rules – minimum gap time (second): 3 ~ 6
7) Desired speed distribution (mph): 40~50, 30~40, 20~30
* The values were based on the table (Flow rate, Throughs, 25 second green) in VISSIM Manual 5.2.6.

7.4.2. Experimental Design for Calibration

Two hundred parameter combinations were generated within the initial parameter ranges determined by using the Latin Hypercube Sampling method.

Five random seeded runs were conducted in VISSIM for each of the 200 cases, for a total of 1000 runs. The reason that only five runs were made for this VISSIM calibration is that VISSIM output usually shows narrower variability than other simulation models. As was done for CORSIM, the average travel time was recorded for each of the 1000 runs. The results from the five multiple runs were then averaged to represent each of the 200 parameter sets.

7.4.3. Feasibility Test

As was done for the CORSIM calibration procedure, a feasibility test was conducted to check whether the simulated outputs based on the current parameter ranges could capture the field data, and to identify the key calibration parameters in VISSIM.

The travel time histograms of the 200 cases are shown in Figure 44. The average field travel time data (56.75 seconds) fall outside of the travel time output distribution. It indicates that the selected parameters and their ranges could not reproduce the field conditions and need adjustments. In order to shift the simulation output distribution to the right side to capture the field condition, key parameters that most affected the results need to be identified. Then, the initial ranges of key parameters are expanded.
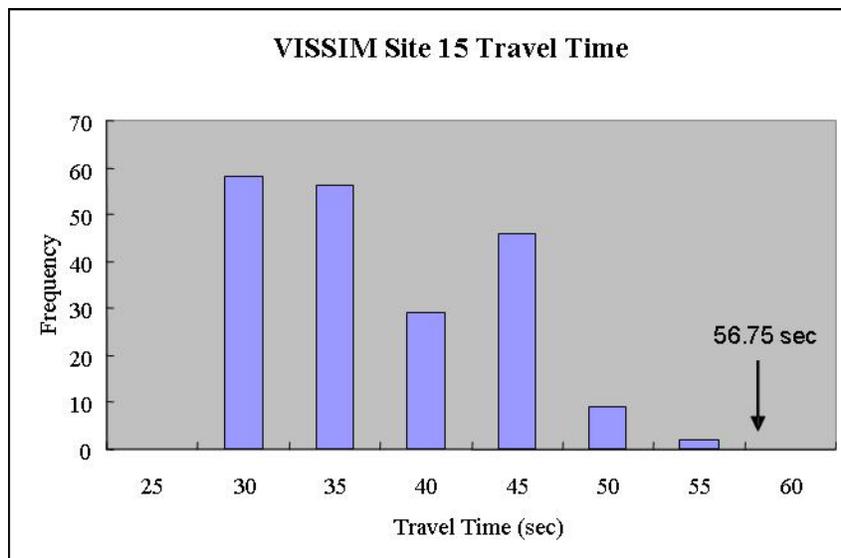

Figure 44. Feasibility Test Results for Site 15 with VISSIM

X-Y plots and statistical test (ANOVA) results are used to identify key calibration parameters. X-Y plots for each parameter versus travel time from the simulation are shown in Figure 45. As seen in these figures, the following calibration parameters showed a relationship with travel time output (performance measure).

- Minimum gap (f)
- Desired speed distribution (g)

According to those X-Y plots, the minimum gap time should be increased and the average desired speed value should be decreased.
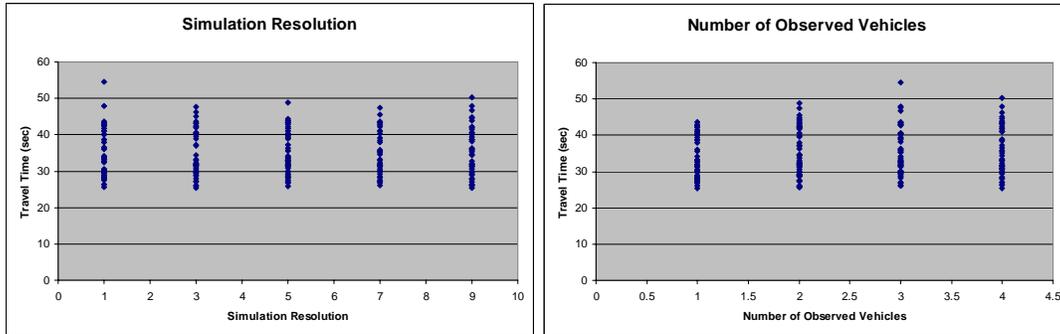
(a) Simulation Resolution   (b) Number of Observed Preceding Vehicles
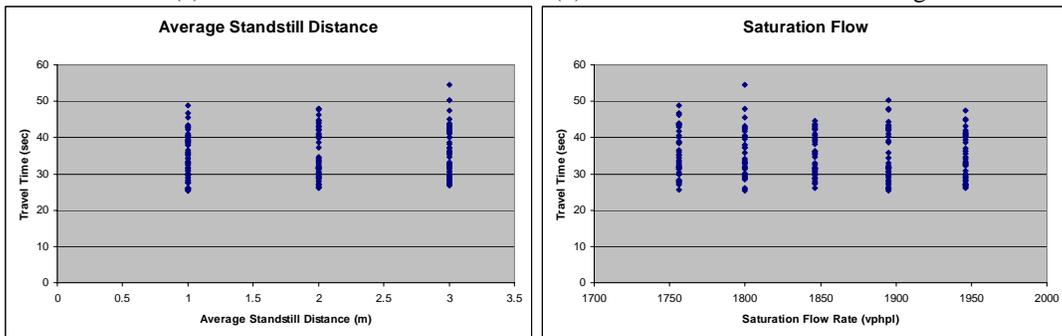
(c) Average Standstill Distance   (d) Saturation Flow Rate
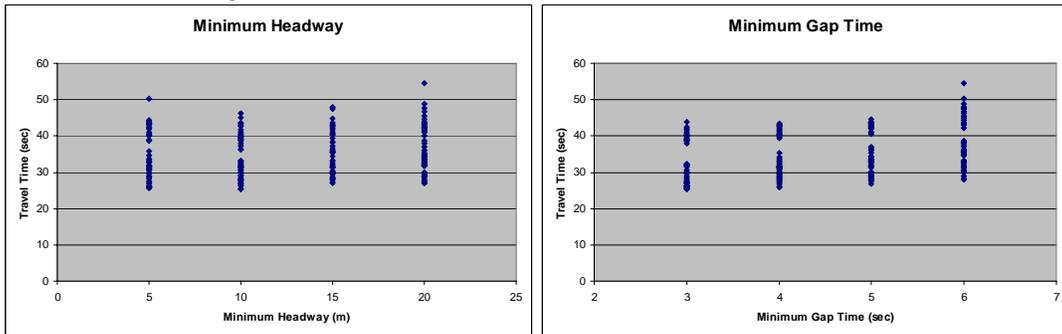
(e) Priority Rules- Minimum Headway Response   (f) Priority Rules- Minimum Gap Time
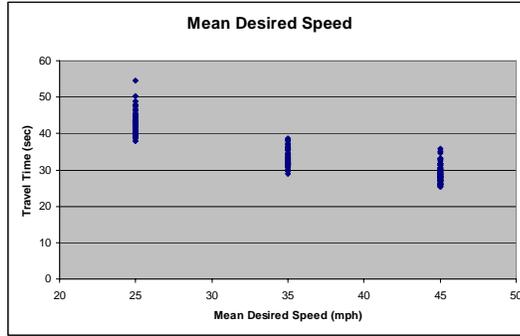
(g) Desired Speed Distribution

Figure 45. X-Y plots of Travel Time vs. Calibration Parameters—CORSIM

It is more rigorous to identify the critical parameters through statistical analysis. The values of each parameter were categorized into several groups and indexed because ANOVA cannot analyze continuous values. Table 24 shows ANOVA results, listing p-values. Since the p-value for this case was also 0.05, if the p-value is less than 0.05, that specific parameter can be considered as a key parameter. Also, if some parameters only show a difference with statistical analysis, joint effect with another key parameter needs to be considered. As shown in the table, minimum gap time and desired speed distribution showed a value less than 0.05 and can be considered key parameters, which is identical with the X-Y plot analysis result. Therefore, the ranges of the two calibration parameters need to be adjusted.

Table 24. ANOVA Results for Site 15 with VISSIM

Site15 – VISSIM	Significance value (p value)
Simulation Resolution	0.989
No. of Preceding Vehicles	0.336
Ave. Standstill Distance	0.369
Saturation Flow	0.910
Min. Headway (meter)	0.220
Min. Gap Time (sec)	**0.000**
Desired Speed Distribution	**0.000**

* Significance value is less than 0.05.

7.4.4. Parameter Range Adjustment

Based on the feasibility test, the desired speed distribution and the minimum gap time were identified as critical parameters and considered expansion from the current ranges. However, the upper bounds of these parameters have already reached the possible maximum that can be considered as a realistic value in this case. There was a possibility that some other factors that might impact the result were not yet identified.

71

Therefore, additional tools like HCM or field data needed to be used to help the clarification of the ranges of some parameters, such as saturation flow and field speed within the intersection. Since the simulation travel time tended to be shorter than the field travel time, it could be possible that the saturation flow in VISSIM or the mean of desired speed distribution was higher than that of the true field value.

Check Speed within the Intersection

Although the posted speed limit of the field was 45 mph, the actual speed within the intersection could be lower than 45 mph due to many different reasons, such as poor sight distance, permitted left-turns, narrow intersection, and other factors like gas stations located around 500 ft upstream of the stop line on the southbound approach.   It is more reasonable to use the field average speed to set the vehicle's desired speed.

Field speed data were retrieved as the tail of the vehicle crossed the stop bar on the intersection by reviewing the videotape. The first few vehicles in the queue were not considered because they were still accelerating and cannot be considered as they were cruising at the driver's desired speed. Figure 46 shows the box plot of field speed within the intersection. This confirms that an average speed was around 35 mph and the range should be 25~45 mph. The modified ranges were: 30-40 mph, 32.5-37.5 mph, and 27.5-42.5 mph.
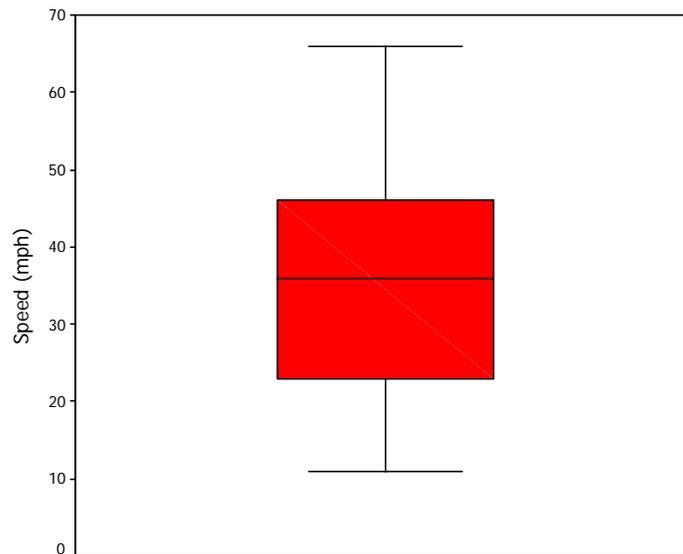

Figure 46. Box Plot of Field Speed within Intersection

Check Saturation Flow

Saturation flow rate is an important factor to determine the intersection capability, which in turn affects the vehicle travel time. It can be obtained from the following three methods:
-   Field data
-   HCM procedure

-   VISSIM simulation

By obtaining saturation flow rate using three different methods, saturation flow rate in a VISSIM model can be verified, whether biased or not.

Calculation of queue discharge headway from the field data
The field saturation flow rate of the intersection was obtained by reviewing the videotape recorded on June 5. First, the time stamps of the third and the last vehicle in the queue passing the stop line in each cycle were recorded. Then the average discharge headway and the saturation flow rate were calculated by using the following equation:

$$\text{Saturation Flow Rate} = \frac{3600}{(t_3 - t_n)/(n-3)}$$

The average headway was 3.13 sec, which resulted in a saturation flow of 1,149 vehicles/hour/lane (vphpl). Generally, field data showed that the southbound approach has relatively low saturation flow rate and should be the main reason accounting for the longer travel times in the previous simulations.

Calculation using HCM procedure
A saturation flow rate can also be computed using the HCM procedure. The ideal saturation flow rate, which is usually 1,900 vphpl, can be adjusted for the prevailing conditions to obtain the saturation flow for the lane group being considered. The adjustment is made by introducing factors that correct for number of lanes, lane width, heavy vehicle percentage, right and left turns, and so on. The adjusted saturation flow rates are computed and summarized in Table 25. The result also confirms that the southbound approach has a low saturation flow.

Table 25. Saturation Flow Rate Using HCM Procedure (Park and Qi, 2004)

6/5	EB			WB			NB			SB		
	L	T	R	L	T	R	L	T	R	L	T	R
Volume	88	98	44	45	77	84	32	172	29	192	300	93
Proportion of LT or RT	0.4		0.2	0.2		0.4	0.1		0.1	0.33		0.2
Basic saturation flow		1900			1900			1900			1900	
Number of lanes		1			1			1			1	
Land width adjustment		1			1			1			1	
Heavy vehicle adjustment		0.962			0.98			0.943			0.952	
Grade adjustment		1			1			1			1	
Parking adjustment		1			1			1			1	
Bus blockage adjustment		1			1			1			1	
Area type adjustment		1			1			1			1	
Lane utilization adjustment		1			1			1			1	
Left-turn adjustment factor		0.791			0.807			0.787			0.726	
Right-turn adjustment factor		1			1			1			1	
Adjusted saturation flow		1445.79			1502.634			1410.068			1313.189	

Calculation of queue discharge headway from VISSIM animation
In VISSIM, two parameters affect saturation flow rate:

- Additive part of desired safety distance
- Multiple part of desired safety distance

The VISSIM manual provides four tables to roughly estimate saturation flow rate based on these two parameters under certain signal timings. However, the relationship between these two parameters and saturation flow rate in VISSIM is not clearly stated in the manual. An accurate way is to record vehicle discharged headways in VISSIM and calculate saturation flow rate. Because of the low saturation flow rate obtained from field data and the HCM procedure, it was envisioned that the current ranges for the additive part of desired safety distance and the multiple part of desired safety distance were not appropriate and overstated. Therefore, the ranges of these two parameters were expanded to have a lower saturation flow rate. The new ranges of the additive part of desired safety distance and the multiple part of desired safety distance were 1.0-5.0 and 1.0-6.0, respectively.

To test the capability of VISSIM to achieve the desired saturation flow rate, two cases with simulation results close to the field data were selected. The VISSIM animations were watched and queue discharge headways were recorded for each cycle. Table 26 shows the observed queue discharge headways and saturation flow rates in VISSIM. The result shows that VISSIM can achieve comparable saturation flow rates to the field condition.

Table 26. Comparison of Saturation Flow

Comparison	Queue Discharge Headway		Saturation Flow (veh/h/l)	Travel Time (sec)
	Ave. (sec)	Stdev (sec)		
Field	3.13	0.82	1149	55.45
VISSIM (Case 1)	2.96	0.69	1216	53.04
VISSIM (Case 2)	3.03	0.83	1190	54.79

Based on the speed and saturation flow rate conditions, the parameter set is modified as follows:
1) Simulation resolution (Time steps/Sim. Sec): 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 headway (meter): 5 – 20
7) Priority rules – minimum gap time (second): 3 – 6
8) Desired speed distribution (mph): 30-40, 32.5-37.5, 27.5-42.5

With the new parameter ranges, another 200 cases were generated using the Latin Hypercube Design method. The new simulated distribution is shown in Figure 47, which covers the field data and indicates that the new parameter ranges were able to capture the field condition. The parameter set with the best performance was selected as the LHD-based parameters in the following evaluation:
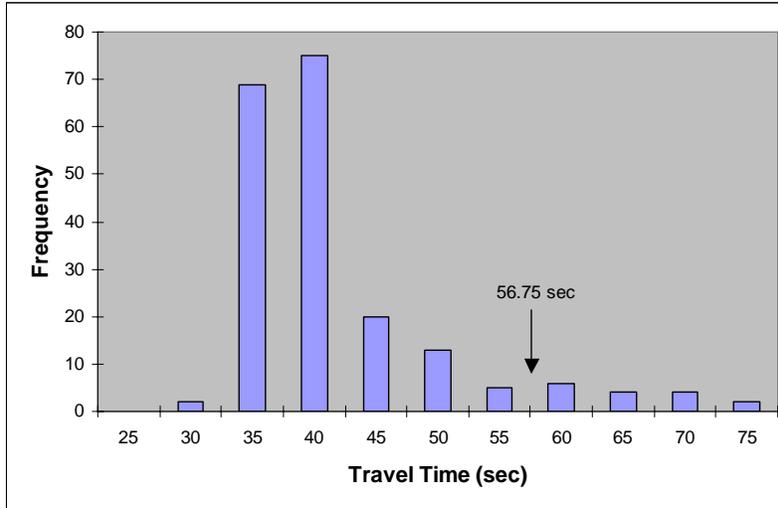
Figure 47. Feasibility Test Results for Site 15 with VISSIM

7.4.5. Parameter Calibration Using Genetic Algorithm

A genetic algorithm was integrated with the VISSIM model to calibrate parameters as well. Figure 48 shows the convergence as the number of generations increases. The results were based on 7 generations of GA evolution and a population size of 20 in each generation. As shown in Figure 48, as the number of generations increases, the fitness value gets smaller, which means that the simulation output gets close to the field-measured travel time.
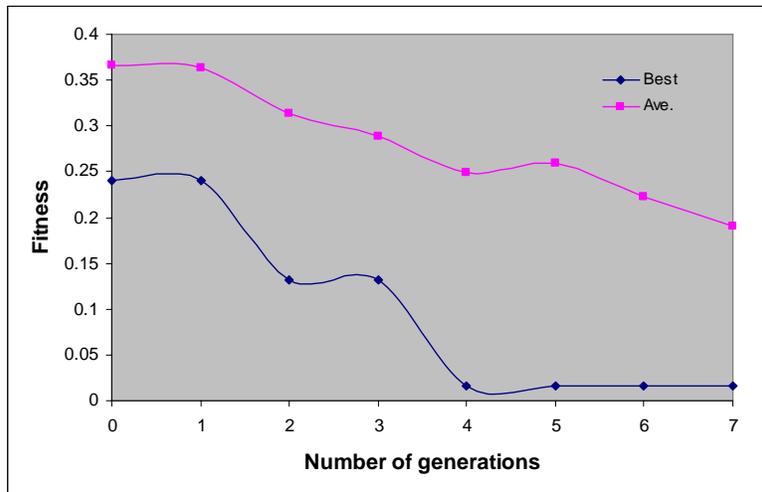
Figure 48. Convergence of Fitness Value with Generation
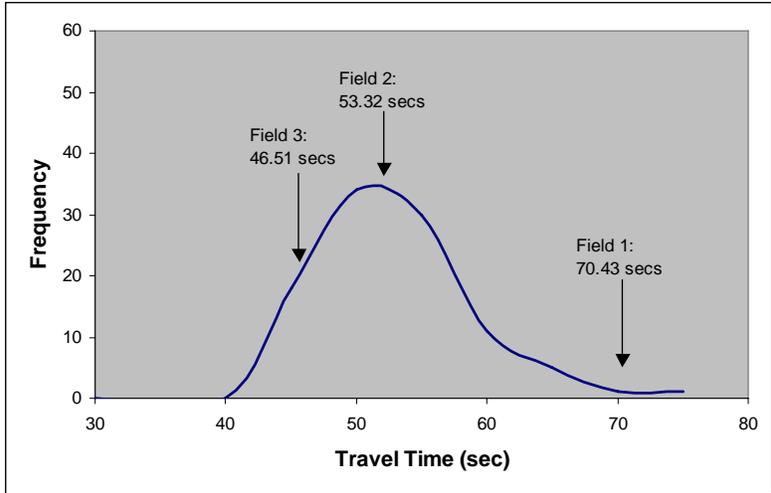
Figure 49. Time Distribution of GA-based Parameter Set

With the calibrated parameter set, VISSIM was run 100 times and an average travel time was recorded for each run. The resulting distribution of travel time is shown in Figure 49. Field travel times from the three days are also shown in the figure and they all fall within the simulated distribution.

To test the consistency of results from GA optimization, the GA process was repeated two more times with 10 generations and 20 populations. Naturally increasing the number of generations or populations would help to reach a better fitness value, but it requires more time. The convergence results are shown in Figures 50 and 51. The parameter set with the best fitness value was selected to represent the local traffic characteristics. These parameter values are summarized in Table 28.
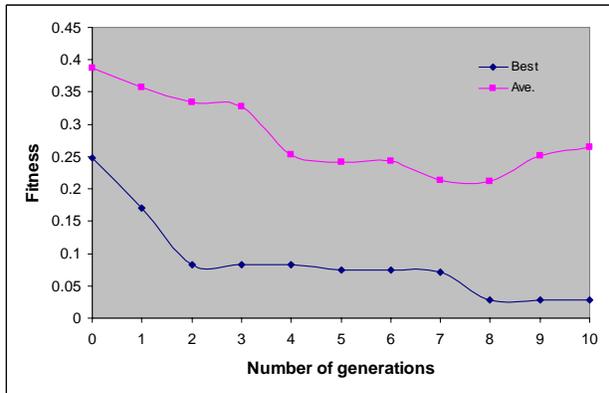
Figure 50. Convergence of Fitness Value with Generation
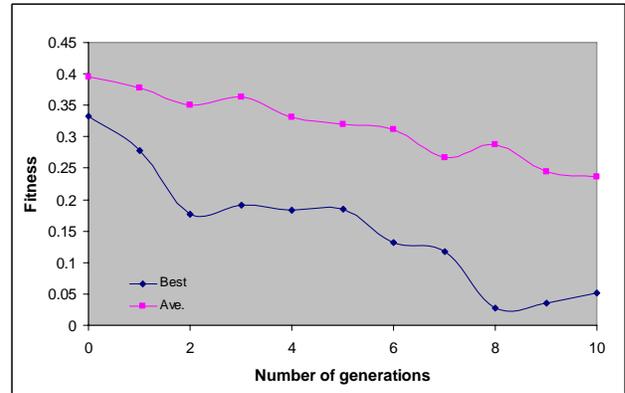
Figure 51. Convergence of Fitness Value with Generation

7.4.6. Evaluation of the Parameter Sets

This section presents the comparison of 100 VISSIM simulation results based on default parameters, LHD-based parameters, and GA-based parameters. The parameter values for each set are listed in Table 27.

Evaluation of default parameters was performed by making multiple VISSIM runs without changing any parameter values. A comparison of the uncalibrated VISSIM models (default parameters), and the calibrated VISSIM model (GA-based parameters) shows the importance of calibration for microscopic simulation models. Travel times along the subject link are compared in Figure 52. As seen in the figure, three average field travel times all fall within the distributions of simulation results using the calibrated model. The uncalibrated models generate much shorter travel times than those observed in the field.

Animations of each parameter set were also viewed in order to determine whether the animations were realistic or unrealistic. For the calibrated parameters, the animations at several travel time percentiles of the distribution were found to be acceptable. For the default parameters, almost all vehicles passed the intersection without waiting, which was not realistic.

Table 27. Two Parameter Sets for Site 15 with VISSIM

| Site15 – VISSIM	Default	Calibrated
Simulation resolution	5	6
Number of observed preceding vehicles	2	4
Maximum look ahead distance (meter)	250.00	215.15
Average standstill distance (meter)	2.00	3.85
Additive part of desired safety distance	3.00	5.0
Multiple part of desired safety distance	3.00	5.3
Priority rules – minimum headway (meter)	5.0	20.0
Priority rules – minimum gap time (second)	3.0	4.0
Desired speed distribution (mph)	Car: 40-50 HV: 30-40	Car: 27.5-42.5 HV: 15.5-18.6
Ave. travel time (second)	23.40	50.33
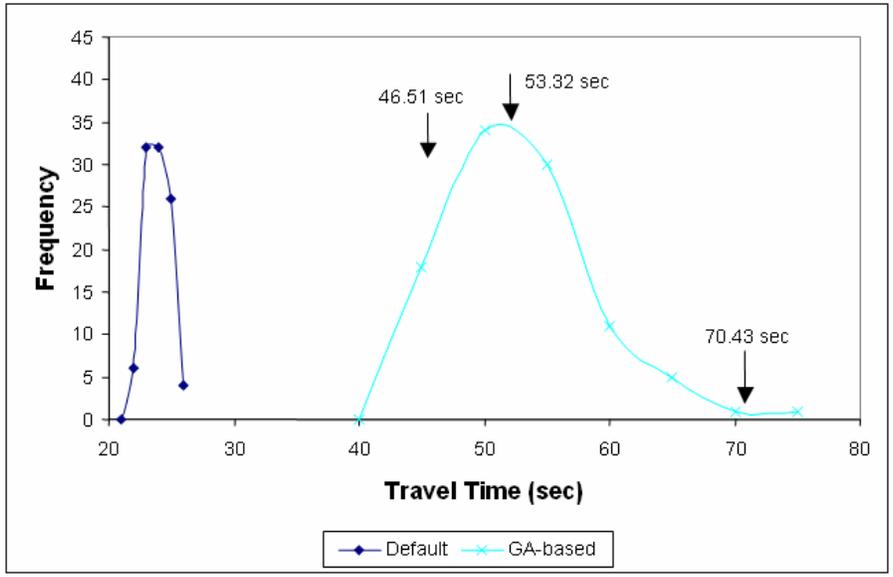
Figure 52. Comparison of Site 15 Travel Time with VISSIM

7.4.7. Validation

The "calibrated" models were then evaluated with a new set of field data under untried conditions, including the input volumes, traffic composition, and other required data. The performance measure was again average travel time on the southbound approach, but on a different day. It was used for evaluation of the four parameter sets in Table 28. The comparison result is shown in Figure 53. Field average travel time, 51.53 sec, was about the mean of the simulated distributions of the calibrated models. And once again, the uncalibrated models generated shorter travel times.
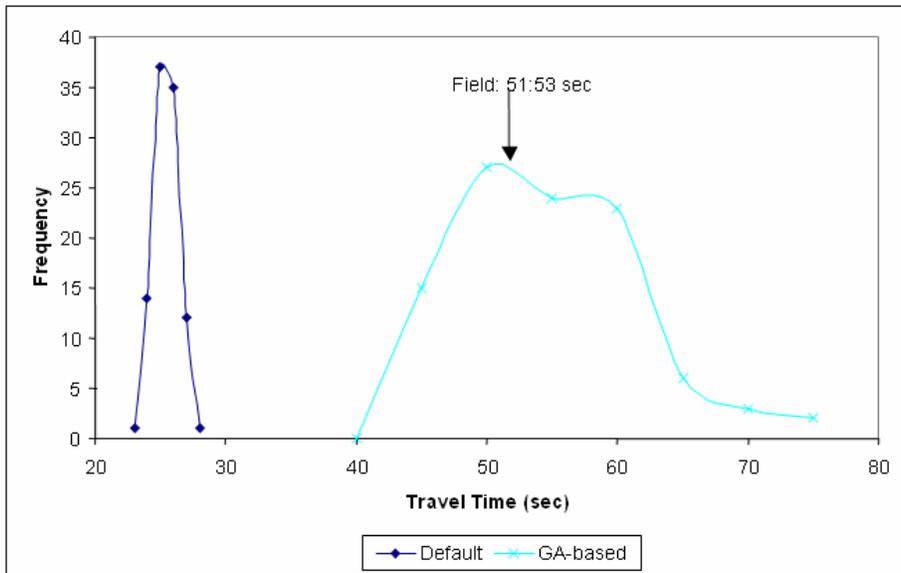
Figure 53. Validation of VISSIM Using Site 15

Thus, it can be concluded that the simulation model with calibrated parameter sets reproduces the field condition effectively.

Chapter 8
Urban Arterial Network

In this chapter:
- Case study of calibration and validation with urban network

---

8.1. Site Description

This test site is an urban network with 4 signalized intersections along Route 29 in Charlottesville, Virginia. Route 29 is one of the major routes in central Virginia because it serves most of the traffic volume from central Virginia to the Washington, D.C., area.

The test site is located at Emmet St. between Hydraulic Rd. and Barracks Rd., and includes a total of four intersections. Among them, two intersections at Emmet St. and Hydraulic Rd. and Emmet St. and Barracks Rd. work as metering intersections to consider boundary effects and only the number of vehicles that enter the network was considered. Another point that has to be considered is the existence of on- and off-ramps to Route 250 that connects Waynesboro to Richmond, Virginia. Because of this characteristic, the geometry of this network is more complex than other networks.

An aerial photo of the study site is shown in Figures 54 and 55. Figure 54 presents the geometry of the whole study site and Figure 55 presents 2 intermediate intersections that are a major interest of this study. Black circles in Figure 55 present signalized intersections.

Figure 54. Emmet Street, Charlottesville, Virginia

Figure 55. Detailed Geometry

During the preliminary site visit as well as data collection, a unique phenomenon was observed by the research team. As mentioned, Route 250 intersects with Route 29 within this network and Route 250 shares its roadway with the Route 29 bypass. Because of its high speed and capacity, a large percentage of vehicles on Route 29 south try to diverge to Route 250 / Route 29 bypass through a one-lane ramp and it forms quite a long queue on the right most-side lane that requires additional attention when building a network. Figure 56 shows the real field condition.
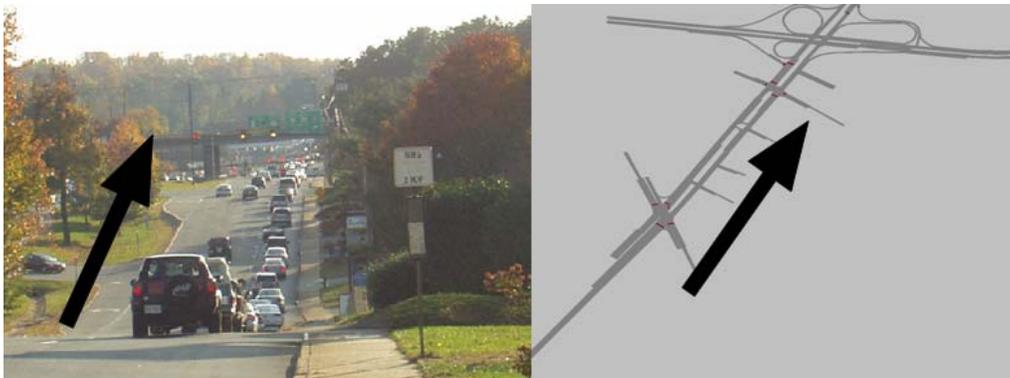
Figure 56. Long Queue in Rightmost Lane

8.2. Data Collection

8.2.1. Required Types of Data

81

The data needed for this case study can be categorized into two major types. First, fundamental traffic data that required building a simulation model network such as traffic counts, heavy vehicle percentage, geometrical characteristics, detector locations, signal timing plan, speed limits, etc. were collected. Then the types of data required are calibration and validation. In this case study, travel time data for two different sections were selected to be used as calibration data and maximum queue length data at two different locations are used as validation data. In the case of maximum queue length data, the unit that has been used from VISSIM is different from the unit that has been collected from the field. VISSIM provides queue length data in length, however, the way that has been collected from the field was by number of vehicles. Therefore, it is the output from VISSIM needed to be converted into the number of vehicles format.

8.2.2. Data Collection

The geometric attributes pertaining to the test network were obtained from SYNCHRO and CORSIM files, developed by VDOT. However, as these networks were developed in 2001, the changes in geometry such as lengths of left- and right-turn lanes or locations of bus-stops were updated using aerial photos and site visits. The current signal timing plans of Emmet St. (Rte. 29) at Angus Rd., Morton Rd. and Barracks Rd, which is in the jurisdiction of the City of Charlottesville, was provided by the traffic engineer in the City of Charlottesville. The signal timing plan of Rte. 29 at Hydraulic Rd. was obtained from VDOT, who manages the intersection. The traffic-related data were collected directly from the field using both manual counting and video recording. The data collection was conducted on a normal weekday, Tuesday through Friday, July 11, 2001, 2:45 p.m. and 4:15 p.m. A group of 17 people performed simultaneous manual counts along the test site as shown in Figure 57.

Traffic volume and heavy vehicle percentage were measured from every intersection and ramp by using an electronic data collection device (TDC-12 Traffic Data Collector) as well as a data collection sheet. Due to the large amount of vehicles entering Route 29 South at Hydraulic Rd (Northern entrance to the network), traffic counts and other required data were collected by using AUTOSCOPE in Smart Travel Van (STV). Four video cameras with synchronized clocks were positioned on the entry point and exit point for each section (2 for location A, 1 for location B and C) and recorded a license plate number. Another two surveyors were located at two approaches at one intersection (QL 1, QL 2) to collect the maximum queue length data by counting the number of stopped vehicles at the end of each red phase.
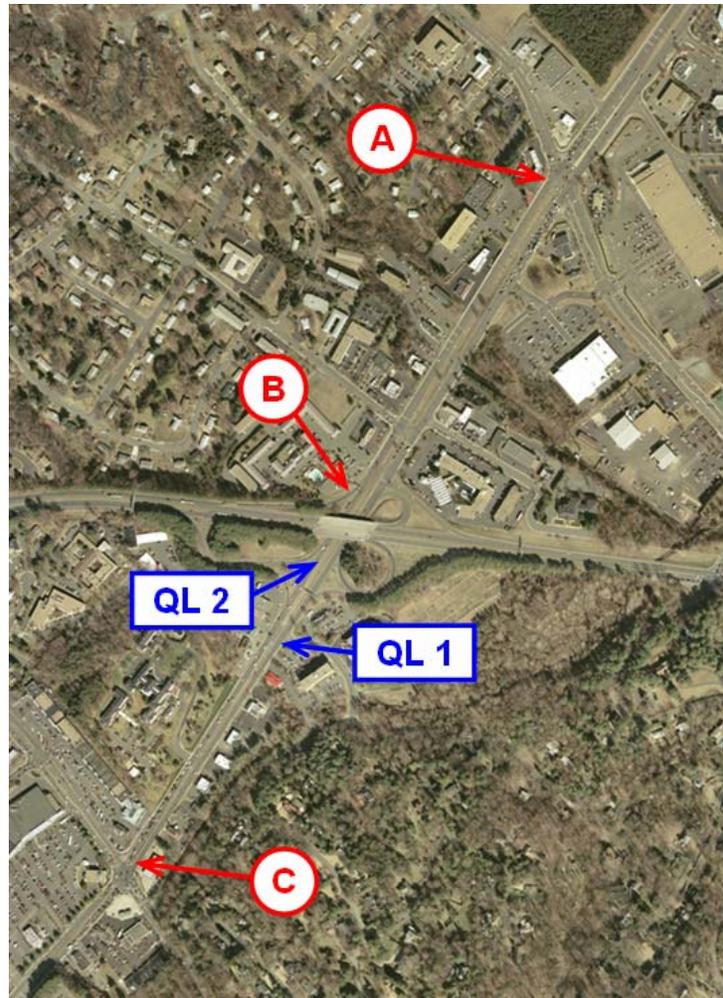
Figure 57. Data Collection Location Map—Charlottesville

8.2.3. Data Reduction

As mentioned previously, traffic volume data of each intersection were collected using an electronic data collection device (TDC-12 Traffic Data Collector). After collecting all data collection devices from surveyors, they were connected to the computer and the traffic volume data in the devices were automatically transferred to computer in Excel format. Also aerial photos were obtained from the City of Charlottesville to be used as a background image for the network building process in all different simulation models.

Video- tapes on different days recorded from three different locations were reviewed and the license plate number of each vehicle was manually recorded and matched to extract the travel time of two sections. Each travel time was determined by subtracting the time when a vehicle passed the subject entry point from the time when the vehicle passed the subject exit point. License plate numbers and times recorded were manually matched. Maximum queue length data were collected manually by writing down the maximum queue length at the end of every signal cycle. Data sheets were manually reviewed and typed in an Excel worksheet.

8.2.4. Data Summary

Traffic data such as volume, heavy vehicle percentage, details on geometry, and signal timing plans are not provided in this case study because of limited space. The only types of data that were required for the calibration and validation procedure, such as travel time and maximum queue length, are described in this part.

Travel Times

The travel times of the following two sections were used as a performance measure for the calibration process:
– Section 1: From Hydraulic Rd.(A) to Barracks Rd.(C)
– Section 2: From Hydraulic Rd.(A) to Ramp 1.(B)

It should be noted that the travel time of section 1 depicts the travel times of vehicles using the southbound leftmost lane on section 1, and that of section 2 is the travel times of vehicles traveling in the southbound rightmost lane on the section. The travel times collected from the two sections are presented in Table 28 and Figure 58.

Table 28. Statistics of Field Travel Time Data

Section	Date	Number of Vehicles	Travel Time (sec)		
			Mean	Standard Deviation	Average
Section 1	7/12/05	36	147.9	26.0	150.3
	7/13/05	29	152.7	28.0	
	7/14/05	51	156.9	27.2	156.9
Section 2	7/12/05	135	132.9	37.5	116.3
	7/13/05	157	99.7	24.4	
	7/14/05	264	138.2	29.8	138.2
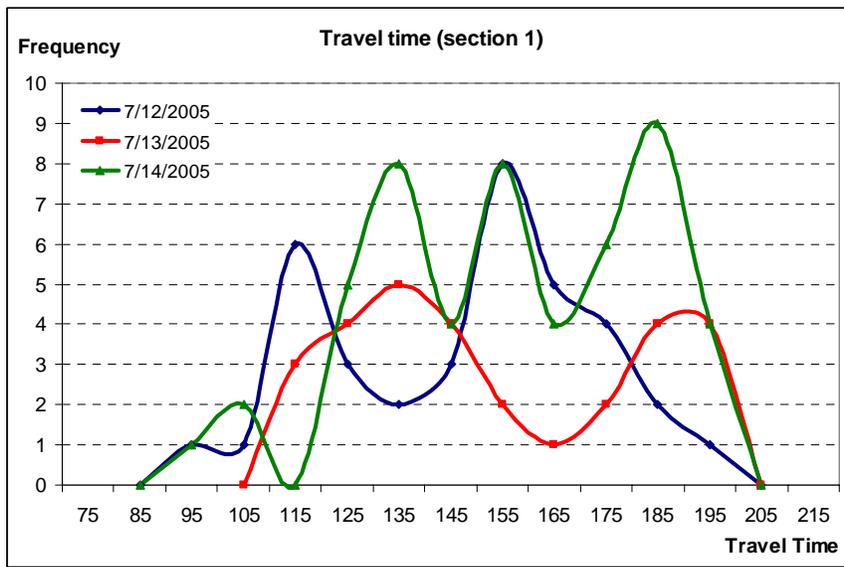
(a) Field Travel Time –Section 1
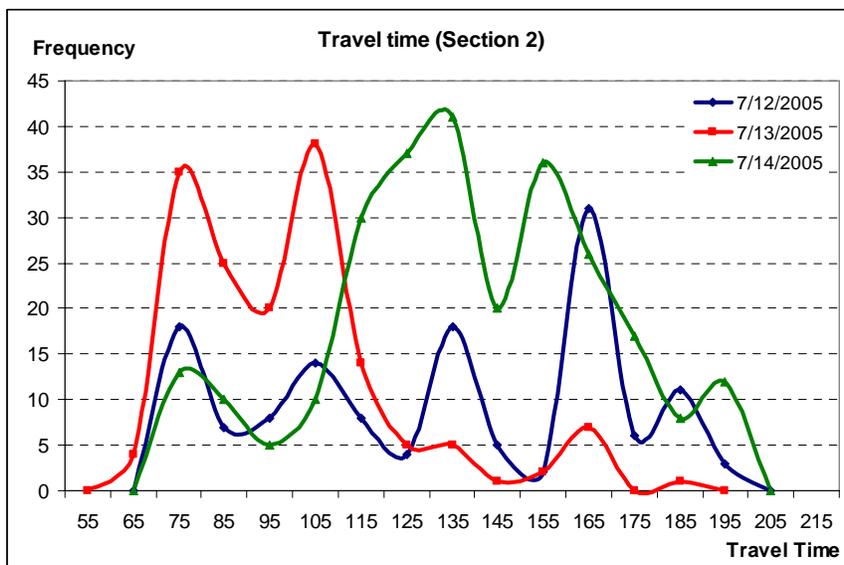
(b) Field Travel Time –Section 2

Figure 58. Field Travel Times—Charlottesville Network

Maximum Queue Length

Maximum queue lengths were measured at the northbound approach of Morton Rd. (QL 1) and the southbound approach of its overlap intersection (QL 2). The queue lengths were determined by counting the number of vehicles in a queue at the end of the red time for each cycle during the data collection period. The longest queue length was designated as the maximum queue length here. The maximum queue length in the two locations was used for validation. Table 29 shows the summary of surveyed queue lengths at the two locations.

Table 29. Queue Length Measured

Location	Date	Queue Length, vehicles		
		Average	Standard Deviation	Max
Location 1	7/12/05	13	7	24
	7/13/05	10	5	24
	7/14/05	13	7	26
Location 2	7/12/05	9	4	15
	7/13/05	7	4	16
	7/14/05	11	5	23

8.3. CORSIM Calibration

8.3.1. Efforts on Network Building

Due to the complexity of this network, many different problems were encountered during the initial test run step. Some of the critical problems or errors that occurred are listed and described in this part. Figure 59 shows the network of Charlottesville in CORSIM.
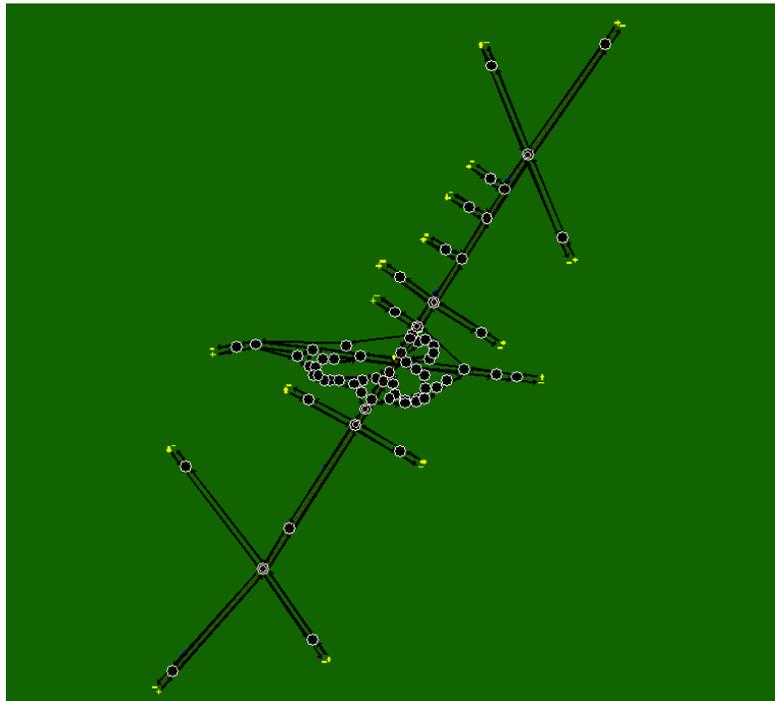
Figure 59. Charlottesville Network—CORSIM

Unrealistic Blocking Phenomenon at Ramp 1

As briefly mentioned in the site description part, CORSIM randomly assigns each vehicle's next destination as they enter the corresponding link. So, if the spacing of the route decision points (e.g., intersections, ramps, etc.) is too close, some vehicles cannot change their lane to the destination lanes due to those narrow spacings. When this kind of vehicle appears on the network, it occasionally blocks the lane and hinders the traffic flow and Figures 60 and 61 show an example of the appearance of blocking vehicles in the network.
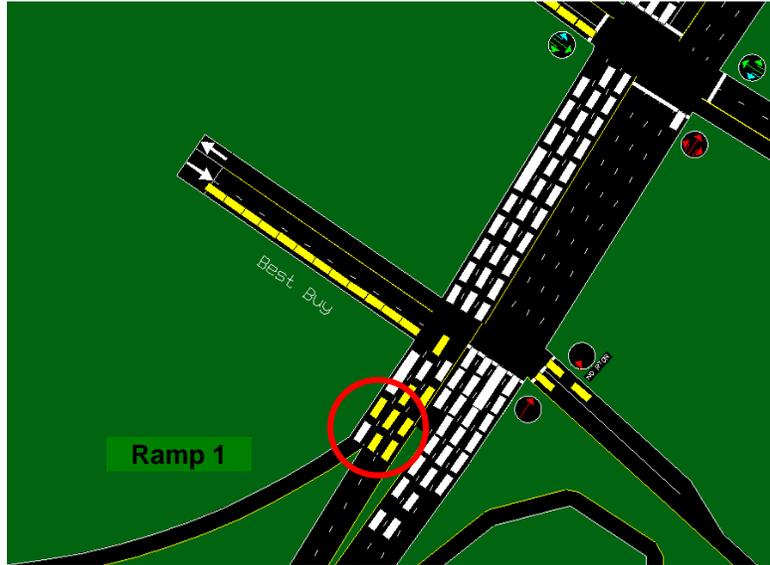
Figure 60. Blocking Vehicle Example 1
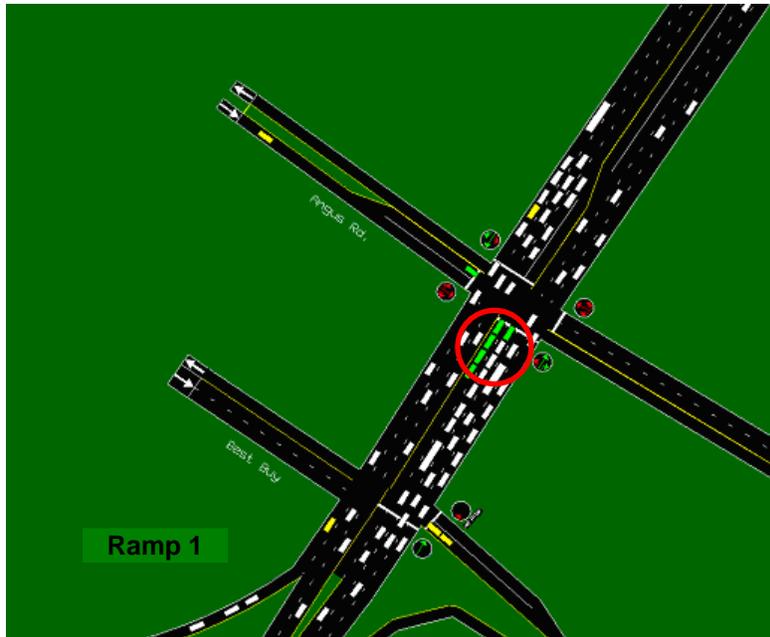
Figure 61. Blocking Vehicle Example 2

The unique traffic pattern between Hydraulic Rd. and the ramp to Rte. 250 that causes the blocking vehicle appearance can be addressed with the following two methods:

- Use of Origin-Destination (O-D) assignment in an interchange
- Use of turning percentage
The two methods were implemented alternatively, and then the latter was finally selected.

Use of Origin-Destination (O-D) Assignment Using an Interchange

By designating a set of links at an interchange, CORSIM can assign an O-D matrix in the interchange to represent a specific traffic pattern. So the southbound links of Emmet St. from Hydraulic Rd. and the ramp 1 were coded as interchanges. For the O-D matrix required by the inclusion of interchanges, QUEENSOD method (Van Aerde et al., 2003) was implemented and followed by a manual fine-tuning process. However, this application produced the following problems:
- Lack of backed-up vehicles on the rightmost lane than real field
- Vehicle still blocked another lane

Therefore, this method was considered unacceptable for this case study.

Use of Turning Percentage

As a default method for assigning traffic on a network, CORSIM uses turning percentages. In other words, CORSIM assigns turning directions to vehicles entering each link randomly. This random assignment has the following limitations in replicating the real traffic patterns:
- No queues in the rightmost lane that enters to ramp 1.
- Relatively shorter travel time in Section 2 than that of real field.

However, this method in the CORSIM network building can avoid the abnormal behaviors illustrated in the first method.

The problems described that took place on the base case network are the primary reason why network calibration and validation processes should be conducted to confirm the microscopic simulation model to represent real traffic patterns.

Unrealistic Turning Movements at Ramp 7

In addition, a lot of traffic from Ramp 7 made left-turns at Angus Rd. An unrealistic turning movement scene was captured and presented in Figure 54. However the method using the turning percentage cannot reflect this traffic pattern because CORSIM randomly assigns turning movements, including left-turns, right-turns or through movements, to vehicles in a link after they enter the link. Therefore, by assigning links located in the area as an interchange, CORSIM could replicate the actual field traffic pattern. Unlike what happened in Figure 62, no abnormal traffic patterns occurred by implementing the O-D assignment in the interchange.
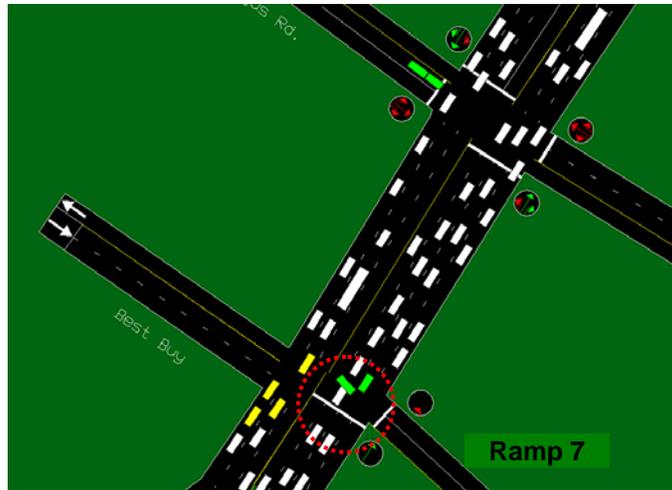
Figure 62. Example of Abnormal Turning Vehicles

In addition, Ramp 7 consists of two lanes. The vehicles turning right were using the right lane whereas the vehicles that wanted to make left-turns at the intersection of Emmet St. and Angus Rd. were using the left lane according to a road sign installed in the middle of the ramp. In practice, the vehicles sitting on the left lane could not freely enter Emmet St. due to high conflict traffic volumes on Emmet St., so they usually waited at the stop bar until the signal head turned to green indication. However, in the simulation, the vehicles on the left lane aggressively entered Emmet St. as illustrated in Figure 54 (refer to green vehicles at the circle). In order to avoid these abnormal turning vehicles in the simulation, right-turn-on-red was placed on the approach from Ramp 7.

8.3.2. Evaluation with Default Parameter Set

First of all, 100 simulation runs with the default calibration parameter set were conducted to check whether the simulated results based on the default parameter set could capture the field condition or not. The averaged result of 100 runs with a default calibration parameter set is tabulated in Table 30.

Table 30. Performance Measure of CORSIM Network with Default Calibration Parameters

Performance Measure	Simulation Results	Actual Values
Travel time of Section 1, seconds	137.7	150.3
Travel time of Section 2, seconds	75.6	116.3
Maximum queue length at SB approach at Morton Rd, vehicles	17	16
Maximum queue length at NB approach at Morton Rd, vehicles	16	24

As shown in the table, the model with default calibration parameter sets could not replicate the field traffic condition in terms of performance measures (Travel time, and Maximum queue length). By and large, travel time and maximum queue length data from simulation output showed consistently lower value than the value collected from the field. For example, travel time of section 1 had a difference of 12.6 sec and needed to be increased.

89

8.3.3. Identification of Calibration Parameters

Due to the complexity of this network and the existence of a long queue at the on-ramp to Rte. 250 west, many different calibration parameters were selected to be calibrated. The lane-changing parameters are included in this calibration to replicate the long queue on Section 1, to adjust for the aggressiveness and/or conservativeness of drivers' behaviors.

However, two calibration parameter settings were predefined and excluded from the set of calibration parameters. First, the spillback probability (Record Type 141 in CORSIM) controls the degree of willingness of a vehicle to join a spillback, when the vehicle faces a spillback condition in its receiving link at the time. Based on the observation on field traffic patterns and the familiarity of the area, the frequency of vehicles joining a spillback and hence blocking intersections was quite rare. Therefore, this case study selected very conservative values for these parameters to match the field traffic patterns, as shown in Table 31.

Table 31. Probability of Vehicles Joining Spillback

Number of Vehicles in spillback		1	2	3	4
Probability of vehicles joining spillback	Default	100	81	69	40
	Selected	20	15	10	5

Second, the driver's familiarity with path distribution was also predetermined and excluded from the calibration process. In CORSIM, whether vehicles actually know their successive turn movements and appropriate lanes for those turn movements depends on the familiarity of the drivers with their path. CORSIM randomly assigns every vehicle a value for a driver familiarity based on the distribution in Table 32. The distribution consists of two percentages: (i) percentage of vehicles knowing the next turn movement, and (ii) percentage of vehicles knowing the next two turn movements. The vehicles knowing only the next turn movement may perform sudden lane changes and would result in the blocking of the other lanes to get enough of a gap for lane changing, especially in closely neighboring intersections linking this test site. Therefore, this case study assumed 100% of vehicles knowing the next two turn movements.

Table 32. Probability of Vehicles Joining Spillback

Number of Turn Movements		1	2
Percentage	Default	10	90
	Selected	0	100

Following is the list of 28 calibration parameters and each parameter's initial range.

NETSIM Link Description (RT 1)
    1) Mean value of start-up lost time (Tenth of seconds): 15, 20, 25, 30
    2) Mean queue discharge headway (Tenth of seconds): 15, 18, 20, 25
    3) Desired free-flow speed (mph): 40, 45
Lane-Change Parameters (RT 81)
    4) Duration of a lane-change maneuver (Seconds): 1, 2, 3, 4, 5, 6, 7, 8

5) Mean time for a driver to react to a sudden deceleration of the lead vehicle (Tenth of seconds): 1, 5, 10, 15, 20, 25, 30
6) Minimum deceleration for lane change (fpss): 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
7) Difference in max and min acceptable deceleration for a mandatory lane change (fpss): 5~15
8) Difference in max and min acceptable deceleration for a discretionary lane change (fpss): 5~15
9) Deceleration rate of lead vehicle (fpss): 10~15
10) Deceleration rate of following vehicle (fpss): 10 ~ 15
11) Driver type factor used to compute driver aggressiveness:
    15, 20, 25, 30, 35, 40, 45, 50
12) Urgency threshold (Tenth of fpss): 0 ~ 5
13) Safety factor X 10: 6 ~ 10
14) Percentage of drivers who cooperate with a lane changer (%):10, 20, 30, 40, 50, 60, 70, 80, 90, 100
15) Headway below which all drivers will attempt to change lanes (Tenth of seconds): 1, 5, 10, 15, 20, 25, 30
16) Headway above which no drivers will attempt to change lanes (Tenth of seconds): 30, 40, 50, 60, 70, 80, 90, 100
17) Mean longitudinal distance over which drivers decide to perform one lane change (ft): 50 ~ 2,500

Left-Turn Jump, Left- and Right-Turning Speeds (RT 140)
18) Left turn jumper probability (%): 10, 20, 30, 40
19) Left turn speed (fps): 13, 19, 25, 31
20) Right turn speed (fps): 13, 17, 21, 25

Probability of Left-Turn Lagger (RT 141)
21) Left-turn lagging within 2 seconds (%): 20, 30, 40, 50
22) Left-turn lagging for 2~4 seconds (%): 5, 10, 15

Amber Interval Response (RT 144)
23) Acceptable deceleration for 10 driver types (fpss):
    Default: 21, 18, 15, 12, 9, 7, 6, 5, 4, 4
    Shift to left: 19, 16, 13, 10, 7, 5, 4, 3, 2, 2
    Shift to right 1: 23, 20, 17, 14, 11, 9, 8, 7, 6, 6
    Shift to right 2: 25, 22, 19, 16, 13, 11, 10, 9, 8, 8

Gaps for Permissive Left-turns and for Right-Turns on Red or at Signs (RT 145)
24) Gap distribution for left turns (sec)
    Default: 7.8, 6.6, 6.0, 5.4, 4.8, 4.5, 4.2, 3.9, 3.6, 2.7
    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
25) Gap distribution for right turns (sec)
    Default: 10.0, 8.8, 8.0, 7.2, 6.4, 6.0, 5.6, 5.2, 4.8, 3.6
    Shift to left 1: 9.0, 7.8, 7.0, 6.2, 5.4, 5.0, 4.6, 4.2, 3.8, 2.6
    Shift to left 2: 8.0, 6.8, 6.0, 5.2, 4.4, 4.0, 3.6, 3.2, 2.8, 1.6

Free-Flow Speed Percentage (RT 147)
26) Distribution of free flow speed by driver types (%)
    Default: 75, 81, 91, 94, 97, 100, 107, 111, 117, 127

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

Link Type Distributions (RT 149)

27) Start-up lost time distribution (%)
Default: 218, 140, 125, 118, 102, 86, 78, 63, 47, 23
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

28) Discharge headway distribution (%)
Default: 170, 120, 120, 110, 100, 100, 90, 70, 70, 50
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

8.3.4. Experimental Design for Calibration

Two hundred parameter combinations were generated within initial parameter ranges determined by using Latin Hypercube Sampling method.

Five randomly seeded runs were conducted in CORSIM for each of the 200 combinations, for a total of 1,000 runs. The average travel time of two travel time measuring sections of each random seeded run was recorded for 1,000 runs and the results from the five multiple runs were then averaged to represent each of the 200 parameter sets.

8.3.5. Feasibility Test

The histogram of travel time output of the 200 combinations of both section 1 and 2 are shown in Figure 63 (a) and (b). Based on the histogram analysis, it was able to conclude that the initial parameter ranges could capture field condition and be ready to be used as a range to be calibrated. Travel time section 1 included the field data, 150.3 seconds, around 75th to 80th percentile location and travel time section 1 included the field data, 116.3 seconds, around the 85th to 90th percentile.
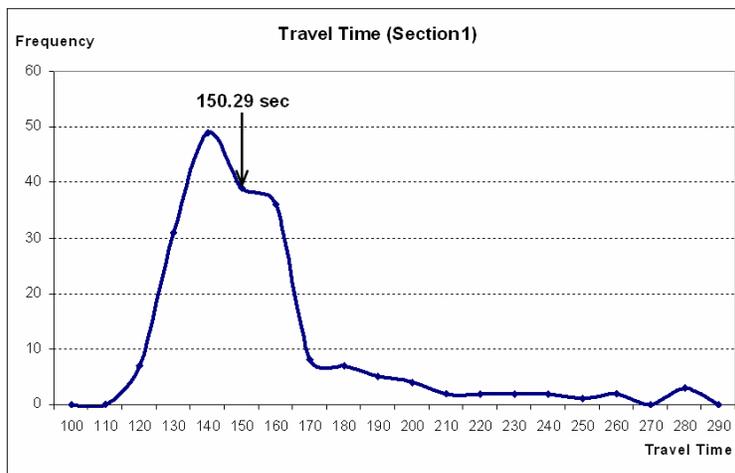
(a) Travel Time of Section 1
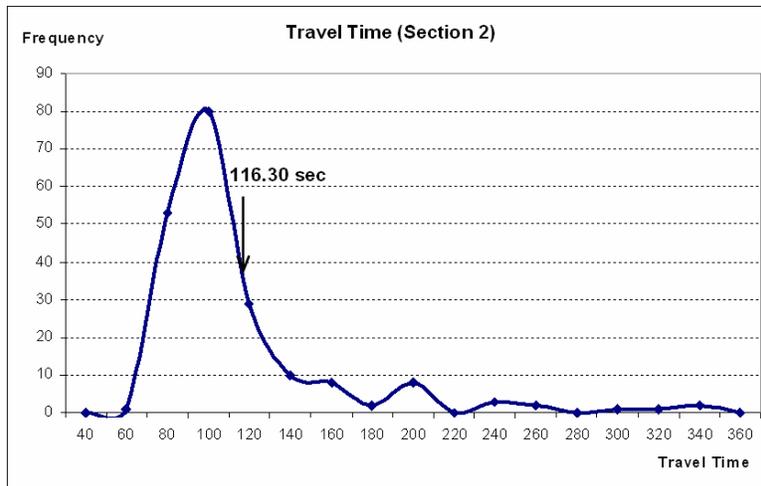
(b) Travel Time of Section 2

Figure 63. Feasibility Test Results of LHD Samples

Even though the initial ranges were acceptable to be used for further calibration purposes, key calibration parameters were tried to be identified for future use. In this case study, statistical analysis (ANOVA) was used for key calibration parameter identification purposes and the result of analysis is shown in Table 33.

Table 33. ANOVA Results (p-value) of LHD Samples Using Travel Time of Sections 1 & 2

Calibration Parameter	Section 1	Section 2
Mean value of start-up lost time	**0.00**	**0.00**
Mean queue discharge headway	**0.00**	**0.00**
Desired free-flow speed	0.07	0.30
Duration of a lane-change maneuver	**0.00**	**0.00**
Mean time for a driver to react to a sudden deceleration of the lead vehicle	**0.00**	**0.00**
Minimum deceleration for lane change	0.24	0.22
Difference in max/min acceptable deceleration for a mandatory lane change	0.50	0.55
Difference in max/min acceptable deceleration for a discretionary lane change	0.90	0.91
Deceleration rate of lead vehicle	0.77	0.86
Deceleration rate of following vehicle	**0.02**	**0.04**
Driver type factor used to compute driver aggressiveness	0.97	0.79
Urgency threshold	0.41	0.51
Safety factor X 10	0.85	0.96
Percentage of drivers who cooperate with a lane changer	0.94	0.80
Headway below which all drivers will attempt to change lanes	0.15	0.31
Headway above which no drivers will attempt to change lanes	0.56	0.68
Mean longitudinal distance over which drivers decide to perform one lane change	0.37	0.12
Left turn jumper probability	0.43	0.52
Left turn speed	0.89	0.69
Right turn speed	0.51	0.13
Left-turn lagging within 2 seconds	0.34	0.31
Left-turn lagging for 2~4 seconds	0.22	0.17
Acceptable deceleration for 10 driver types	0.61	0.80
Gap distribution for left turns	0.48	0.85
Gap distribution for right turns	0.79	0.91
Distribution of free flow speed by driver types	0.14	0.23
Start-up lost time distribution	0.07	0.15
Discharge headway distribution	0.05	0.08

* Note: level of significance 0.05 was used.

According to the results of the statistical analysis, the key parameters were as follows:
− Mean value of start-up lost time
− Mean queue discharge headway
− Desired free-flow speed
− Duration of a lane-change maneuver
− Mean time for a driver to react to a sudden deceleration of the lead vehicle
− Start-up lost time distribution

8.3.6. Parameter Calibration Using Genetic Algorithm

A genetic algorithm was integrated with the CORSIM model to calibrate each parameter value. This algorithm was designed to minimize the result of the following equation that minimizes the difference between the field-measured travel time and simulation travel time outputs.

$$\frac{\sum (\text{Travel Time}_{Field} - \text{Travel Time}_{Simulation\,(i)})}{\text{Travel Time}_{Field}}$$

In this case, individual travel time output of each simulation run was considered rather than considered as an averaged value. The simulation travel time output was the value averaged from the five CORSIM runs of each parameter set. 10 generations and 10 populations were adopted in the GA process. The convergence of fitness value with generations is shown in Figure 64. The parameter set with the best fitness value was selected to represent local traffic conditions and selected in the final evaluation.
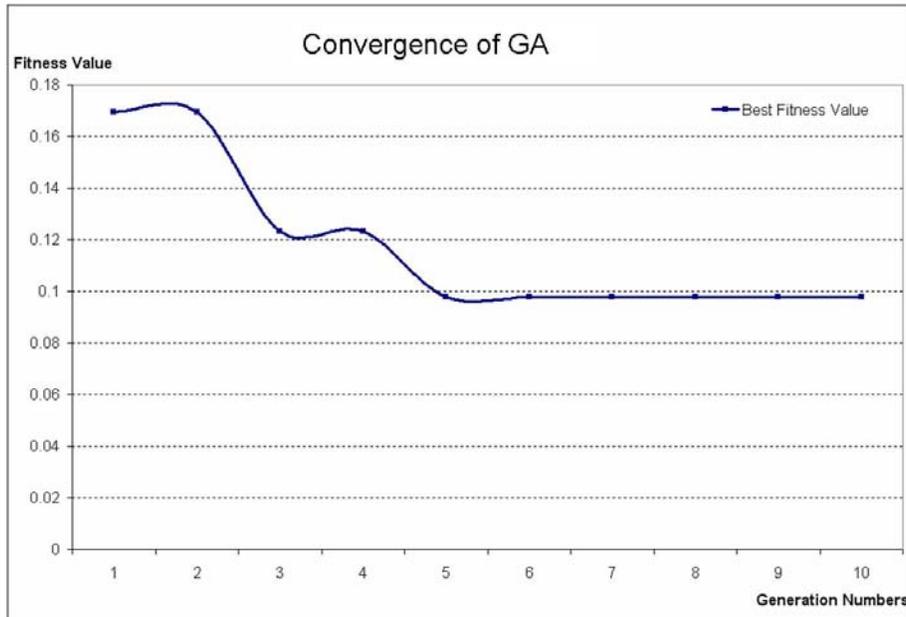
Figure 64. Convergence Curve—CORSIM, Charlottesville

8.3.7. Evaluation of the Parameter Sets

Since the calibrated parameter set was obtained from the previous step, 100 multiple runs with the calibrated parameter set was conducted to check the performance of the model with a calibrated parameter set. Figure 65 and Table 34 summarize the evaluation results of 100 travel time outputs for both travel time measuring sections 1 and 2.
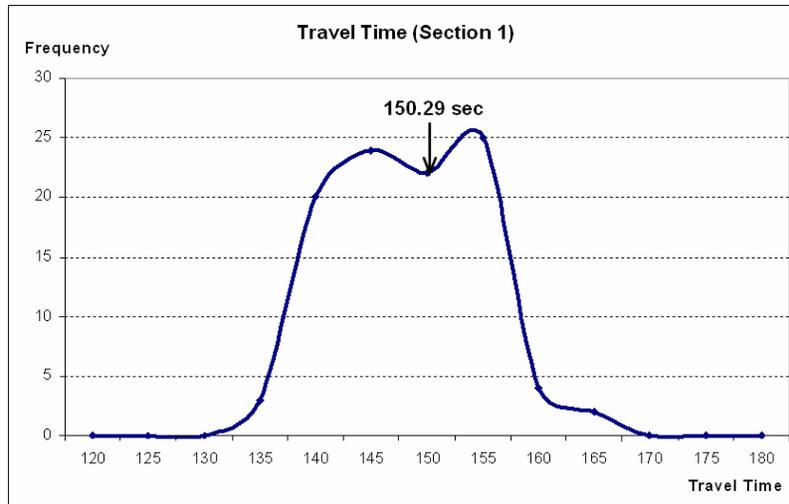
(a) Travel Time of Section 1
(b)
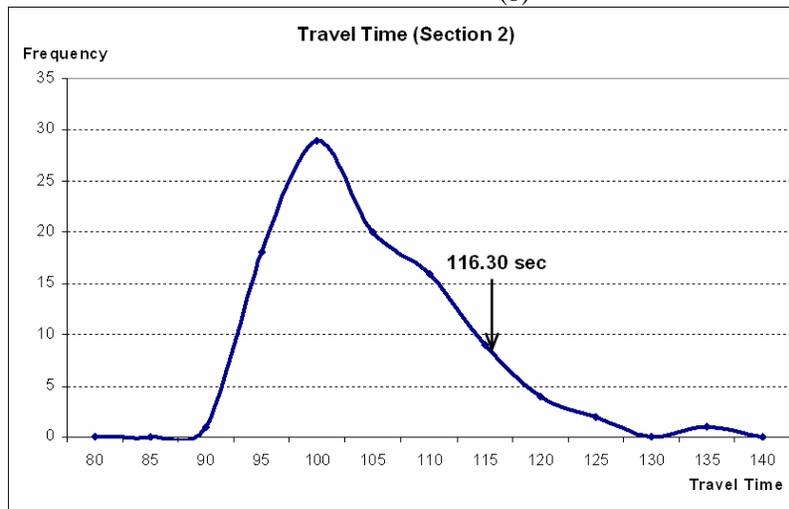
(b) Travel Time of Section 2

Figure 65. Evaluation Result of Calibrated Parameter Set

Table 34. Average Value of Performance Measure with Calibrated Parameters

MOEs	Simulation Results	Actual Values
Travel time of Section 1, seconds	145.7	150.3
Travel time of Section 2, seconds	101.8	116.3
Maximum queue length at SB approach at Morton Rd, vehicle	16	16
Maximum queue length at NB approach at Morton Rd, vehicle	15	24

8.3.8. Validation

In order to validate the model with a calibrated parameter set, a new data set, which was not used in the calibration, was applied. Even though the traffic pattern on that day (Friday) was quite different from the other days (Wednesday and Thursday), the simulated travel time was similar to the actual travel time. Especially, during the day, there were very long queues in the rightmost

96

lane from Hydraulic Rd. to the first entrance ramp to Rte. 250 and thus the travel time of Section 2 shows a relatively bigger discrepancy than that of Section 1. Also, the sizes of maximum queues are not very well matched. Table 35 shows the result of the validation procedure.

Table 35. Validation Result

MOEs	Standard Deviation	Actual Values
Travel time of Section 1, seconds	117.7 (27.8)	138.2
Travel time of Section 2, seconds	158.8 (14.7)	157.0
Maximum queue length at SB approach at Morton Rd. vehicles	17 (1.9)	23
Maximum queue length at NB approach at Morton Rd. vehicles	22 (9.6)	26

After completing the calibration and validation procedure using single performance measure, the necessity of the procedure using multiple performance measures was evaluated by verifying simulation outputs that have been generated with the calibrated model using single performance measure. In other words, if the calibrated model satisfies the criteria provided in the enhanced procedure, then it is not necessary to conduct the calibration procedure using multiple performance measures. This is because the enhanced procedure would produce similar results. Figure 12 shows the result from multiple simulation runs with the calibrated model under single performance measure. The results were presented in both histograms and X-Y plots to determine the need for using multiple performance measures.

(a) Travel Time (Section 1)   (b) Travel Time (Section 2)

(c) Travel Time (Section 1) vs. Traffic Count   (d) Travel Time (Section 2) vs. Traffic Count
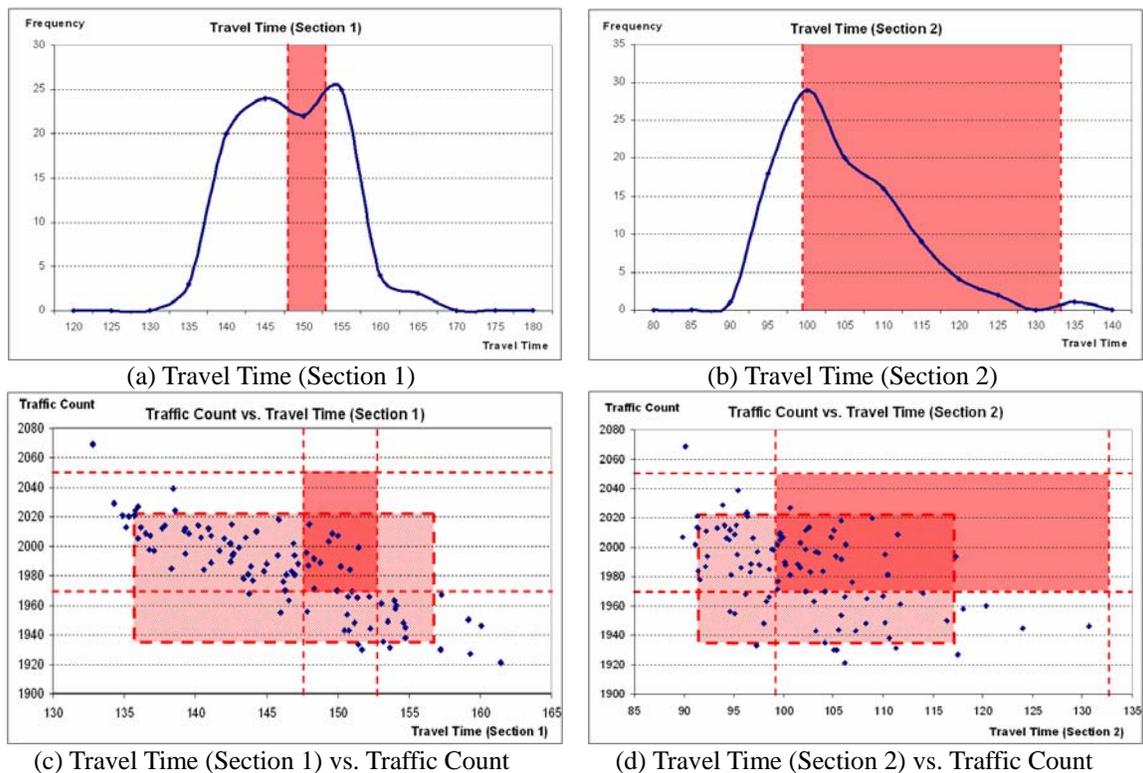
Figure 66. Performance of Calibrated Model using Single Performance Measure—CORSIM

As shown in Figure 66 (a) and (b), the distributions of travel time outputs with the calibrated model could include the field-measured travel time ranges. Based on two histograms, it could be concluded that the model is well-calibrated.

The traffic count data were added to the consideration and presented as X-Y plots and compared with the field-collected data ranges. As shown in Figure 66 (c) and (d), 90% confidence interval region of simulation output data were overlapping with the field-collected performance measure data ranges. Therefore, it could be concluded that the enhanced procedure is not necessary.

8.4. VISSIM Calibration

As mentioned earlier, this network has a complex geometry condition as well as a congested traffic condition. Due to those characteristics, the use of multiple performance measures was considered as a calibration method. However, trial with the single performance measure, travel time data, was conducted ahead of multiple performance measures consideration and the result turned out to be not valid by overfitting other types of data that have not been used for the calibration. Therefore, multiple types of data, i.e., travel time, traffic volume, and maximum queue length data, were used for the calibration with VISSIM.

8.4.1. Efforts on Network Building

Due to the similar reasons described in the CORSIM case study section (complexity of the network), it is required to modify the network to mimic the real field condition. The problem that we had with VISSIM was similar to the problem that we had with CORSIM. Initially, the VISSIM network was built without any special considerations on the long queue on the rightmost side lane on Rte. 29 south. As a result, the vehicles that needed to diverge to the ramp toward Rte. 250 west did not generate a long queue but blocked the whole section because some of the vehicles changed their lane right in front of the diverging point. As a result, everywhere in the network was congested, which is not realistic at all.

So, our research team assumed that the vehicles select their lane before they enter the network. In order to create that kind of condition, vehicle entry points for two different destination groups were separated with different lane groups:
- Vehicle group that makes a right turn up to on-ramp toward Rte. 250 west
- Other vehicles

The concept of separate entries by lane is shown in Figure 67.
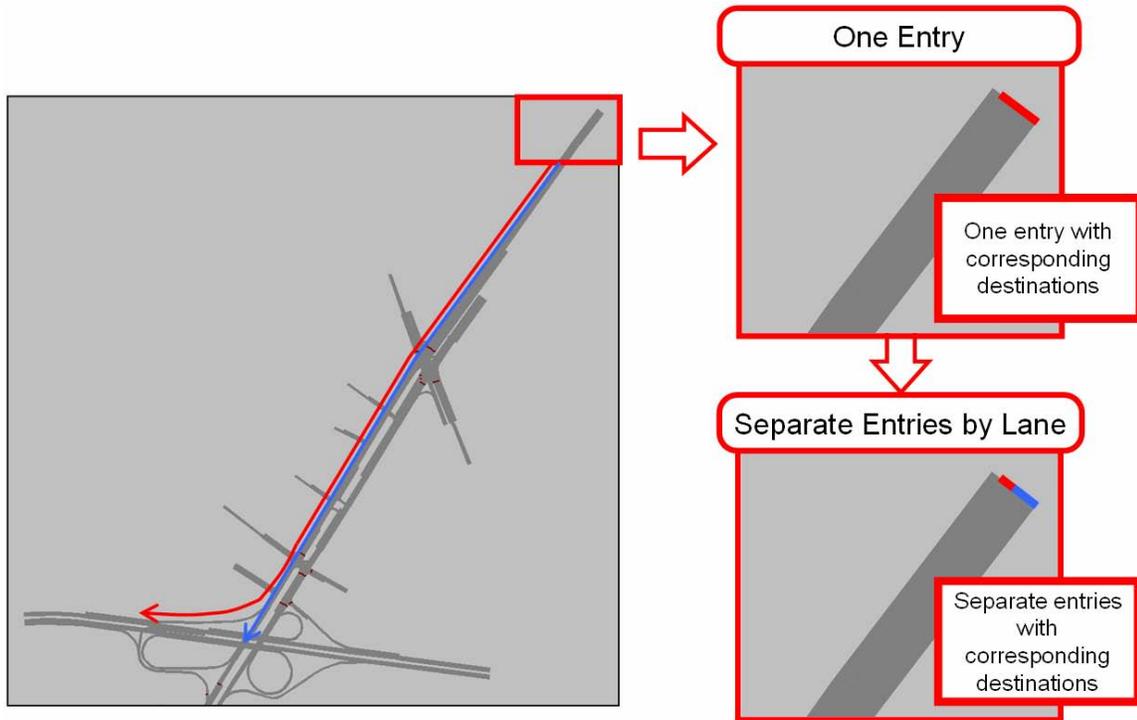
Figure 67. Modifications on Vehicle Entry Point

8.4.2. Evaluation with Default Parameter Set

100 randomly seeded runs with the default parameter set were conducted to investigate the feasibility of the model with the default parameter set. As shown in Figure 68, the distribution of travel time output with the default parameter set for both sections did not include the field-measured travel time. Therefore, a further procedure of calibration is considered to be required.
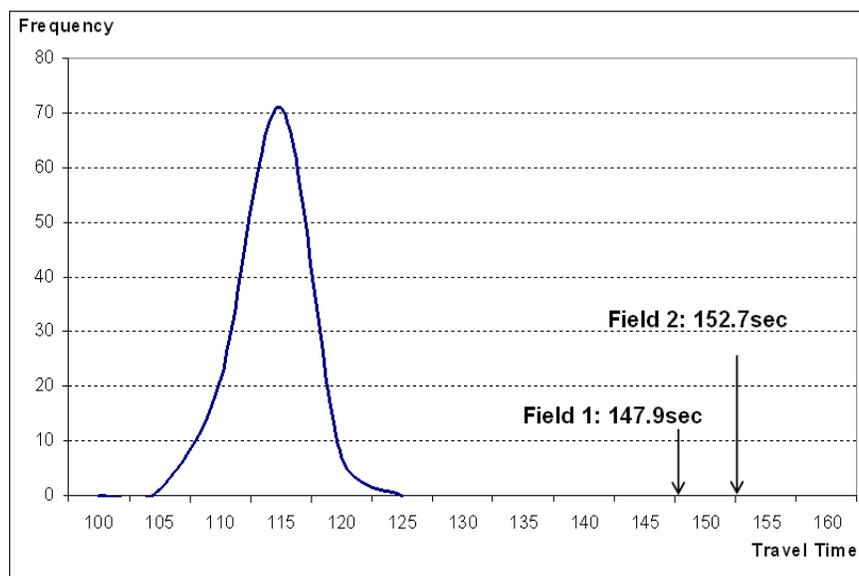
(a) Travel Time Result (Section 1)

99
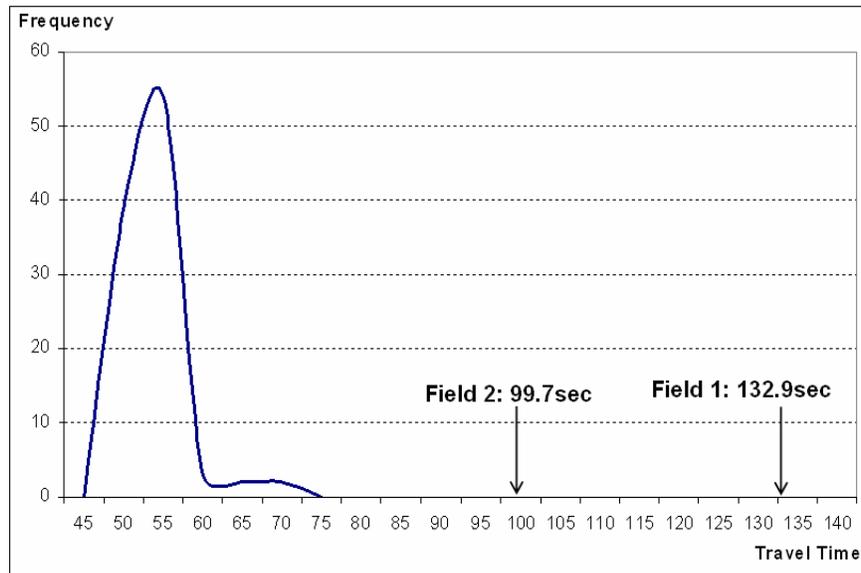
(b) Travel Time Result (Section 2)

Figure 68. Distribution of Travel Time Output with Default Parameter Set

8.4.3. Identification of Calibration Parameters

Since this network is in an urban area, the Wiedemann 74 model is selected as a car-following model in VISSIM. Also some other calibration parameters on lane-changing behavior are included as well as the desired speed distributions.

Following is the list of calibration parameters and each parameter's initial range**:**
  2) Simulation resolution (Time step/sim.sec) : 1 ~ 9
  3) Number of preceding vehicles (veh): 1 ~ 6
  4) Maximum look ahead distance (m): 200 ~ 300
  5) Average standstill distance (m): 1 ~ 3
  6) Additive part of desired safety distance: 2.5 ~ 5.5
  7) Multiple part of desired safety distance: 3.5 ~ 6.5
  8) Waiting time before diffusion (sec): 30 ~ 90
  9) Minimum headway (m): 0.5 ~ 1.5
  10) Maximum deceleration (m/s^2): -5.0 ~ -2.0
  11) Reduction rate (m): 50 ~ 150
  12) Accepted deceleration (m/s^2): -1.5 ~ -0.5
  13) Desired speed distribution (mph)
    - Rte. 29 / Car: 35~45, 40~50, 45~55
    - Rte. 29 / Heavy Vehicle: 30~40, 35~45, 40~50
    - Ramps / Car: 15~25, 20~30, 25~35
    - Ramps / Heavy Vehicle: 15~25, 20~30, 25~35

8.4.4. Experimental Design for Calibration

100

Two hundred parameter combinations were generated within initial parameter ranges determined by using Latin Hypercube Sampling method.

Five randomly seeded runs were conducted in VISSIM for each of the 200 cases, for a total of 1,000 runs. The average travel time was recorded for each of the 1,000 runs. The results from the five multiple runs were then averaged to represent each of the 200 parameter sets.

8.4.5. Feasibility Test

A feasibility test was conducted to check whether the simulated outputs based on the current parameter ranges could capture the field data or not. If it does not include the field-measured travel time, then additional effort of finding key parameters needs to be made. As shown in Figure 69, the minimum and maximum values of field-collected performance measure values were used as a boundary of each performance measure and the region where two ranges are intersecting is considered as an acceptable region and presented as a shaded box. Three X-Y plots are presented, from Figure 69 (a) to (c), to check the combination of three independent performance measure data. As shown in the later two X-Y plots, Figure 68 (b) and (c), combinations of travel time of section 2 versus traffic counts and travel time of section 1 versus section 2 were acceptable because the acceptable region, shaded box, could include a certain number of data points. However, in the case of the combination of travel time of section 1 versus traffic count, no dots were located within the acceptable region, which indicates that there is no or a small chance of finding a parameter set that produces the field condition and needs further calibration procedure.
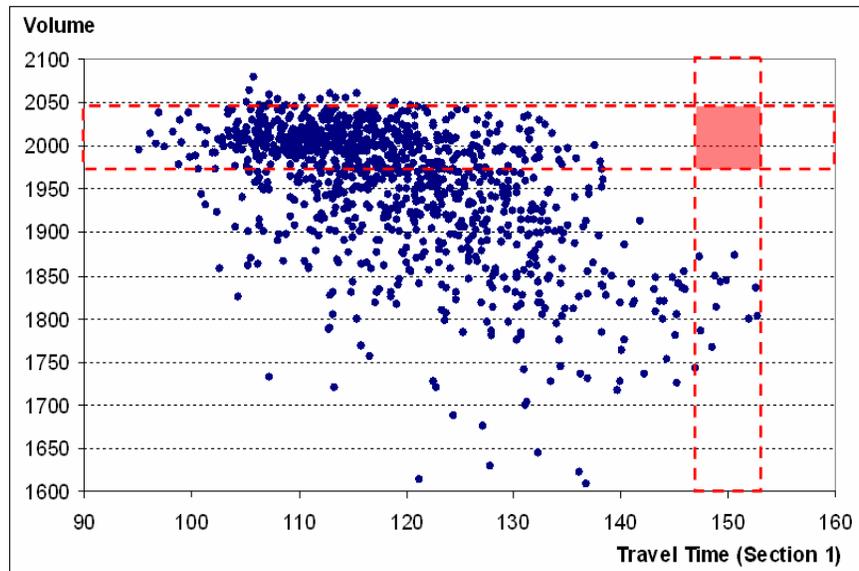
(a) Travel Time (Section 1) vs. Traffic Count
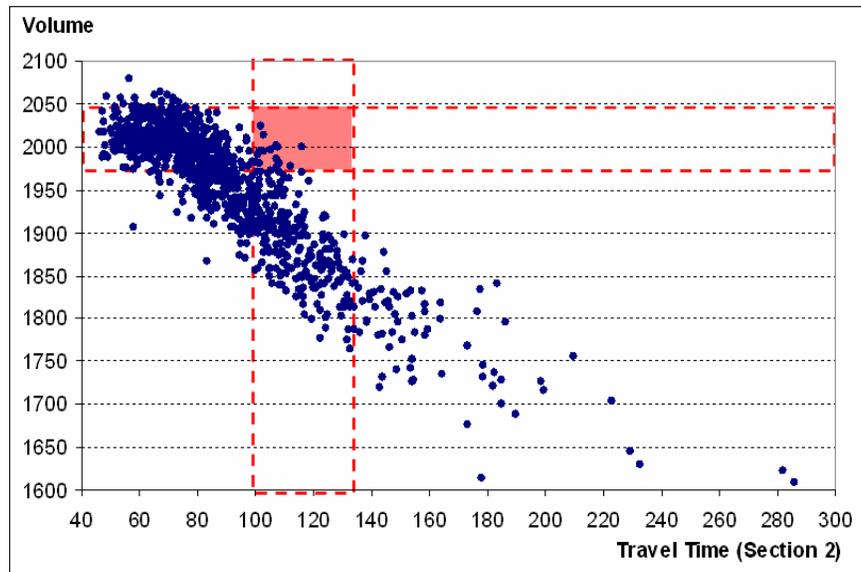
(b) Travel Time (Section 2) vs. Traffic Count
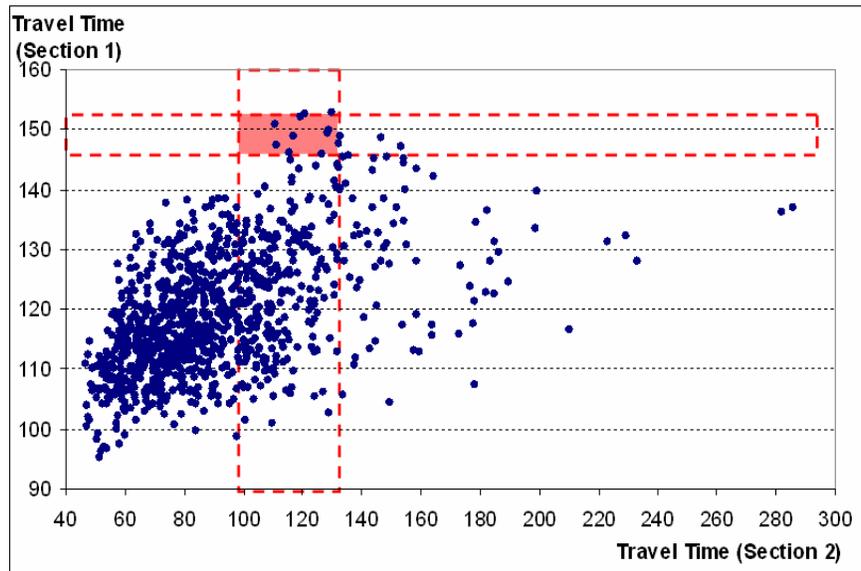
(c) Travel Time (Section 1) versus Travel Time (Section 2)

Figure 69. Feasibility Test Result—VISSIM, Charlottesville

8.4.6. Parameter Range Adjustment

Before conducting parameter range adjustment, it is essential to identify the key parameter that affected the performance measures, travel time of two sections, and traffic counts. Two different methods have been applied to identify key parameters, which are interval plot and statistical analysis (ANOVA).

As a first step, the interval plot of each parameter versus travel time and traffic count were drawn and the trend of travel time changes as each calibration parameter alterations were examined. Figures 70 and 71 show two examples of interval plot of calibration parameter versus

performance measure. Both figures were presented as an example of a calibration parameter that shows significant effect to the performance measure value.

Table 36 shows ANOVA result with p-value. The small p-value indicates that the corresponding parameter affected the performance measure value significantly. The significance level applied in this case study was 0.05.
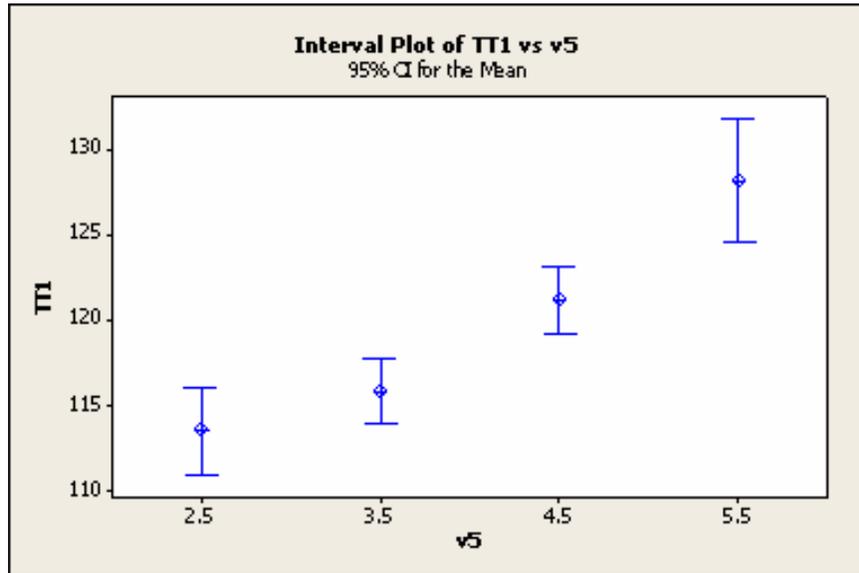
Figure 70. Additive Part of Desired Safety Distance vs. Travel Time (Section 1)
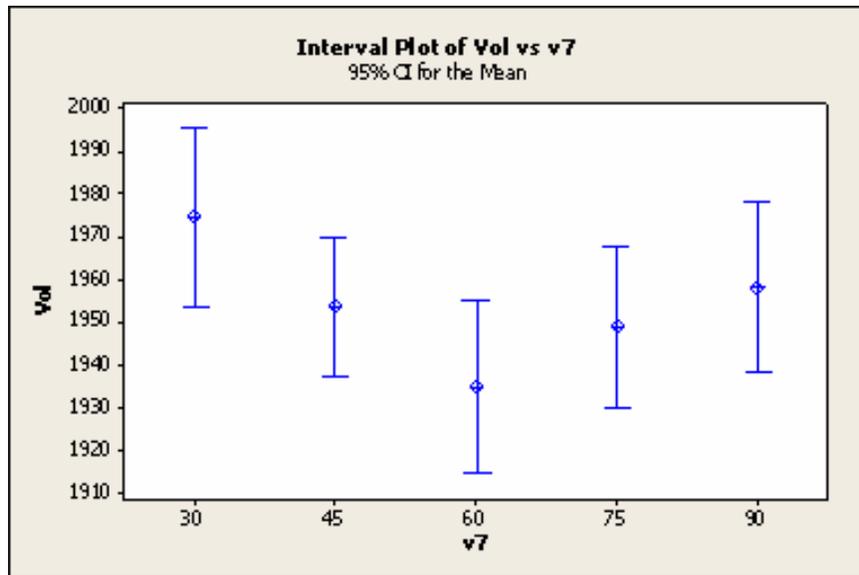
Figure 71. Waiting Time Before Diffusion vs. Traffic Count

Table 36. ANOVA Result for Charlottesville Network with VISSIM

Parameters		p-value		
		Travel Time (Section 1)	**Travel Time (Section 2)**	**Volume**
Simulation resolution		**0.003**	0.138	0.347
Number of observed preceding vehicles		0.555	0.670	0.473
Maximum look ahead distance		0.593	0.759	0.861
Average standstill distance		**0.004**	0.101	**0.015**
Saturation Flow Rate	Additive part of desired safety distance	**0.000**	**0.000**	**0.000**
	Multiple part of desired safety distance	0.192	**0.002**	**0.001**
Waiting time before diffusion		0.134	0.075	0.112
Min. headway		0.340	**0.023**	**0.037**
Max. deceleration		0.256	0.085	0.499
Reduction rate		0.518	0.189	0.219
Accepted deceleration		**0.018**	0.469	0.485
Des. Speed Dist.(Rt29/Car)		**0.000**	0.098	**0.036**
Des. Speed Dist.(Rt29/HV)		0.613	0.633	0.456
Des. Speed Dist.(Ramp/Car)		0.637	**0.000**	**0.002**
Des. Speed Dist.(Ramp/HV)		0.366	0.334	0.496

Based on the analyses, the following seven parameters were identified as a key parameter to different performance measures: simulation resolution, average standstill distance, additive/multiple part of desired safety distance, minimum headway, and desired speed distributions for car.

Among those key parameters, the following three parameters showed significantly low p-value (0.00): (1) additive part of desired safety distance, (2) desired speed distribution of Rte. 29 for car, and (3) desired speed distribution of ramp segment for car.

Combined Effect of Non-Key Parameters

In addition to the key parameter identification, the interaction of key parameters and other calibration parameters were investigated by using a 3D contour plot. Once key parameters are identified, it is required to understand the interaction of key parameters and other parameters. The interaction of two parameters can be presumed by looking at the color as two parameter values alter. Figures 72 and 73 show an example of the 3D contour plot for the key parameters versus other calibration parameters.
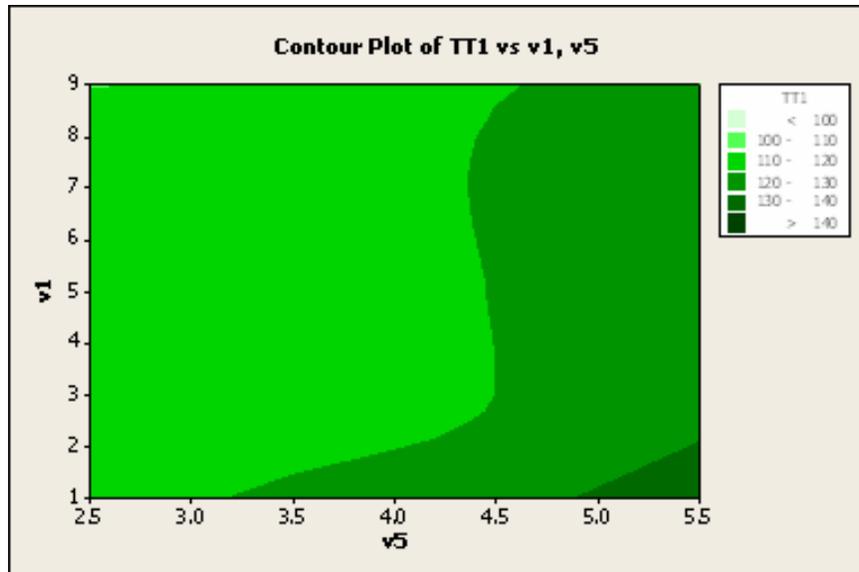
Figure 72. Travel Time (Section 1, TT1) vs. Simulation Resolution (V1) vs. Additive Part of Desired Safety Distance (v5)
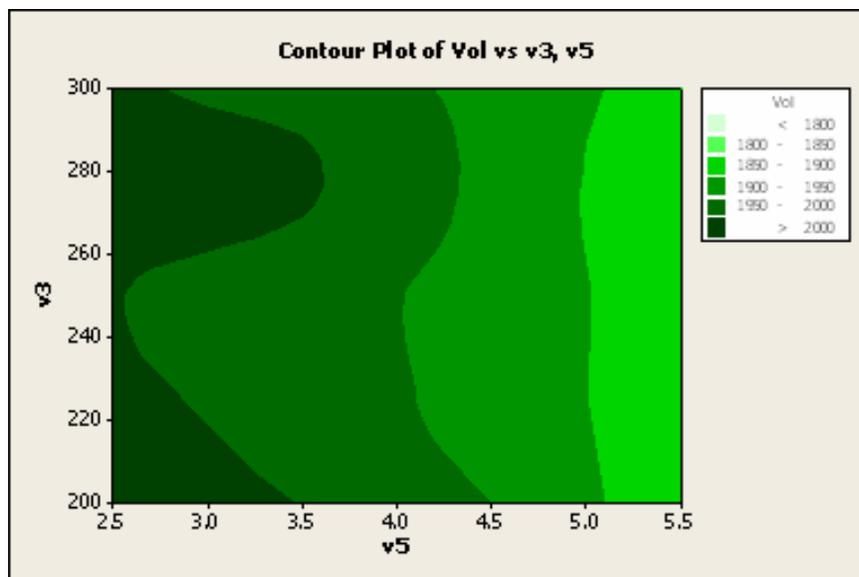
Figure 73. Traffic Count (vol) vs. Max. Look Ahead Distance (v3) vs. Additive Part of Desired Safety Distance (v5)

The interactions between three key parameters and others were examined by using the 3D contour plot. It also indicated that certain parameters need to be shifted or tightened to maximize the impact of key parameter modification. For instance, as shown in Figure 72, simulation resolution should have a lower value since it showed higher travel time output with an identical additive part of desired safety distance value.

Based on analyses, ranges of all calibration parameters were either shifted or tightened. Above all, even though it was set up based on the speed limit data obtained from the field, desired speed distributions were considered to be higher than appropriate value and minimum speeds were

reduced by 10 mph. On the other hand, the ranges of additive/multiple part of desired safety distance, which is known to affect capacity the most, were tightened to have lower value than before. The set of calibration parameter ranges after adjustment is shown in Table 37 as well as initially determined ranges.

Table 37. Two Calibration Parameter Ranges for Charlottesville Network

Parameters			Unit	Initial		Adjusted	
				Min	Max	Min	Max
Simulation resolution			Time steps/Sim. Sec.	1	9	1	4
Number of observed preceding vehicles			veh	1	6	2	4
Maximum look ahead distance			m	200	300	150	250
Average standstill distance			m	1.0	3.0	1.5	2.5
Saturation Flow Rate	Additive part of desired safety distance		-	2.5	5.5	2.5	4.5
	Multiple part of desired safety distance		-	3.5	6.5	3.5	5.5
Waiting time before diffusion			sec	30	90	45	55
Min. headway			m	0.5	1.5	0.75	1.25
Max. deceleration			m/s^2	-5.0	-2.0	-3.5	-2.5
Reduction rate			m	50	150	75	125
Accepted deceleration			m/s^2	-1.5	-0.5	-1.5	-0.75
Desired Speed Distribution	Rte29/Car	Min.	mph	30~40, 35~45, 40~50, 45~55		25, 30, 35	
		Var.				10, 15, 20	
	Rte29/HV	Min.		30~40, 35~45, 40~50		20, 25, 30	
		Var.				10, 15, 20	
	Ramp/Car	Min.		15~25, 20~30, 25~35		15, 20, 25	
		Var.				5, 10, 15	
	Ramp/HV	Min.		15~25, 20~30, 25~35		10, 15, 20	
		Var.				5, 10, 15	

After the adjustment of calibration parameter ranges, 200 combinations were generated again and 5 random seeded runs were conducted for each combination, for a total of 1,000 runs. The average travel time of each simulation run was recorded and presented as a scatter plot of two performance measures as it was in the previous feasibility step. Figure 74 shows three scatter plots of two performance measures on each axis. As a result, all three feasible regions (shaded box) included dots and were considered to be feasible. Therefore, the adjusted parameter ranges were considered to be feasible and can be used for the calibration procedure.
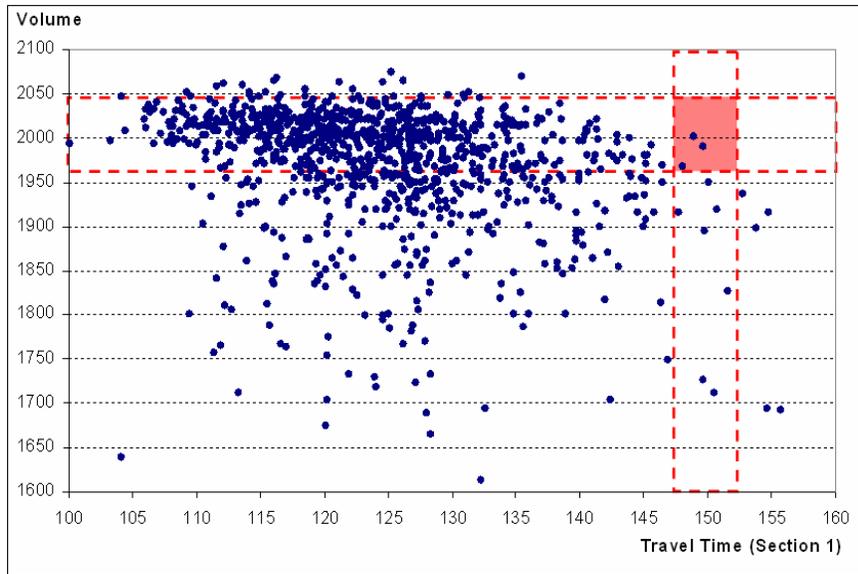
(a) Travel Time (Section 1) versus Traffic Count
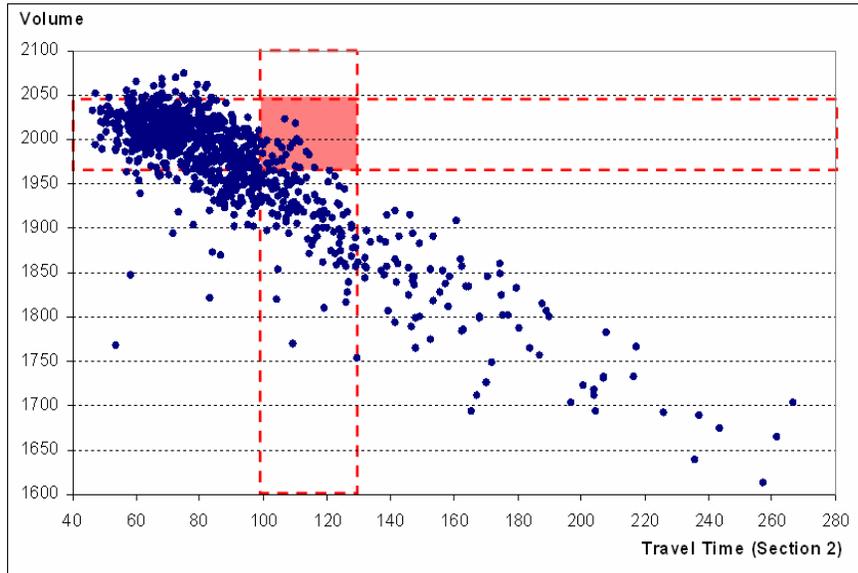
(b) Travel Time (Section 2) versus Traffic Count
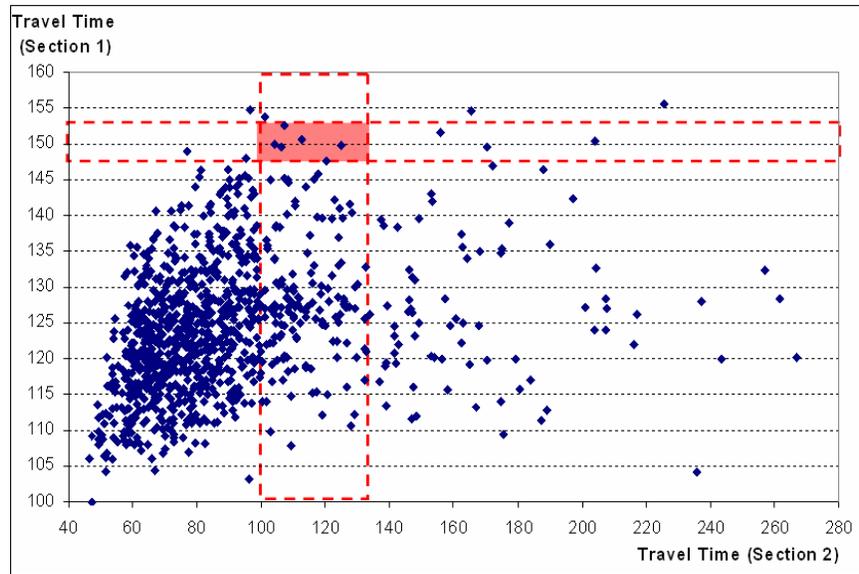
(c) Travel Time (Section 1) versus Travel Time (Section 2)

Figure 74. -Y Plots of Two Performance Measure Values—After Range Adjustment

8.4.7. Parameter Calibration Using Genetic Algorithm

A genetic algorithm was integrated with the VISSIM model to calibrate parameters. In this case study, two different methods that can take multiple performance measures into consideration are all used and presented.

Two methods used in this part utilized the traffic volume data as an additional performance measure. The convergence result of a trial with a log transformation method is shown in Figure 75(a) and the result with a constraint insertion method is shown in Figure 75(b). The results were produced based on 22 generations of GA evolution and a population size of 10 in each generation. The best fitness value was found in the 13th and 16th generations and continued until the 22nd generation; however, both calibration processes were continued to the 22nd generation to have an identical calibration condition.
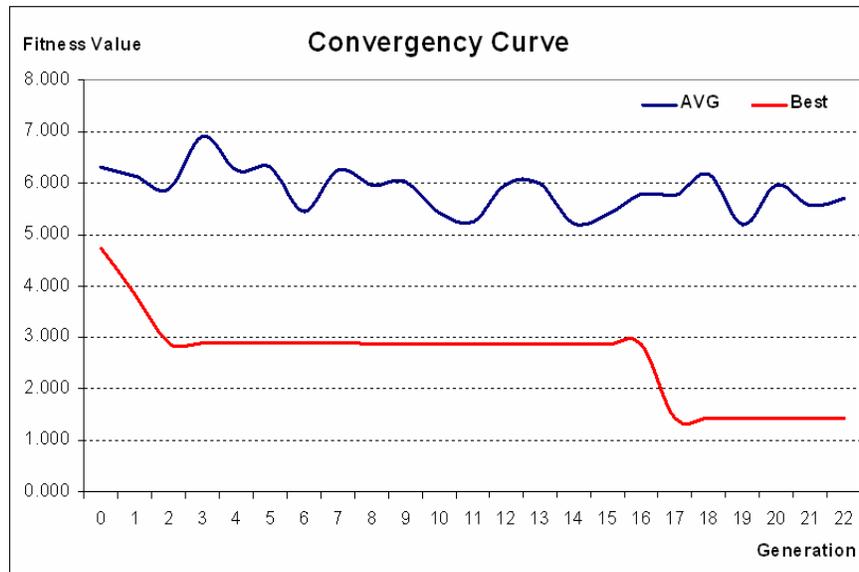
(a) Calibration Method with Multiple Performance Measures (Log Transformation)
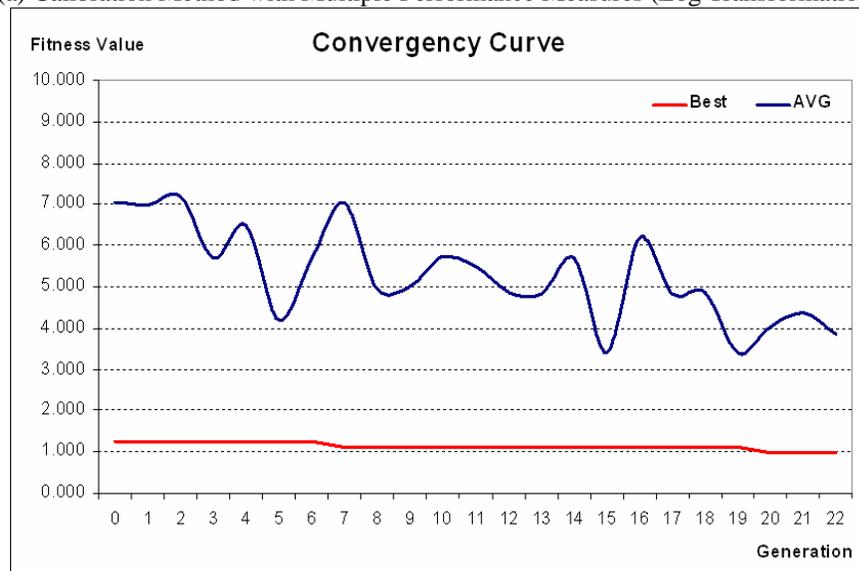
(b) Calibration Method with Multiple Performance Measures (Individual Simulation Output Consideration)

Figure 75. Convergence of Fitness Value with Generation—Charlottesville Network, Existing

8.4.7. Evaluation of the Parameter Sets

This section compares the performance of three calibration parameters sets: calibrated with multiple measures using log transformation method and calibrated with multiple measures using the constraint insertion method. Comparisons between field-observed measures and simulation output measures from each of the two sets were shown in Figure 76 (a) and (b). It is noted that the dark shaded box indicates ranges of field measures, while the light shaded box shows a 90th percentile confidence interval of simulation output measures.

It is ideal to see that both dark- and light-shaded boxes in each chart should overlap each other indicating that the model can replicate field conditions for the performance measures considered.

The results showed that both calibrated models (under multiple performance measures) well overlap between the 90th percentile confidence interval region and field-measured performance measure ranges. Therefore, it can be concluded that the use of multiple performance measures can significantly enhance the accuracy of calibration results. It is noted that visualization testing conducted for both calibrated models with multiple performance measures indicated that their animations are reasonable.
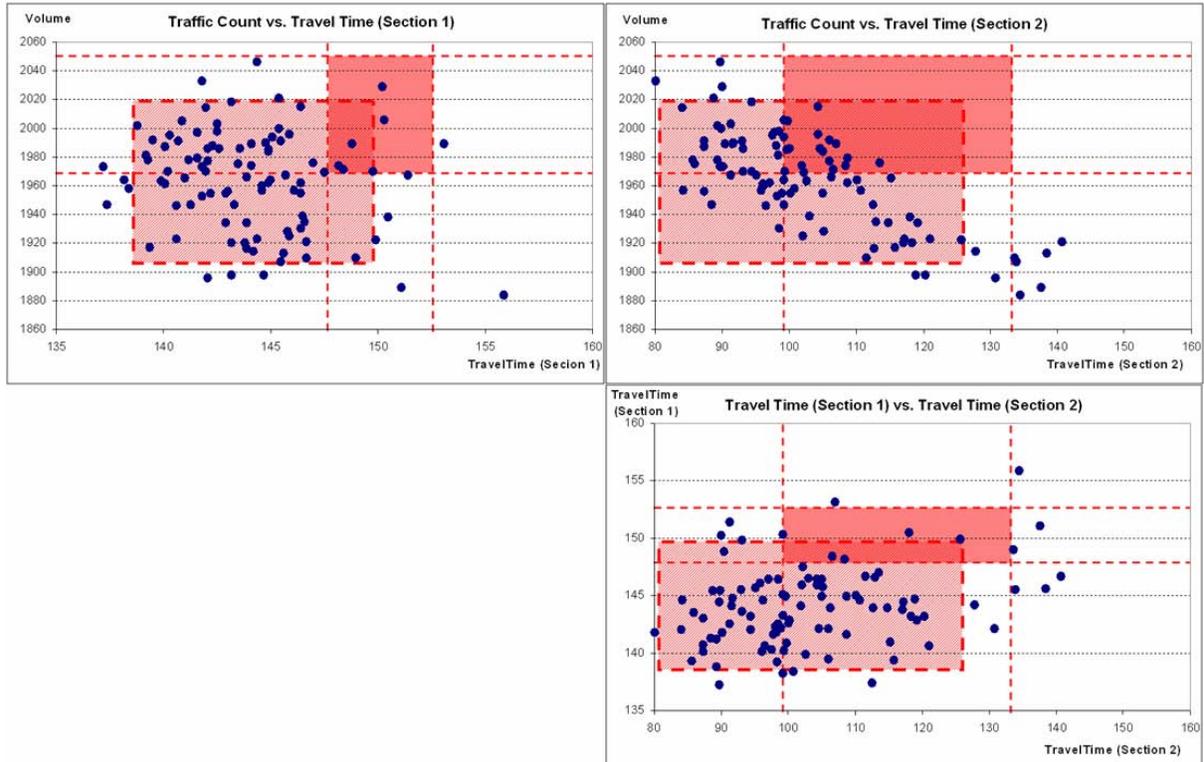
(a) Calibrated with Multiple Performance Measure (Log Transformation Method)
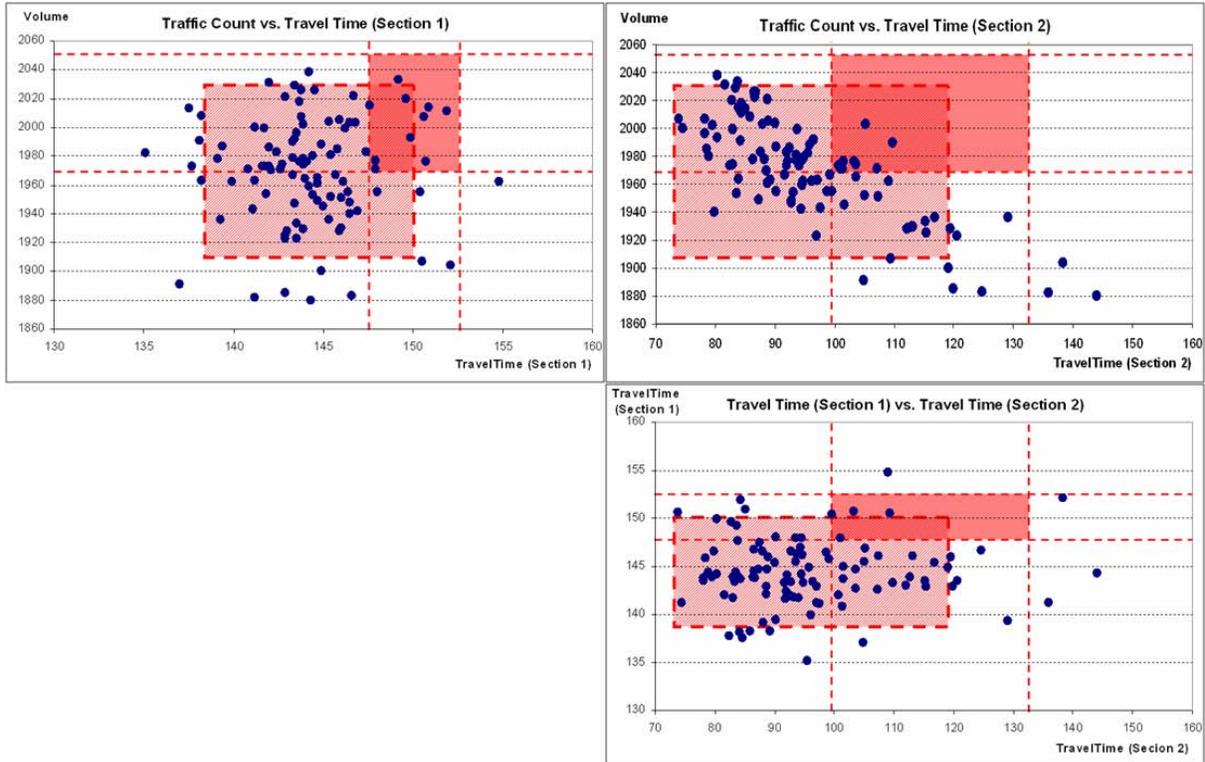
(b) Calibrated with Multiple Performance Measure (Constraint Insertion Method)

Figure 76. Evaluation Result of VISSIM Network with Charlottesville Network

Animations of the model with the calibrated parameter set were viewed in order to determine whether they were realistic or not. For the calibrated parameters, the animations at several travel time percentiles of the distribution were found acceptable. For default parameters, almost all vehicles passed the intersection without waiting, which was not realistic.

In order to validate the two calibrated parameter sets identified from the proposed methodology, new field data that were not exposed during calibration were used and another 100 simulation runs made with the new data set. In addition to maximum queue length data, the validation performance measure, the travel times and traffic counts measures were also extracted for comparison purposes.

As shown in Figure 77 (a), the distributions of maximum queue length obtained from two proposed calibrated methods, log transformation and constraint insertion, include the field-measured maximum queue length. In addition, travel times and traffic counts collected for the validation data set were compared with those from simulation outputs. It is noted that since the validation data were collected on a single day, no ranges of field measurements were available. As such, the 90th percentile confidence intervals of travel times and traffic counts were compared with the corresponding single field measurement. Figures 77 (b) and (c) show that two travel times and traffic counts were within the confidence intervals for the calibrated models under the proposed procedure that incorporates multiple performance measures. Therefore, it can be concluded that the calibrated models are considered to be validated.

111
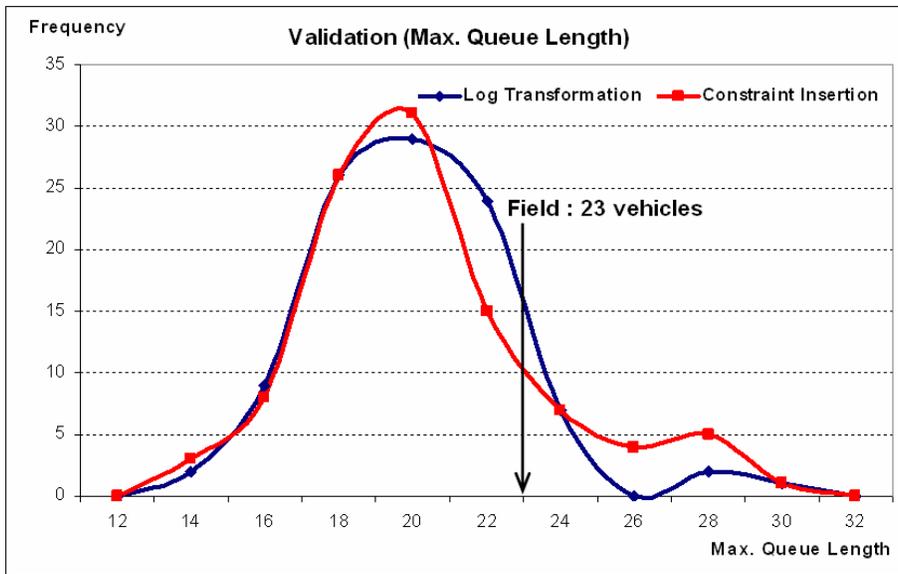
(a) Maximum Queue Length
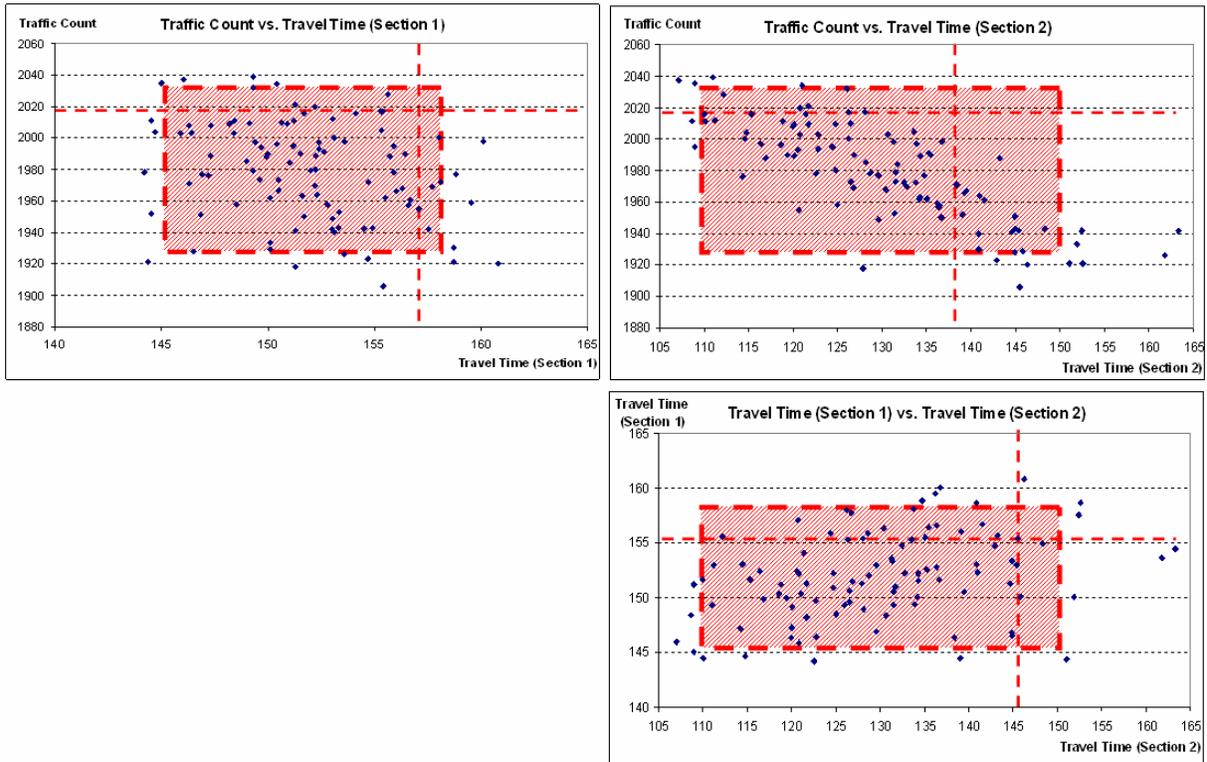
(b) Log Transformation Method
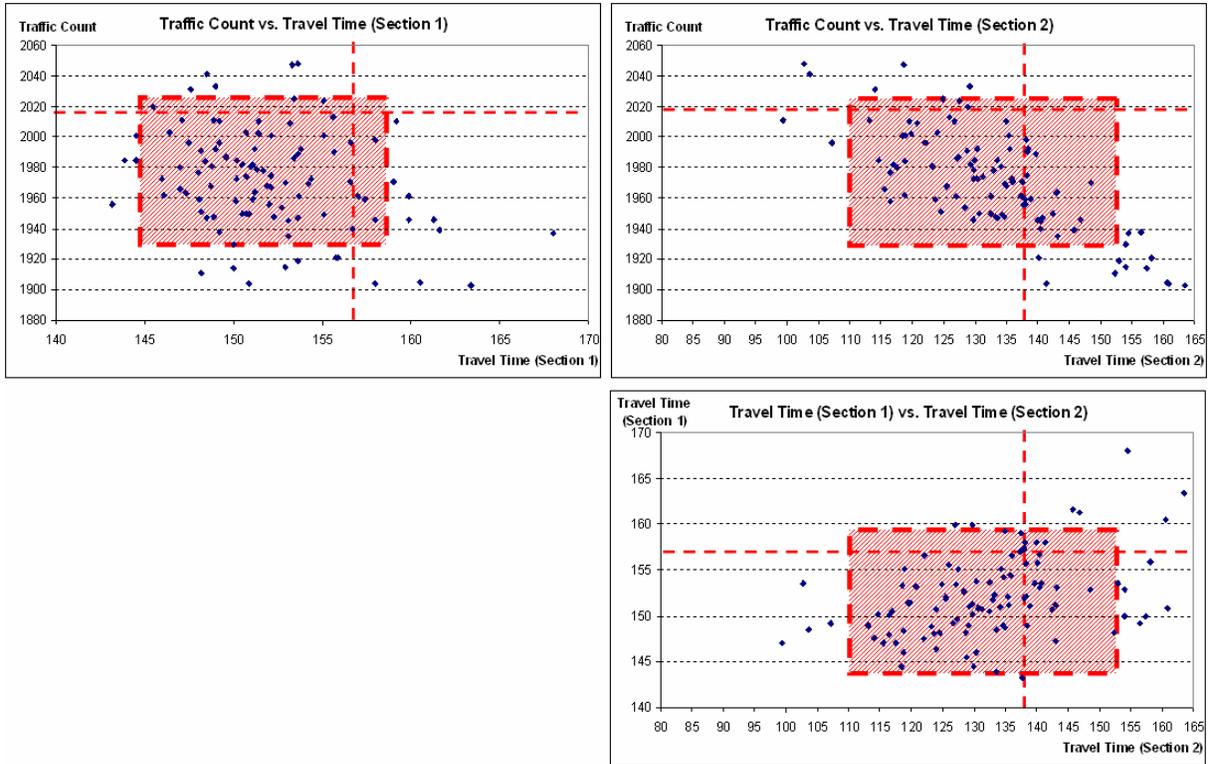
(c) Constraint Insertion Method

Figure 77. Validation Result of VISSIM with Charlottesville Network

Chapter 9
Highway Basic Segment

In this chapter:
- Case study of calibration and validation with highway basic segment and work-zone

---

9.1. Site Description

The test site selected in this case study was a lane-closure work zone located on Interstate Highway 64 in Covington, Alleghany County, Virginia. The selected 5-mile highway segment plies longitudinally in the east-west direction and the westbound direction is of interest in this study. The test site is referred to as "Covington network" in the remaining part of this case study. A unique part of this network is that it involves a relatively large work zone, which has a specific driving phenomenon such as slowing down before the merge area and speeding up after the work zone. The work zone consists of a bridge reconstruction, which takes a fairly long time and is favorable for site visits and multiple-day data collections. There are two lanes at the beginning of the network but only one lane at the work zone area. Moreover, there are three different posted speed limits along the network. The normal posted speed limit for Interstate Highways is 65 mph. At the beginning of the network, there is a posted speed limit sign of 55 mph and further downstream the posted speed limit is 45 mph. Figure 78 is an aerial photo showing the Covington network alignment and location. The gray line represents Interstate Highway 64 and the black line represents the Jackson River in the Covington area.

Figure 78. Schematic of Covington Network

9.2. Data Collection

9.2.1. Data Category

First of all, several types of data are required to be collected from the field: the number of lanes, the length of work zone, speed limits, and traffic volume. The performance measure used for the calibration and validation were travel time of the entire section of the network, and travel speed at the merge area. These two performance measures are selected because of ease of collection from both the field and the simulation tools. In addition, the calculation methods for each are consistent by the simulation models.

9.2.2. Data Collection

Prior to data collection, the research team visited the site twice to determine the locations for collecting data and a safe location to place the data collection equipment. To collect the required traffic data, such as traffic volume, heavy vehicle percentage, travel time, etc., three video cameras were positioned separately at the beginning, merge area, and end of the network. The distance between the first two cameras was around 3 miles while the distance between the second and the third cameras was around 2 miles.

To account for day-to-day variability, multiple days of data sets were conducted. A group of five people collected data during the evening peak hour between 5 and 6 pm on June 10, 24, 25, and 26 in 2003, which were all weekdays. Three of them were responsible for the video cameras and the other two collected speeds using a radar gun at the merging area. Synchronization between clocks of all the equipment was performed before data collection to ensure data quality.

9.2.3. Data Reduction

Data reduction through videotapes consisted of obtaining traffic counts, travel times to pass the network and travel speeds at the merge area. Traffic counts and heavy vehicle counts were obtained through reviewing the videotape of the upstream camera because there was no ramp within the segment. However, there was an on-ramp just around the beginning of the network and the traffic from that entrance was also included. For the Covington network, a vehicle description matching method was used to extract travel times instead of using a license plate matching method because the travel speed in this highway was relatively high and it was difficult to obtain the clear license plate number for every vehicle. The process was to note down the vehicle description (brand, name, color, size, etc.) and record the time appearing in the beginning and end of the network, then match them one by one and extract the time differences, which were the travel times passing the network.

Travel speed collection was relatively easy and time efficient as the software attached to the radar gun automatically recorded the speed and time during data collection.　The only issue was to remove multiple speed records for the same vehicle obtained at different distances because of the fast speed of radar reflection. Therefore, a simple filtering method was manually implemented before these speeds were finally used.

9.2.4. Data Summary

Traffic Counts

Table 38 summarizes traffic counts on four individual days. Data on June 10, 2003 were reserved for validation. The compliance rate was calculated as the percentage of the drivers following the posted speed limits according to the collected speed data on each day.

Table 38. Field Traffic Volume and Composition for Covington Network

Covington	Traffic Counts (veh)	HV%	Compliance Rate (%)
06/10/03*	601	17	30
06/24/03	744	15	18
06/25/03	700	14	23
06/26/03	851	13	30

Travel Time

Table 39 shows the statistics of travel times including mean, standard deviation, and number of vehicles. Covington network travel time data show a small variation of mean travel time on different days. As seen in Figure 79, travel time distributions on different days are quite similar and it may be due to the characteristics of this network that is uninterrupted flow.
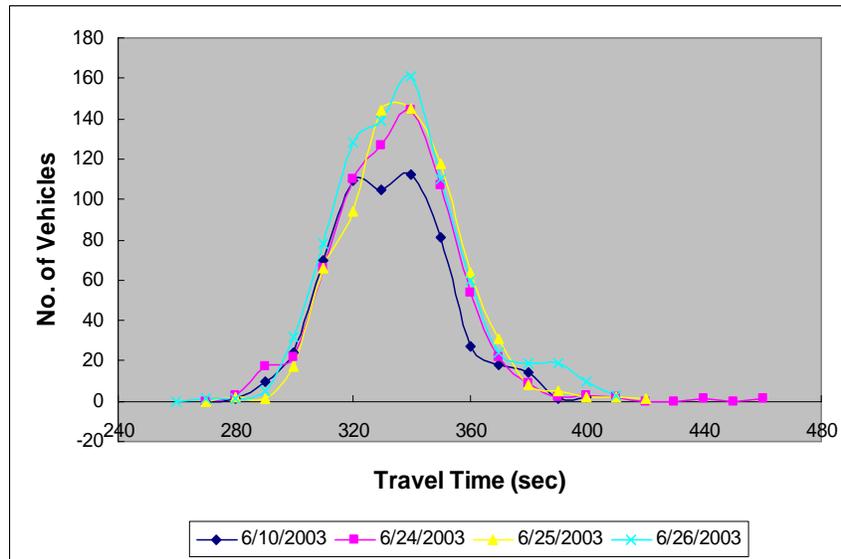
Figure 79. Covington Network Field Travel Time

Table 39. Field Travel Time for Covington Network

Covington	Mean Travel Time (sec)	Stdev	Size
06/10/03*	328.21	19.53	574
06/24/03	330.04	20.91	690
06/25/03	332.73	19.26	700
06/26/03	332.19	21.83	790

116

Travel Speed

Table 40 summarizes the statistics of travel speeds including mean, standard deviation and number of vehicles. The mean speed also shows a small variation over four days.

Table 40. Field Travel Speed for Covington Network

Covington	Mean Speed (mph)	Stdev.	Size
6/10/03	51.54	6.84	234*
6/24/03	52.9	5.65	405
6/25/03	52.33	5.76	634
6/26/03	51.42	6.83	695

* The data were collected between 5:10 to 6:00 and interrupted by the police for about 10 minutes.

Table 41 lists the statistics of vehicle free flow speed, which were filtered from the raw data of four days only when the headway between two continuous vehicles was greater than 4 seconds. The distributions of speed data on different days are shown in Figure 80 while the histogram of free flow speed is shown in Figure 81. For speed data, it should be noted that the distribution of the collected speeds underestimated field conditions as some high speeds in the fast lane were not recorded during data collection.

Table 41. Field Free Flow Speed for Covington Network

Covington	Mean Free Flow Speed (mph)	Stdev.	Size
6/10/03	51.67	6.71	109
6/24/03	53.30	5.97	240
6/25/03	52.98	5.96	282
6/26/03	52.12	6.60	272
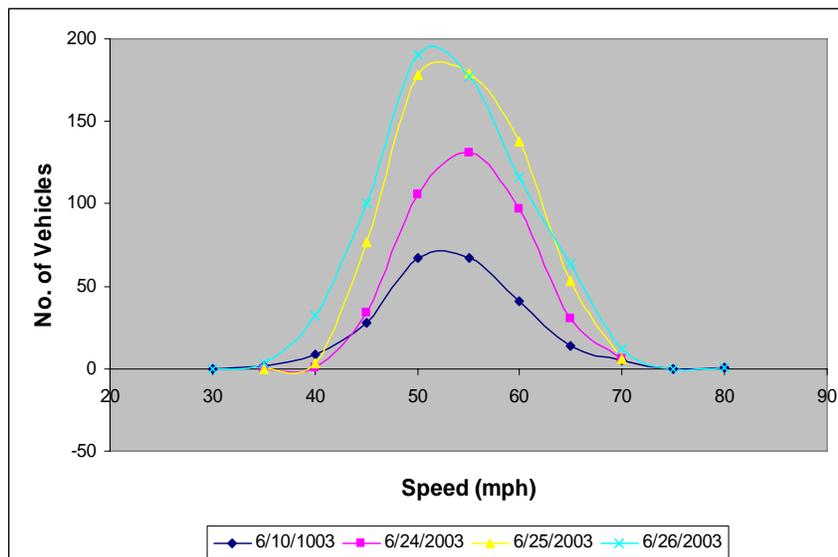
Figure 80. Covington Network Field Speed

117
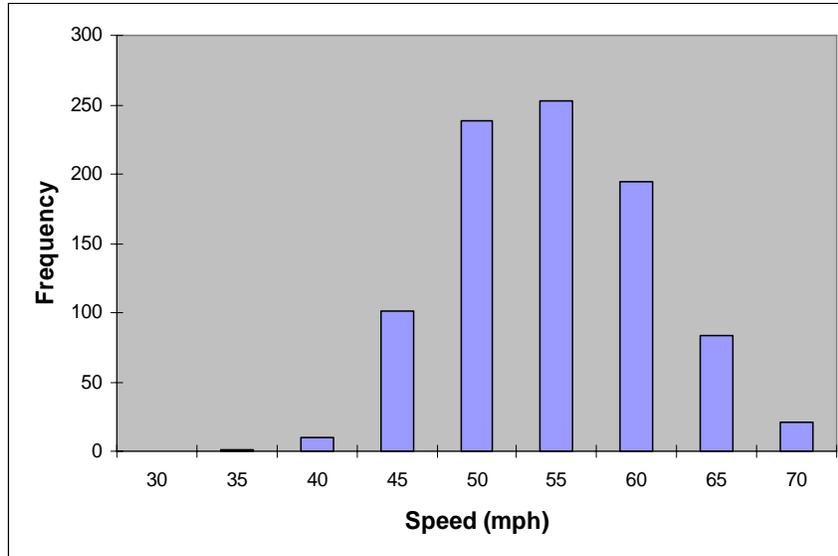
Figure 81. Covington Network Field Free Flow Speed

9.3. CORSIM Calibration

9.3.1. Identification of Calibration Parameters

Among many different calibration parameters, most of the parameters that are related to the car-following and lane changing behavior were selected. For those parameters having a distribution such as car-following sensitivity, several options were created to represent different driver behaviors. Parameters 1 to 3 set the link mean free flow speeds for links with different field posted speed limits. Index 1 represents 65 mph while index 3 represents 45 mph.

The initial parameters with acceptable ranges and distribution alternatives are listed as follows:
1) Link mean free flow speed 1 (mph): 65-70
2) Link mean free flow speed 2 (mph): 55-60
3) Link mean free flow speed 3 (mph): 45-50
4) Car Following Sensitivity:
   Default (1): 1.25, 1.15, 1.05, 0.95, 0.85, 0.75, 0.65, 0.55, 0.45, 0.35
   Small1 (2): 1.15, 1.05, 0.95, 0.85, 0.75, 0.65, 0.55, 0.45, 0.35, 0.25
   Large1 (3): 1.35, 1.25, 1.15, 1.05, 0.95, 0.85, 0.75, 0.65, 0.55, 0.45
   Narrow (4): 1.16, 1.08, 1.00, 0.92, 0.84, 0.76, 0.68, 0.60, 0.52, 0.44
   Wider (5): 1.34, 1.22, 1.10, 0.98, 0.86, 0.74, 0.62, 0.50, 0.38, 0.26
5) Pitt car following constant (feet): 3-10
6) Lag acceleration (second): 0.3 – 2.0
7) Lag deceleration (second): 0.3 – 1.5
8) Time to complete a lane-change maneuver (second): 1 – 4
9) Gap acceptance parameter: 1 – 6
10) Percent of drivers desiring to yield to merging vehicles (%): 5 – 30
11) Multiplier for desire to make a discretionary lane change: 0.1 – 0.9
12) Advantage threshold for discretionary lane change: 0.1 – 0.9

13) Minimum separation for generation of vehicles (sec): 0.2 – 2.0
14) Distribution of free flow speed by driver type
    Default: 88, 91, 94, 97, 99, 101, 103, 106, 109, 112
    Narrow (0.8): 90, 93, 95, 98, 99, 101, 102, 105, 107, 110
    Wide (1.2): 86, 89, 93, 96, 99, 101, 104, 107, 111, 114

9.3.2. Experimental Design for Calibration

The Latin Hypercube Design consisted of 200 cases using selected parameters and the values within acceptable ranges presented.

Ten randomly seeded runs were conducted in CORSIM for each of the 200 cases, for a total of 2,000 runs. The average travel time was recorded for each of the 2,000 runs and the results from the 10 multiple runs were then averaged to represent each of the 200 parameter sets.

9.3.4. Feasibility Test

The field data were compared with the distribution of simulation travel time output of 200 cases in Figure 82. The figure shows that all the travel times of simulation outputs are higher than the field data, which indicates that the initial ranges for the selected parameters were not sufficient to achieve the field conditions and needed adjustment.
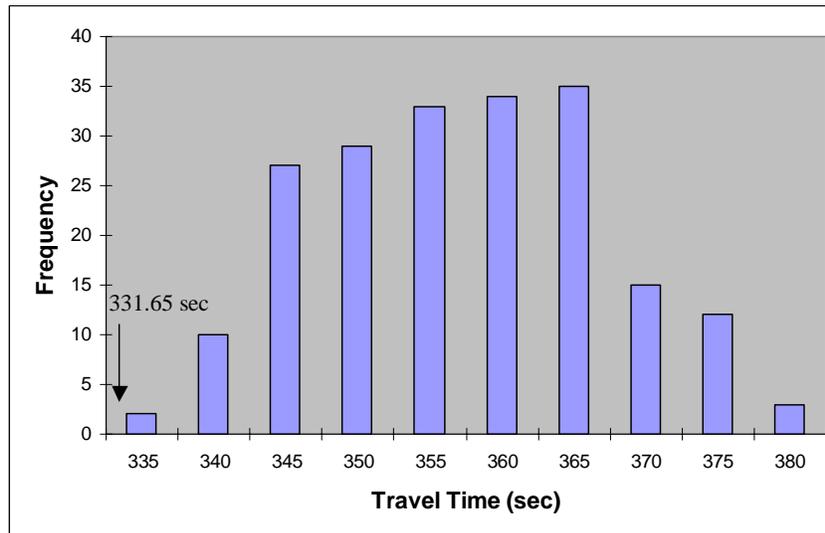
Figure 82. Feasibility Test Results for Covington Network with CORSIM

According to the X-Y plots in Figure 83 on the next page, the following two parameters showed a certain relationship with travel time output.
- Link mean free flow speed 3
- Link mean free flow speed 2
- Distribution of free flow speed

119
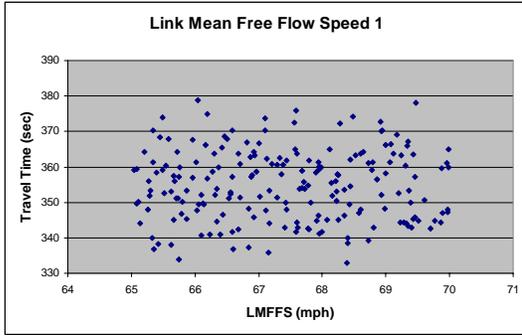
(a) Link Mean Free Flow Speed 1
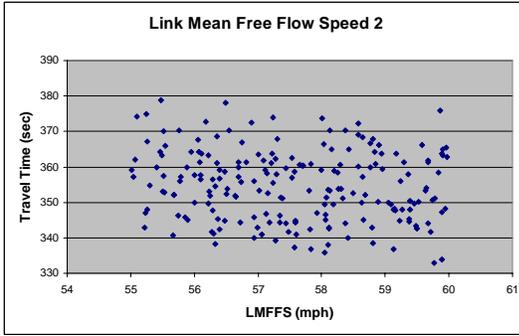
(b) Link Mean Free Flow Speed 2
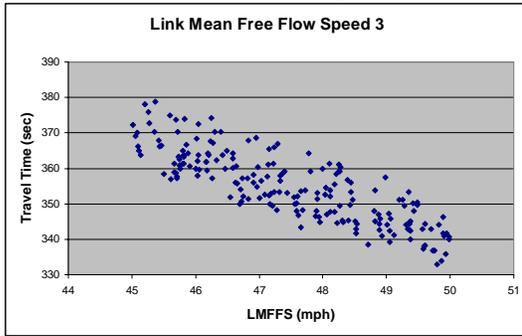
(c) Link Mean Free Flow Speed 3
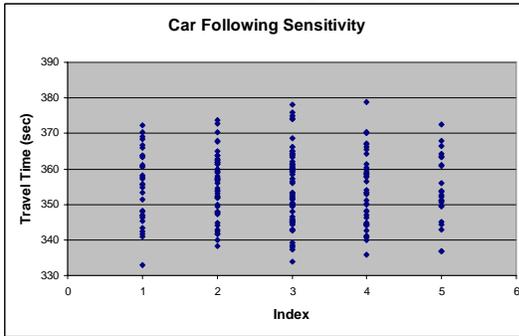
(d) Car Following Sensitivity
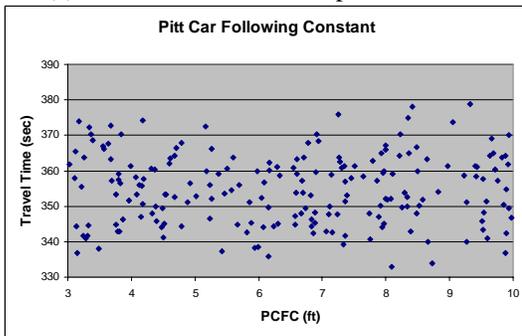
(e) Pitt Car Following Constant
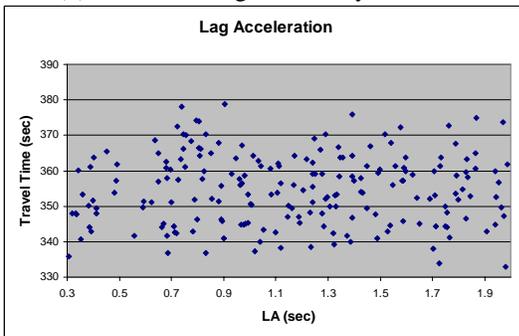
(f) Lag Acceleration
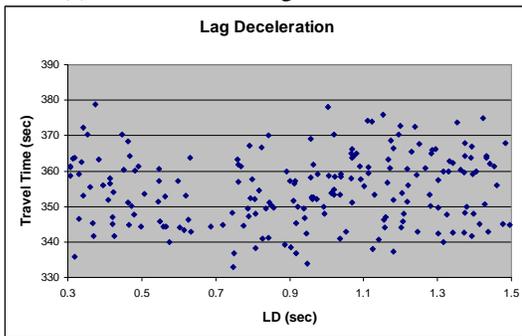
(g) Lag Deceleration
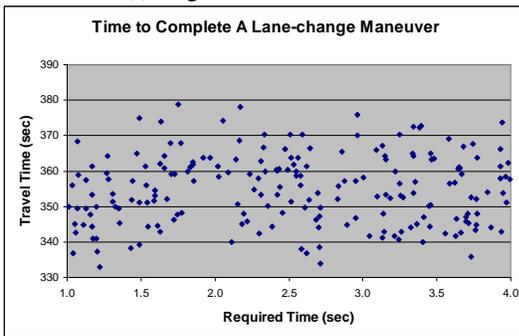
(h) Time to Complete A Lane-Change Maneuver

120
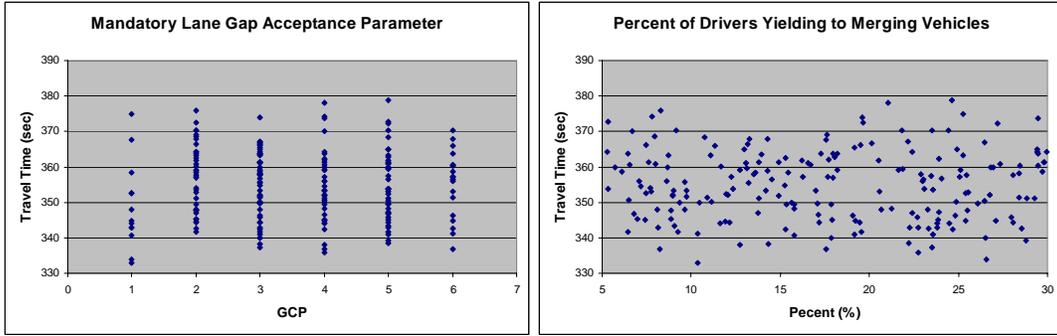
(i) Mandatory Lane Gap Acceptance Parameter

(j) Percent of Drivers Yielding to Merging Vehicles
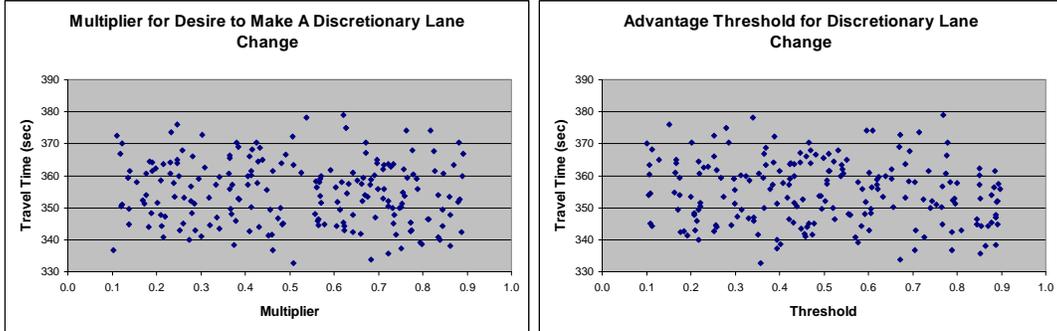
(k) Multiplier for Desire to Make
a Discretionary Lane Change

(l) Advantage Threshold for
Discretionary Lane Change
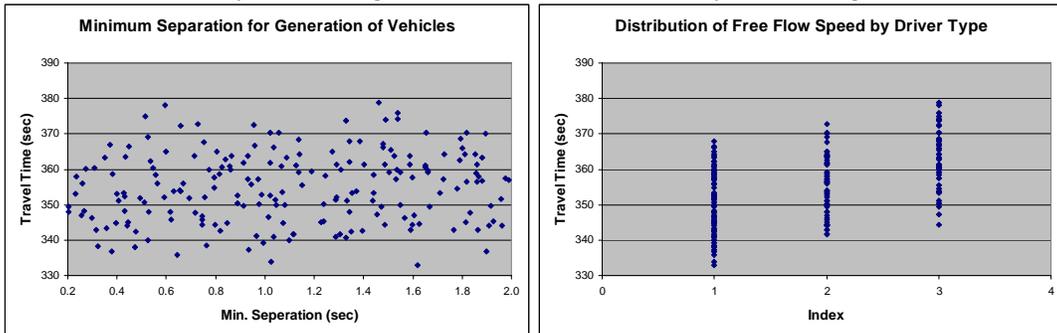
(m) Minimum Separation for
Generation of Vehicles

(n) Distribution of Free Flow Speed
by Driver Types

Figure 83. Feasibility Test Results for Covington Network with CORSIM

The result of statistical analysis is tabulated in Table 42. As seen from the table, two calibration parameters' p-value was less than 0.05, which is the criterion of the acceptance. Two parameters were also selected from previous X-Y plot analysis; however, the p-value of link mean free flow speed 2 was 0.085, which exceeded the maximum allowable value. Despite this exceeded value, link mean free flow was considered as a key parameter because of its relevance in X-Y plot analysis and relatively low p-value.

121

Table 42. ANOVA Results for Covington Network with CORSIM

Covington – CORSIM	p-value
Link mean free flow speed 1	0.858
Link mean free flow speed 2	0.085
Link mean free flow speed 3	**0.000**
Car following sensitivity	0.958
Pitt car following constant	0.222
Lag acceleration	0.683
Lag deceleration	0.102
Time to complete a lane-change maneuver	0.431
Gap acceptance parameter	0.963
Percent of drivers desiring to yield to merging vehicles	0.730
Multiplier for desire to make a discretionary lane change	0.665
Advantage threshold for discretionary lane change	0.366
Minimum separation for generation of vehicles	0.117
Distribution of free flow speed by driver type	**0.000**

In order to shift the simulated distribution to the left (decrease the simulation travel time output value) to capture the field travel time, key parameters such as mean free flow speeds need to be adjusted toward higher values. Finally, link mean free flow 2 and 3 were selected to be modified because it is definite that free flow speed affects the travel time dominantly. As a result, link mean free flow speed 2 and 3 were modified as shown here.

- Link mean free flow speed 2: 50~60 mph ➔ 57.5~62.5 mph
- Link mean free flow speed 3: 45~50 mph ➔ 47.5~52.5 mph

Two hundred combinations are generated again with modified parameter ranges by using Latin Hypercube Design and the result of 2,000 runs (10 random seeded runs for 200 combinations) is shown in Figure 84, in which field data fall within the acceptable range of the distribution. Therefore, the new parameters were considered to be able to produce the realistic traffic condition.
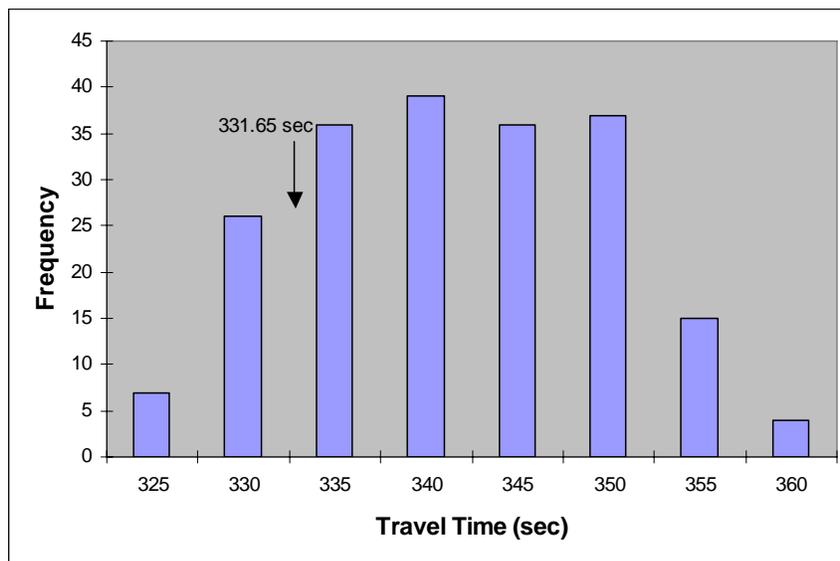
Figure 84. New Feasibility Test Result for Covington Network with CORSIM

9.3.5. Parameter Calibration Using Genetic Algorithm

A genetic algorithm (GA) was integrated with the CORSIM model to calibrate parameter values. 10 simulation runs were made for each trial within each generation and the output from those 10 CORSIM runs were averaged. 10 generations and 20 populations were adopted in the GA process and 2,000 runs (10 runs x 10 generation x 20 generations) were made in total. In order to get close to the field travel time, the following equation was used to calculate the fitness of simulation travel time outputs with a certain parameter setting.

$$\frac{\left| \text{Travel Time}_{Field} - \text{Travel Time}_{Simulation} \right|}{\text{Travel Time}_{Field}}$$

Convergence of calibration procedure (GA) is shown in Figure 85. The blue line represents the best fitness value (the value calculated by using the equation) that has been obtained from each generation and the pink line represents an average fitness value of each generation. As shown in the figure, the blue line converges toward the point zero, which means that the parameter set is being calibrated so that the parameter set alters toward the set that can reproduce field condition.
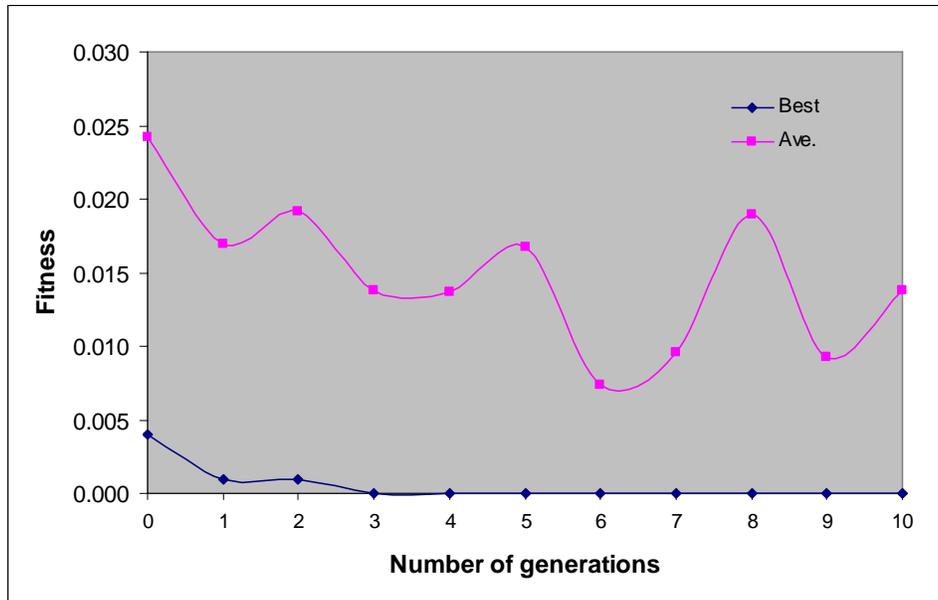
Figure 85. Convergence of Fitness Value with Generation

9.3.6. Evaluation of the Parameter Sets

To evaluate a calibrated parameter set, 100 runs for the model with default, LHD best (the parameter set that showed the closest match with field-measured travel time), and calibrated parameter set were made respectively. Table 43 lists these parameter values. The average travel time of those 100 runs are presented at the last row of Table 40. With the average travel time, it could be found that the model-calibrated parameter set outperformed the model with a default parameter set.

Table 43. Two Parameter Sets for Covington Network with CORSIM

Covington -CORSIM	Default	GA-Final
Link mean free flow speed 1 (mph)	65	66
Link mean free flow speed 2 (mph)	55	57
Link mean free flow speed 3 (mph)	45	51
Car following sensitivity Index	1	1
Pitt car following constant (ft)	10	3
Lag acceleration (sec)	0.3	1.2
Lag deceleration (sec)	0.3	0.5
Time to complete a lane-change maneuver (sec)	2.0	1.0
Gap acceptance parameter	3	4
Percent of drivers desiring to yield to merging vehicles (%)	20	20
Multiplier for desire to make a discretionary lane change	0.5	0.4
Advantage threshold for discretionary lane change	0.4	0.8
Minimum separation for generation of vehicles (sec)	1.6	1.3
Distribution of free flow speed by driver type Index	1	2
Ave. Travel Time (sec)	372.23	331.67

Figure 86 shows the same result with Table 43; however, it presents the travel time output of 100 runs for each parameter set with a histogram. It is clearer that the model with calibrated parameters outperformed the model with default parameters and all the field data fall within the distributions of calibrated models.
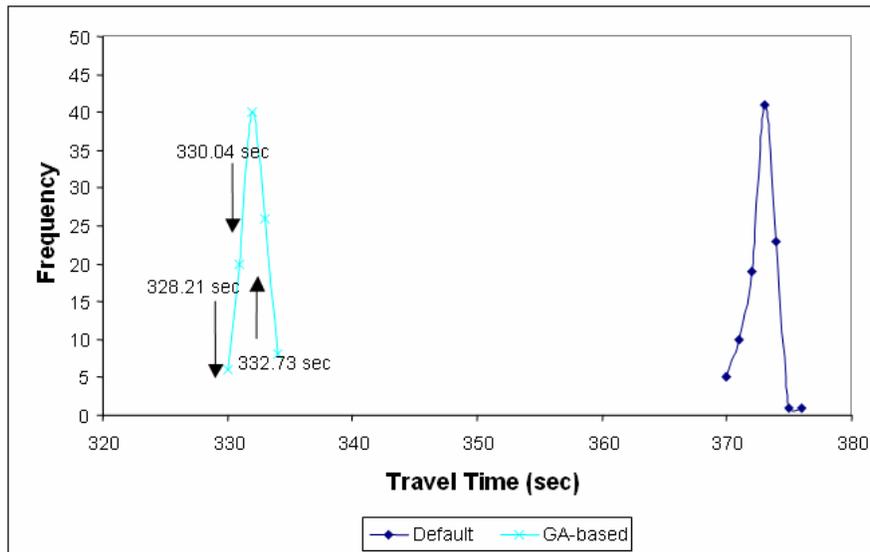
Figure 86. Comparison of Covington Network Travel Time with CORSIM

9.3.7. Validation

Another 100 runs were made (or the result that has been used for previous parameter evaluation can be used again) and compared with other field travel time data that have been saved for the validation procedure. As shown in Figure 87, the model with a calibrated parameter set was able to reproduce the field condition.
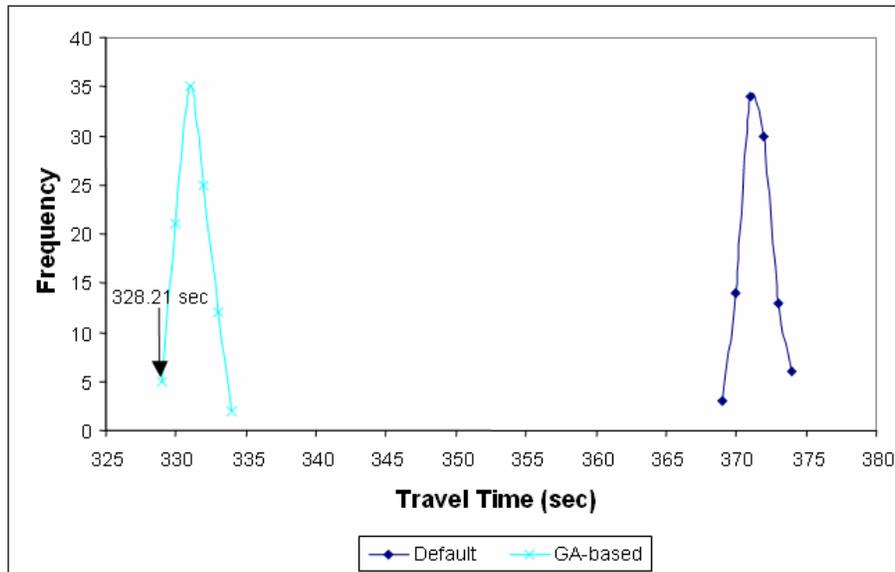
Figure 87. Validation of CORSIM Using Covington Network

9.4. VISSIM Calibration

9.4.1. Identification of Calibration Parameters

As stated previously in the site description part, this network is highway with work-zone. Also this network has three different speed limit zones due to the existence of the work-zone. They were indexed for the convenience of experimental design later. For instance, speed index 1 has six options to define the travel speed on the freeway where the posted speed limit is 65 mph. Because of the uniqueness of this network, three different desired speed distributions were included in the calibration parameter list. Also, the VISSIM model provides two different types of driving behavior model and according to the user manual, the Wiedemann 99 model is appropriate for this kind of highway network. Therefore, all calibration parameters in the Wiedemann 99 model are selected as a calibration parameter. Also some of the lane-changing parameters are selected due to the characteristics of this highway network.

The following is the initial set of parameters and acceptable ranges used in the calibration process:

1) Speed Index 1: 1 ~ 6
   (65.0~70.0, 62.5~72.5, 60.0~75, 67.5~72.5, 65.0~75.0, 62.5~77.5 mph)
2) Speed Index 2: 1 ~ 6
   (55.0~60.0, 52.5~62.5, 50.0~65.0, 57.5~62.5, 55.0~65.0, 52.5~67.5 mph)
3) Speed Index 3: 1 ~ 6
   (45.0~50.0, 42.5~52.5, 40.0~55.0, 47.5~52.5, 45.0~55.0, 42.5~57.5 mph)
4) Simulation Resolution: 1 ~ 9
5) Waiting time before diffusion (second): 30 ~ 90
6) Min. Headway (front/rear, meter): 0.1 ~ 0.9
7) Max. Deceleration (m/s^2): -5.00 ~ -1.00

125

8) Reduction Rate (meter per $1m/s^2$): 20 ~ 80
9) Accepted Deceleration (m/s^2): -3.0 ~ -0.2
10) Number of observed preceding vehicles: 1 ~ 5
11) Maximum look ahead distance (meter): 200 ~ 300
12) CC0 – average standstill distance (meter): 1.0 ~ 2.0
13) CC1 – headway at a certain speed (second): 0.5 ~ 3.0
14) CC2 – longitudinal oscillation (meter): 0.0 ~ 15.0
15) CC3 – start of the deceleration process (second): -30.0 ~ 0
16) CC4 – minimal closing Δv (m/s): -1.0 ~ 0.0
17) CC5 – minimal opening Δv (m/s): 0.0 ~ 1.0
18) CC6 – $\pm dv/dx$ (10-4 rad/s): 0.0 ~ 20.0
19) CC7 – car following activities $\pm b$ (m/s2): 0.0 ~ 1.0
20) CC8 – acceleration behavior when starting (m/s2): 1.0 ~ 8.0
21) CC9 – acceleration behavior at v ~ 80 km/h (m/s2): 0.5 ~ 3.0

9.4.2. Experimental Design for Calibration

Two hundred combinations of 21 calibration parameters within the initial parameter ranges were generated by using Latin Hypercube Sampling method.

Five random seeded runs were conducted in VISSIM for each of the 200 cases, for a total of 1,000 runs. The average travel time was recorded for each of the 1,000 runs.   The results from the five multiple runs were then averaged to represent the travel time output of each of the 200 parameter sets.

9.4.4. Feasibility Test

A feasibility test was conducted to check whether the simulated outputs based on the current parameter ranges could capture the field data and identify the key calibration parameters in VISSIM.

The travel time histograms of 200 cases are shown in Figure 88. The field travel time data (331.65 seconds) fall inside of the simulation travel time output distribution. They indicate that the selected parameters and their ranges were able to reproduce the field condition and can be directly used for further calibration procedure.
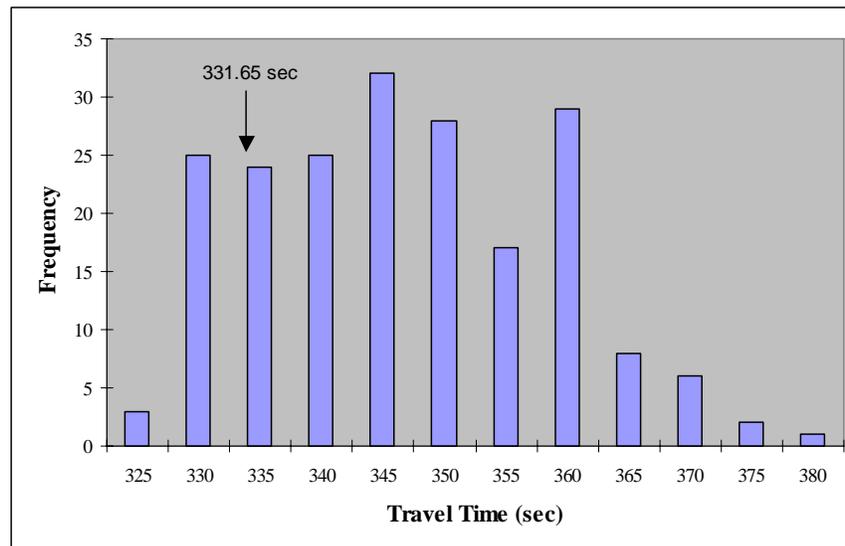
Figure 88. Feasibility Test Results for Covington Network with VISSIM

Even though the parameter adjustment procedure is not necessary, key parameters were found for the purpose of presentation. To identify the critical parameters, each parameter versus travel time from simulations was plotted and can be seen in Figure 89. An apparent trend was observed for the following calibration parameters in each X-Y plot:
- Speed index 3
- Average standstill distance, CC0
- Headway at a certain speed, CC1
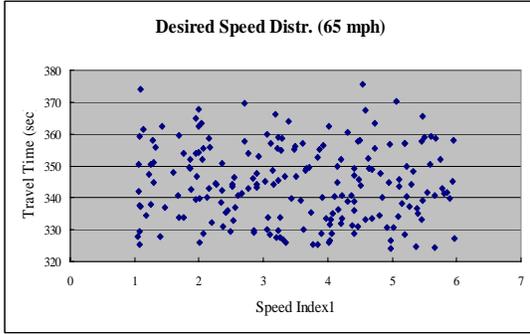
(a) Desired Speed Distribution (65mph)
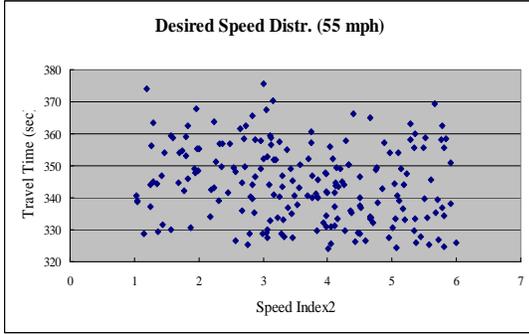
(b) Desired Speed Distribution (55mph)
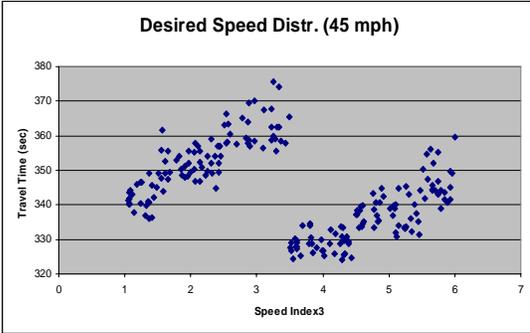
(c) Desired Speed Distribution (45mph)
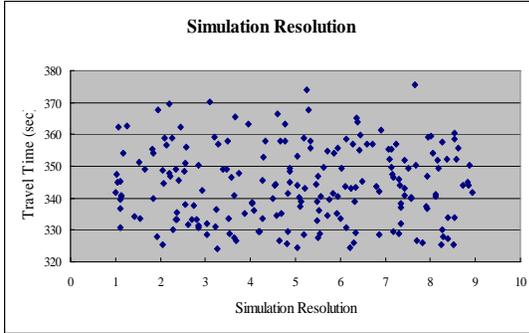
(d) Simulation Resolution
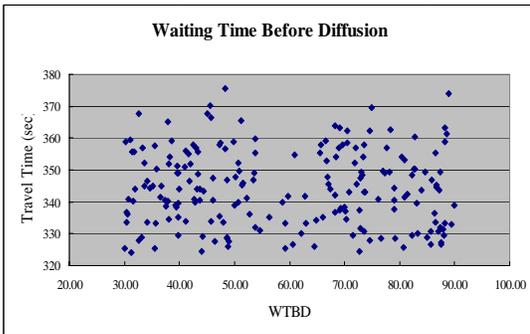
(e) Waiting Time before Diffusion
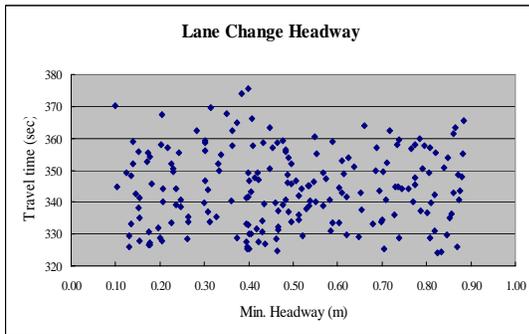
(f) Lane Changing Headway
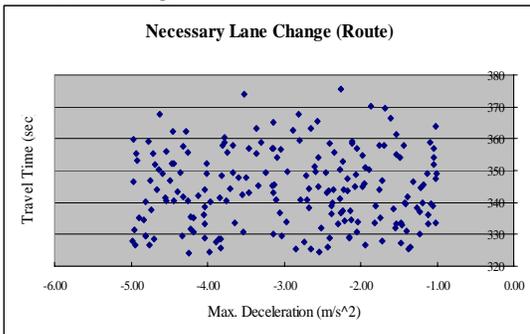
(g) Maximum Deceleration

(h) Reduction Rate

128
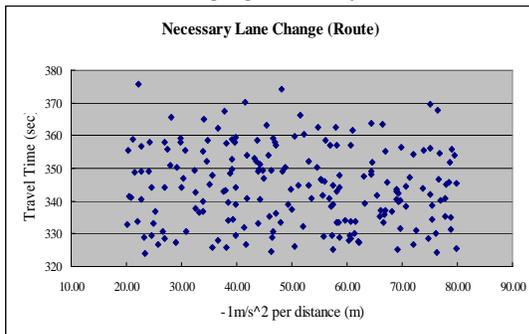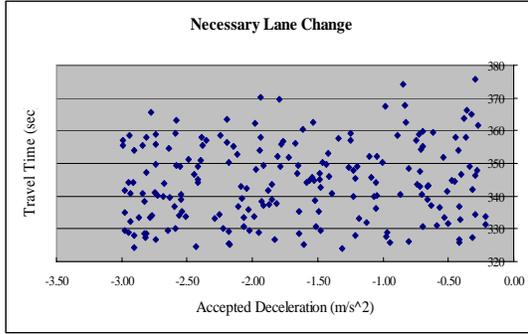
(i) Accepted Deceleration
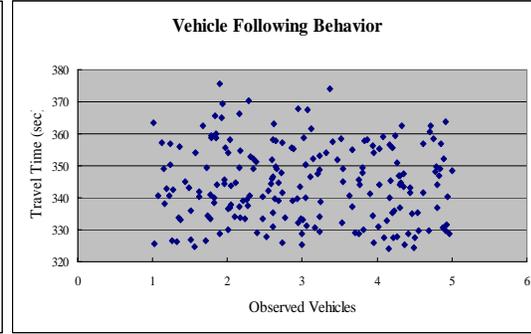
(j) Number of Observed Vehicles Ahead
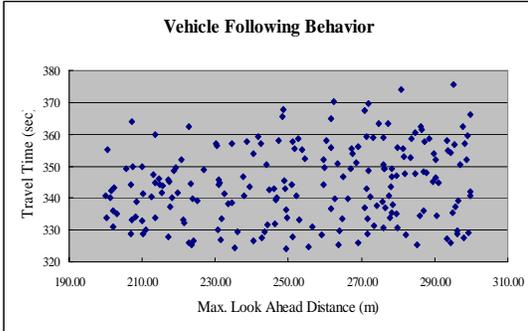
(k) Maximum Look Ahead Distance

(l) CC0

(m)CC1

(n) CC2

(o) CC3

(p) CC4

129

(q) CC5

(r) CC6

(s) CC7

(t) CC8

(u) CC9
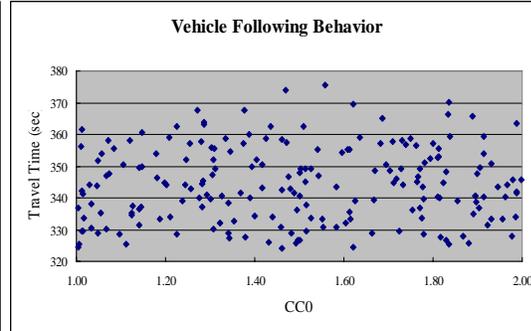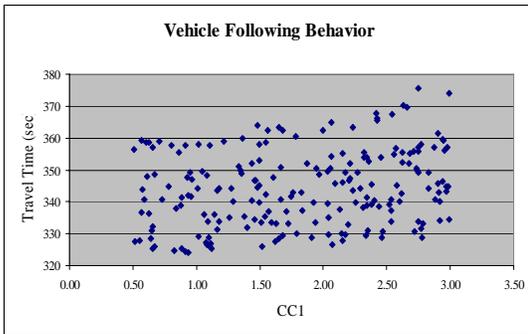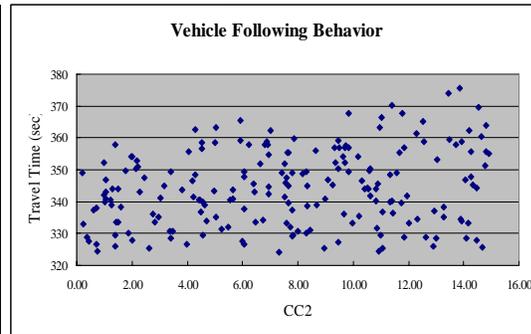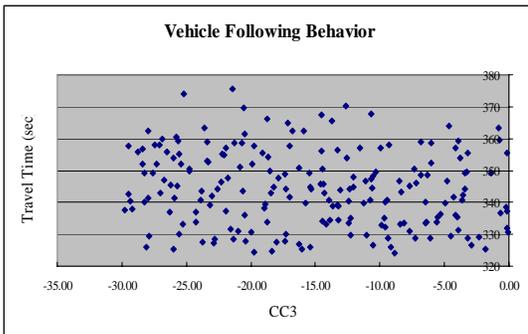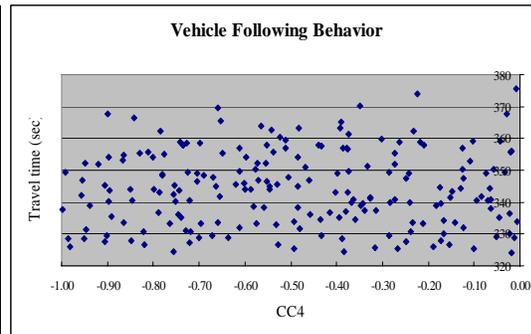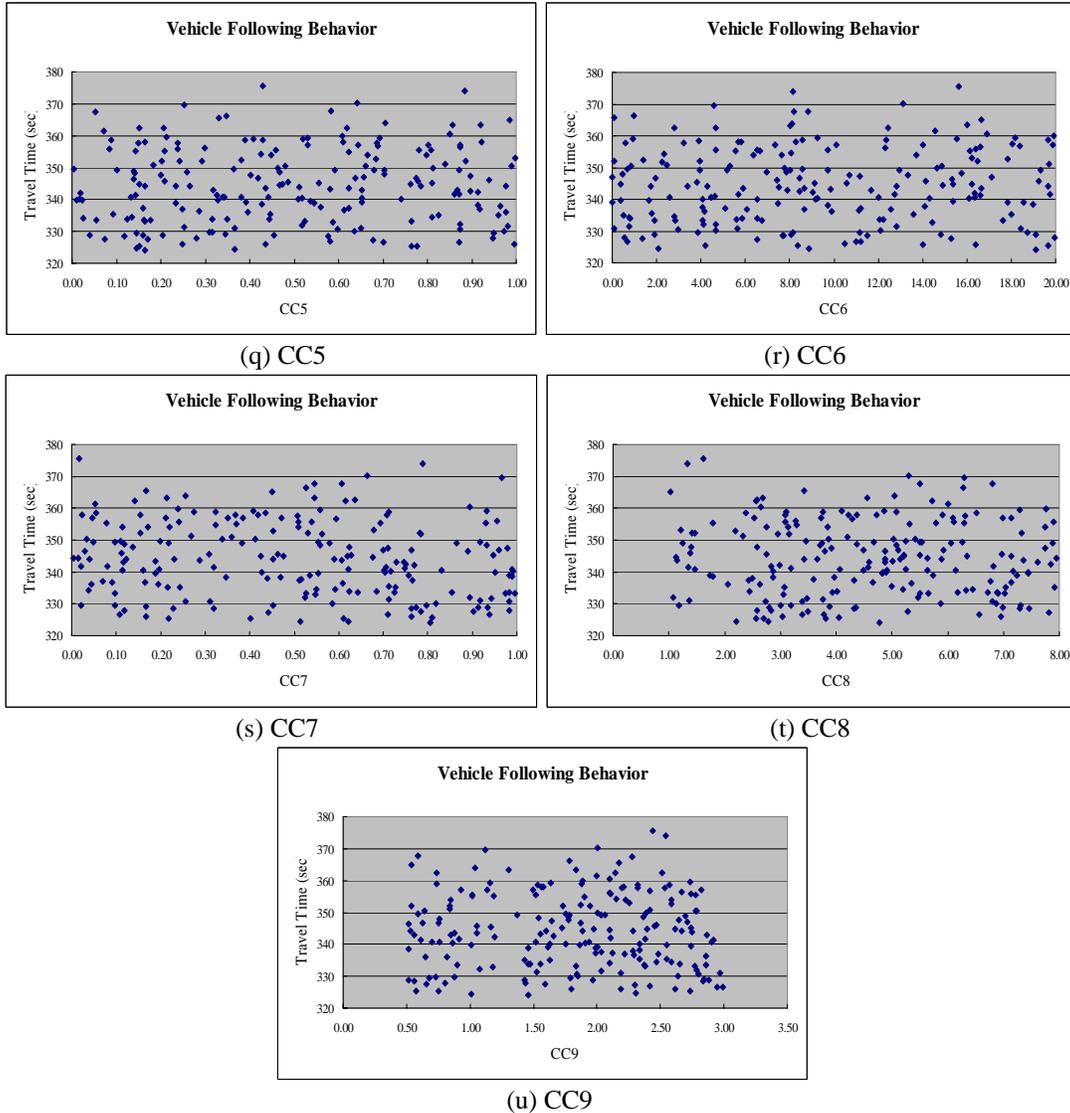
Figure 89. Feasibility Test Results for Covington Network with VISSIM

In addition, ANOVA was used to identify the key parameters in a statistical manner. Table 44 shows the statistical test analysis results. Since the significant value is 0.05, parameters with a p-value less than 0.05 were selected as a key calibration parameter and the following parameter showed a p-value less than 0.05:

- Speed Index 2
- Speed Index 3
- Look Ahead Distance
- Headway at a certain speed, CC1
- 'Following' variation, CC2
- Threshold for entering 'Following,' CC3
- Oscillation acceleration, CC7

Table 44. ANOVA Results for Covington Network with VISSIM

Covington –VISSIM	p-value	Covington –VISSIM	p-value
Speed Index 1	0.510	CC0	0.815
Speed Index 2	**0.035**	**CC1**	**0.016**
Speed Index 3	**0.000**	**CC2**	**0.011**
Simulation Resolution	0.809	**CC3**	**0.037**
Waiting Time Before Diffusion (sec)	0.757	CC4	0.335
Min. Headway (front/rear)	0.672	CC5	0.406
Max. Deceleration	0.621	CC6	0.321
-1m/s^2 per Distance	0.565	**CC7**	**0.022**
Accepted Deceleration	0.142	CC8	0.538
Observed Vehicles	0.757	CC9	0.582
Look Ahead Distance (max)	**0.011**		

* Significant value is less than 0.05.

The calibration parameters that have been selected from both X-Y plot and statistical analysis can be considered as independent key calibration parameters and other calibration parameters that have been selected from a single method can be considered as a parameter that has a joint effect with other parameters.

9.4.5. *Parameter Calibration Using Genetic Algorithm*

With the parameters and acceptable ranges identified in the previous step, Genetic Algorithm (GA) was integrated with VISSIM to calibrate the parameter set for the Covington network. In this case, GA calibration procedure was conducted twice to ensure the performance of the model with calibrated parameter set. The convergence graphs of two trials are shown in Figures 90 and 91. For each calibration, the number of generations and number of populations are determined to be 10 and 20, respectively. The blue line represents the parameter set that showed the best fitness that has been calculated with the following equation. Also, the pink line represents the average fitness value of all 20 populations. Based on the convergence curve that is presented here, it can be expected that the GA converges to the field value quickly at the beginning generations.

$$\frac{\left| \text{Travel Time}_{Field} - \text{Travel Time}_{Simulation} \right|}{\text{Travel Time}_{Field}}$$
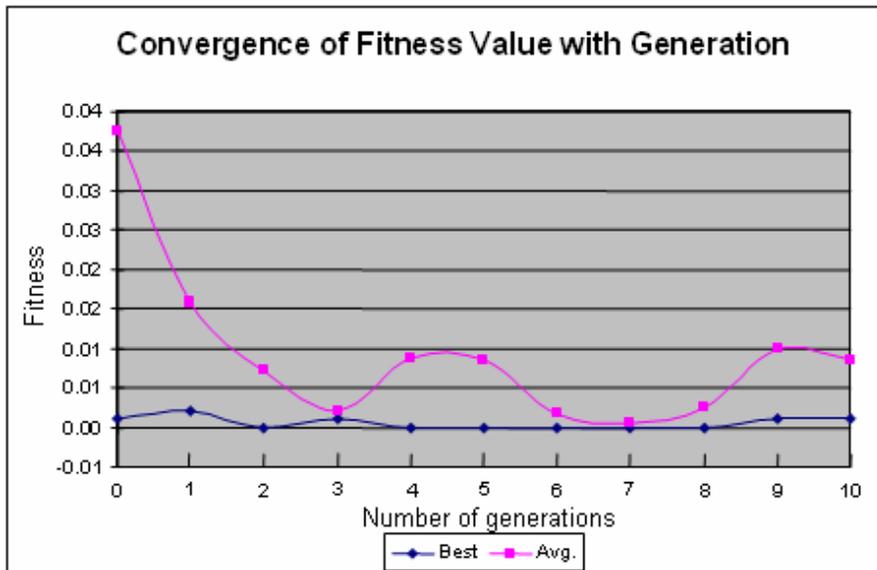
Figure 90. Convergence of Fitness Value with Generation
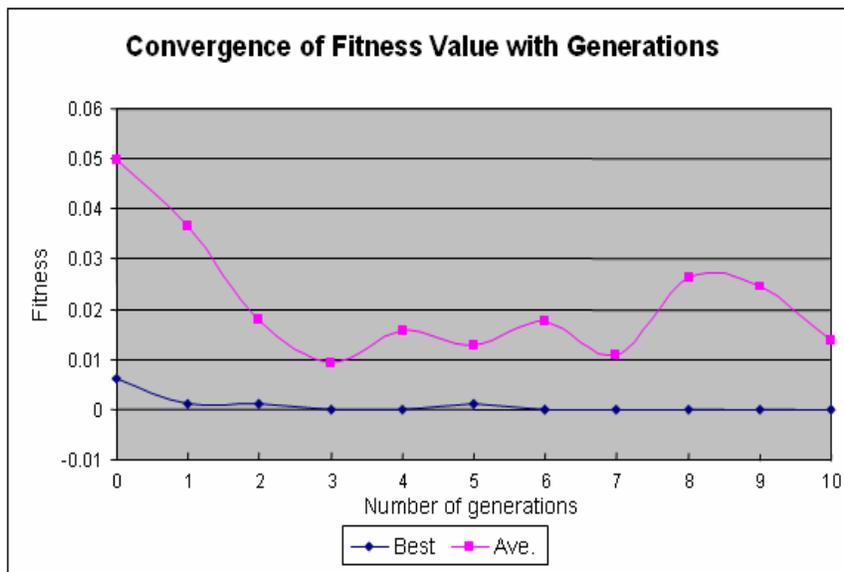
Figure 91. Convergence of Fitness Value with Generation

9.4.6. Evaluation of the Parameter Sets

For evaluation purposes, 100 randomly seeded runs were made for the model with default, LHD-based, and calibrated parameter set. The list of each parameter values for each parameter set and the average travel time with each parameter set is presented at the last row of Table 45. As can be seen from the table, the model with the default parameter set was away from the average value of field-measured travel time. However, the model with a calibrated parameter set was close to the average field-measured travel time.

The travel time output result of each model with a different parameter set is presented in Figure 92. The distribution of travel time output from calibrated parameter set included three different

field travel time values and the variation of the distribution was narrow as it was observed from the field. On the other hand, the distribution of travel time output with a default parameter set showed larger travel time than field-measured travel time. Animations of the calibrated VISSIM were viewed and deemed acceptable.

Table 45. Parameter Sets for Covington Network with VISSIM

Covington-VISSIM	Default	GA-based
Speed Index 1	62.5-67.5 mph	65.0-75.0 mph
Speed Index 2	52.5-57.5 mph	52.5-62.5 mph
Speed Index 3	42.5-47.5 mph	47.5-52.5 mph
Simulation Resolution	5	9
Waiting Time Before Diffusion (sec)	60.00	68.2
Min. Headway (front/rear)	0.50	0.72
Max. Deceleration	-3.00	-1.04
-1m/s^2 per Distance	50.00	60
Accepted Deceleration	-1.00	-0.7
Observed Vehicles	2	3
Look Ahead Distance (max)	250.00	257.58
CC0	1.50	1.74
CC1	0.90	2.77
CC2	4.00	4.09
CC3	-8.00	-0.91
CC4	-0.35	-0.97
CC5	0.35	0.86
CC6	11.44	10.7
CC7	0.25	0.67
CC8	3.50	2.06
CC9	1.50	2.77
Ave. Travel Time (sec)	358.27	332.767
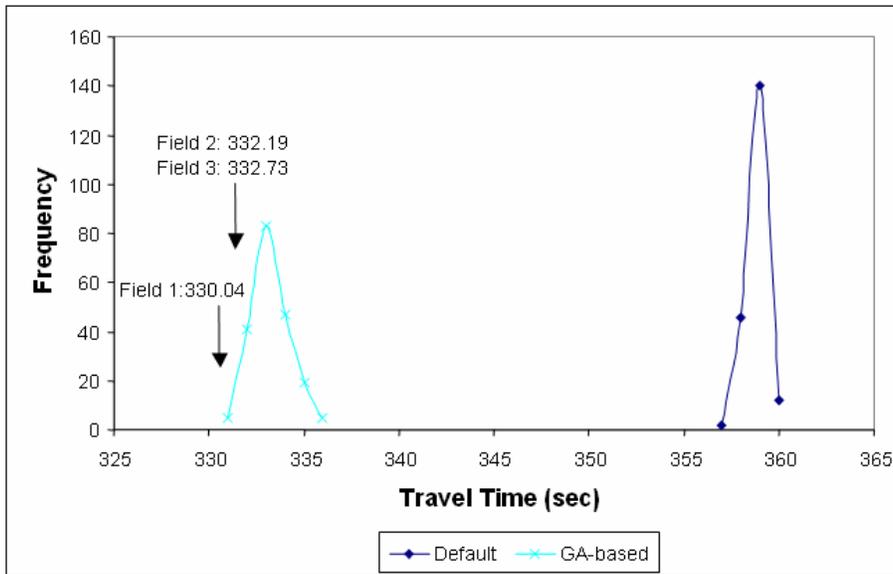
Figure 92. Covington Network Travel Time with VISSIM

9.4.7. Validation

Traffic data collected on a different day including fundamental traffic data and calibration data were used for the validation of parameter sets obtained from the calibration procedure. The result is shown in Figure 93. The field-measured travel time was compared to the distributions of 100 runs using three parameter sets (Default, LHD-based, and Calibrated). Field data were a bit outside (roughly 2 seconds) of the distributions of the calibrated parameters; however such a small difference could be ignored and considered an acceptable case. As a result, the distribution with a calibrated parameter set and LHD-based parameter set were much closer to the validation data (field measured travel time on the 4th day) than that of the distribution with the default parameter set.
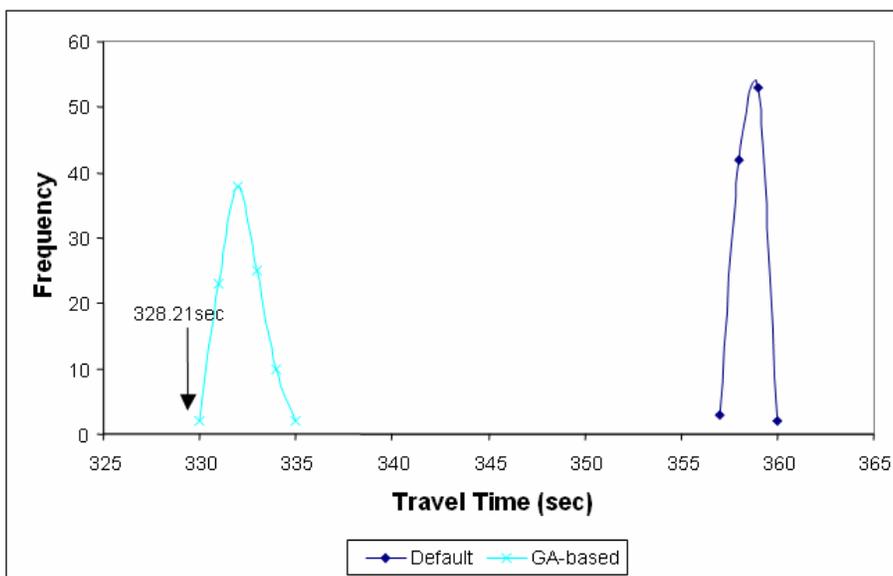
Figure 93. Validation of VISSIM Using Covington Network

Chapter 10
Highway Merging Section

In this Chapter:
- Case study of calibration and validation with highway merging section

---

10.1. Site Description

The test site selected in this case study was a merging segment located on Interstate Highway 66 in Arlington, Virginia. The site plies in the east-west direction and the direction of interest is eastbound toward the Washington D.C. area. The site is referred to as "Arlington network" in the remaining part of this case study. Especially, Glebe Rd. intersects Interstate Highway 66 toward the D.C. area. This site is unique because it contains a merging area within the network where the driving behavior may differ from basic segments. Mainline has two lanes and the single lane ramp section merges to the mainline in this area. The site can be considered as an important location in the Arlington area, Virginia because many vehicles that travel from Arlington downtown area to the D.C. area use this ramp to merge onto Interstate Highway 66. Posted speed limit in this area is 55 mph and, based on the preliminary site visit, most of the vehicles were traveling at a higher speed than the speed limit. Figure 94 shows an aerial photo of the site.

Figure 94. Aerial Photo of Arlington Network

10.2. Data Collection

10.2.1. Data Category

In order to conduct a calibration procedure, many different types of data needed to be collected from the field such as geometry, traffic count, heavy vehicle percentage, and performance measure data. The performance measure data collected for this case study were travel time for two sections including mainline and merging section and speed on the mainline entry point. Two types of performance measure data were selected because those were easier to collect from the field as well as the simulation model. Furthermore, those are known to directly reflect the field traffic condition.

10.2.2. Data Collection

Prior to the data collection step, the research team went to the site a couple of times to determine the locations where the data collection was easy and safe. As shown in Figure 95, two video cameras (B and C) were positioned at the entry and exit point of mainline and the smart travel van (STV) was parked at the shoulder of the ramp segment to record the ramp entry to collect ramp volume and heavy vehicle percentage as well as travel time data. Furthermore, one additional camera (A) was positioned at the location where a merging segment could be recorded and another AUTOSCOPE of STV was recording an overall ramp segment to record the vehicle's movement that entered the ramp segment.

Figure 95. Data Collection Location Map—Arlington

To account for day-to-day variability, multiple days of data sets were conducted. A group of four people collected data during the evening peak hour between 4:30 to 6:00 pm on August 10, 11, and 12 in 2005 from Wednesday to Friday. Three people were responsible for the video recording and two of them were also needed to measure speed data by using a radar gun. The last surveyor was positioned inside the STV. Synchronization between clocks of all the equipment was performed before data collection to ensure data quality.

10.2.3. Data Reduction

Recorded videotapes were reviewed by the research team to extract the traffic count, heavy vehicle percentage and time stamp of each vehicle that could be distinguished from other vehicles. Speed data were converted to MS Excel format for future use. For this case study, due to the high speed of passing vehicles, a vehicle description matching method was used to extract travel times instead of using a license plate matching method. The descriptions on each vehicle including brand, name, color, size, etc. were time stamped and recorded. After this step, the descriptions tried to be matched and the time stamp of exit and entry as well as ramp entry point were extracted to measure travel time of each vehicle.

10.2.4. Data Summary

Traffic Counts

Table 46 summarizes traffic counts on three individual days. Data collected on the last day were reserved for validation. Vehicles from ramp entry were entered from southbound and northbound of Glebe Rd. and left turn (LT) and right turn (RT) represent southbound and northbound traffic respectively.

Table 46. Field Traffic Volume and Composition for Arlington Network

		Location B			Ramp (LT)			Ramp (RT)		
		Car	**Bus**	**Total**	**Car**	**Bus**	**Total**	**Car**	**Bus**	**Total**
Calibration Data	2005/8/10	2,863	9	2,872	165	3	168	553	11	564
	2005/8/11	2,898	11	2,909	173	2	175	548	1	549
	Average	2,881	10	2,891	169	3	172	551	6	557
Validation Data	2005/8/12	2,765	9	2,774	133	2	135	478	0	478
	Average	2,765	9	2,774	133	2	135	478	0	478

Travel Time

Table 47 and Figure 96 show the travel time data of two sections: mainline and from ramp to mainline. As shown in the table and figure, the travel time of each section over three days was quite similar due to the characteristics of the network which is uninterrupted flow.

Table 47. Field Travel Time Data—Arlington Network

		Mainline	**Merging Traffic from Ramp**
Calibration Data	2005-08-10	24.8	N/A*
	2005-08-11	24.7	31.5
	Average	24.8	31.5
Validation Data	2005-08-12	25.4	32.0
	Average	25.4	32.0

*: Time stamp on the camera was missing on that day
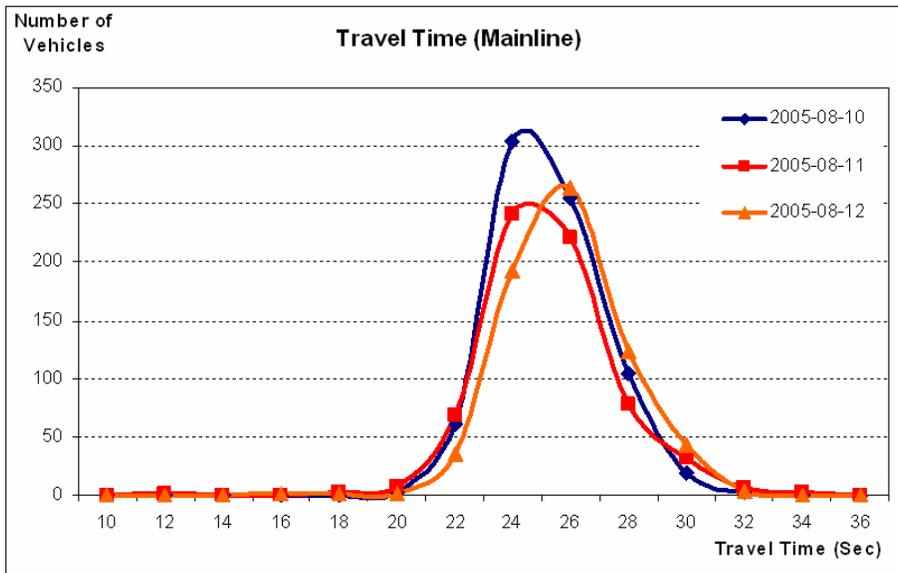
(a) Field Travel Times – Mainline
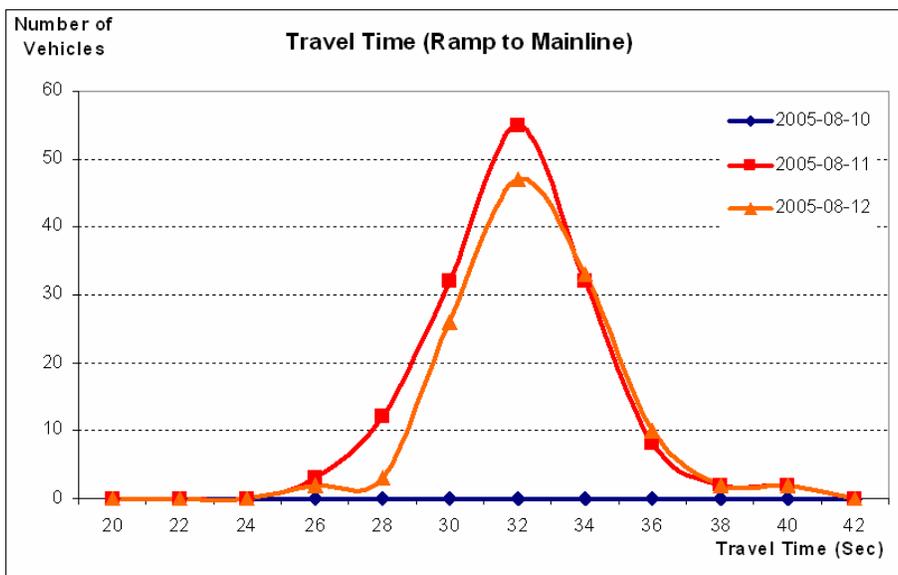
(b) Field Travel Times – Ramp Entry to Mainline

Figure 96. Field Travel Times—Arlington Network

Travel Speed

Table 48 summarizes the statistics of travel speeds including average, standard deviation, minimum and maximum value. Overall speed statistics on the last day were slightly lower than the other two days.

Table 48. Field Speed for Arlington Network

	Calibration Data		Validation Data
	2005/8/10	**2005/8/11**	**2005/8/12**
Average (mph)	57.2	57.5	56.9
Minimum (mph)	41.0	41.0	36.0
Maximum (mph)	76.0	77.0	73.0
Standard Deviation	4.51	4.65	4.70

The distributions of speed data on different days are shown in Figure 97. Based on the distribution, the speed distribution over three days was very similar and did not have a significant difference. Note that the speed data were collected from both lanes on the mainline.
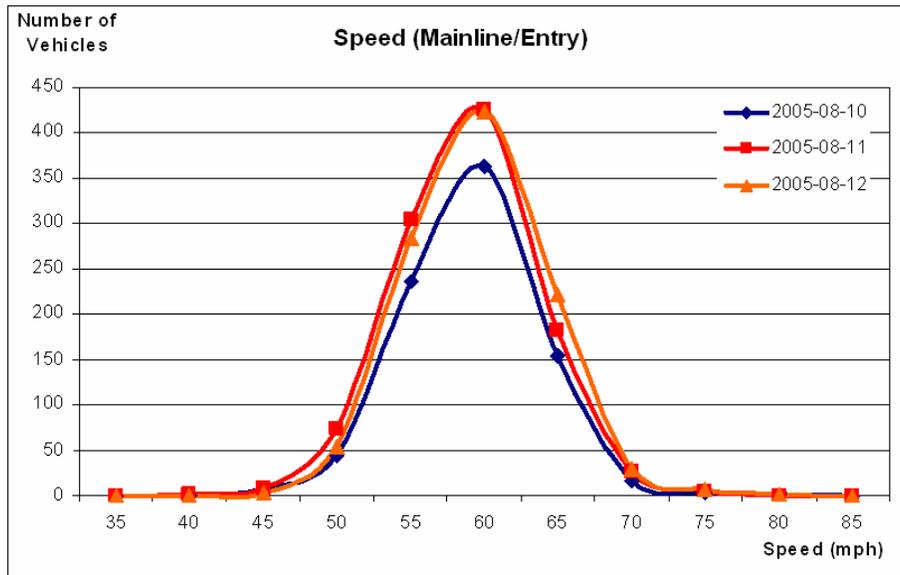
Figure 97. Field Speed Distribution for Arlington Network

10.3. CORSIM Calibration

10.3.1 Evaluation with Default Parameter Set

100 random seed runs with default calibration parameters were conducted to verify the feasibility of the default model. As shown in Figure 98, travel time of mainline falls within the distribution; however, it fell by the margin of the distribution. Furthermore, the distribution of travel time for the ramp section did not include field-measured travel time data. Therefore, a further procedure of calibration is considered to be required.
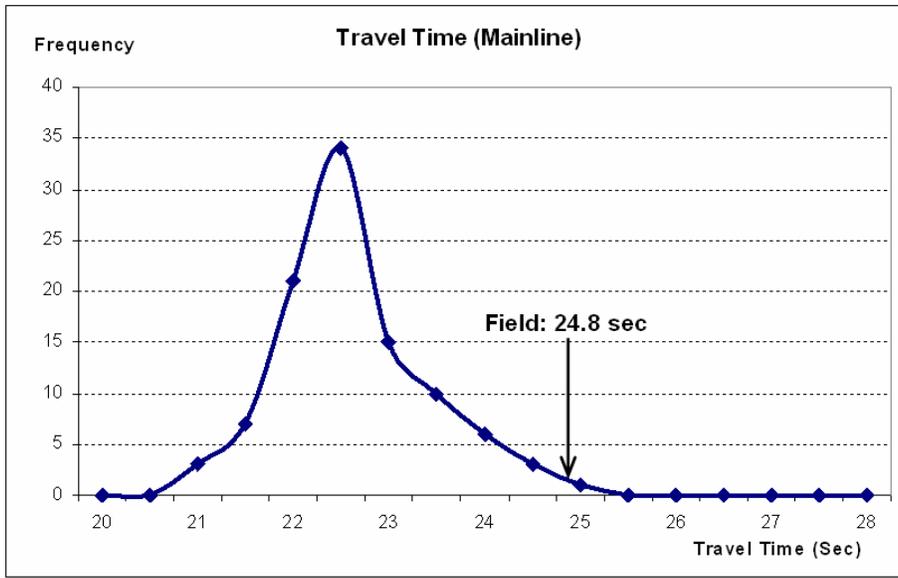
(a) Travel Time (Mainline)
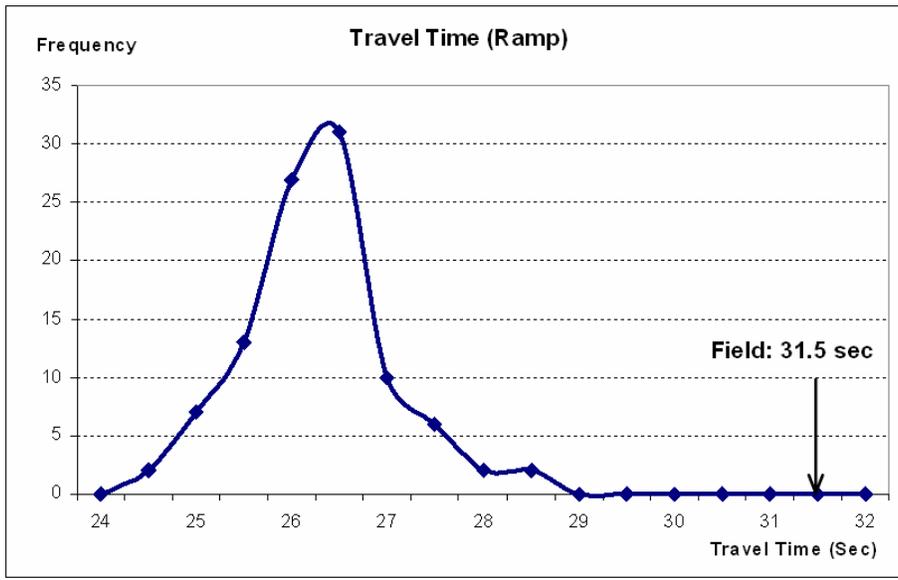
(b) Travel Time (Ramp)

Figure 98. Distribution of Travel Time Output with Default Parameter Set—CORSIM

10.3.2. Identification of Calibration Parameters

Since this is a highway network, almost every calibration parameter that has been selected from 9.3 was selected which was related to the car-following and lane changing behavior. For the calibration parameters that needed to have distribution. they were created as several scenarios such as small, large, and default. Link mean free flow speed range was determined based on the field-collected speed data and used as the first calibration parameter.

The initial parameters with acceptable ranges and distribution alternatives are listed as follows:

1) Link mean free flow speed 1 (mph): 45-70
2) Car Following Sensitivity:
        Default (1): 1.25, 1.15, 1.05, 0.95, 0.85, 0.75, 0.65, 0.55, 0.45, 0.35
        Small1 (2): 1.15, 1.05, 0.95, 0.85, 0.75, 0.65, 0.55, 0.45, 0.35, 0.25
        Large1 (3): 1.35, 1.25, 1.15, 1.05, 0.95, 0.85, 0.75, 0.65, 0.55, 0.45
3) Pitt car following constant (feet): 3-10
4) Lag acceleration (second): 0.3 – 2.0
5) Lag deceleration (second): 0.3 – 1.5
6) Time to complete a lane-change maneuver (second): 1 – 4
7) Gap acceptance parameter: 1 – 6
8) Percent of drivers desiring to yield to merging vehicles (%): 5 – 30
9) Multiplier for desire to make a discretionary lane change: 0.1 – 0.9
10) Advantage threshold for discretionary lane change: 0.1 – 0.9
11) Minimum separation for generation of vehicles (sec): 0.2 – 2.0
12) Distribution of free flow speed by driver type
        Default: 88, 91, 94, 97, 99, 101, 103, 106, 109, 112
        Narrow (0.8): 90, 93, 95, 98, 99, 101, 102, 105, 107, 110
        Wide (1.2): 86, 89, 93, 96, 99, 101, 104, 107, 111, 114

10.3.2. Experimental Design for Calibration

Two hundred combinations of each calibration parameter values were generated by using LHD based on the determined parameter ranges.

Ten random seeded runs were conducted for 200 combinations, for a total of 2,000 runs and an average travel time for each run was stored.

10.3.4. Feasibility Test

Simulation outputs obtained from the previous step were compared with field-collected travel time data. As shown in Figure 99, both travel time output distributions were able to include field-collected travel time data. Therefore, the set of parameter ranges that have been determined in the previous step can be considered as feasible ranges and used for further calibration procedure.
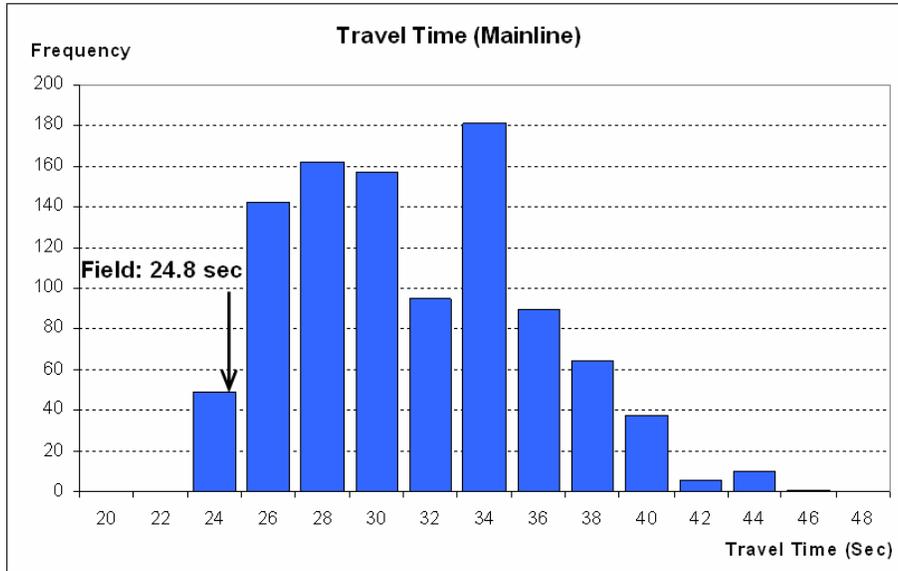
(a) Travel Time – Mainline
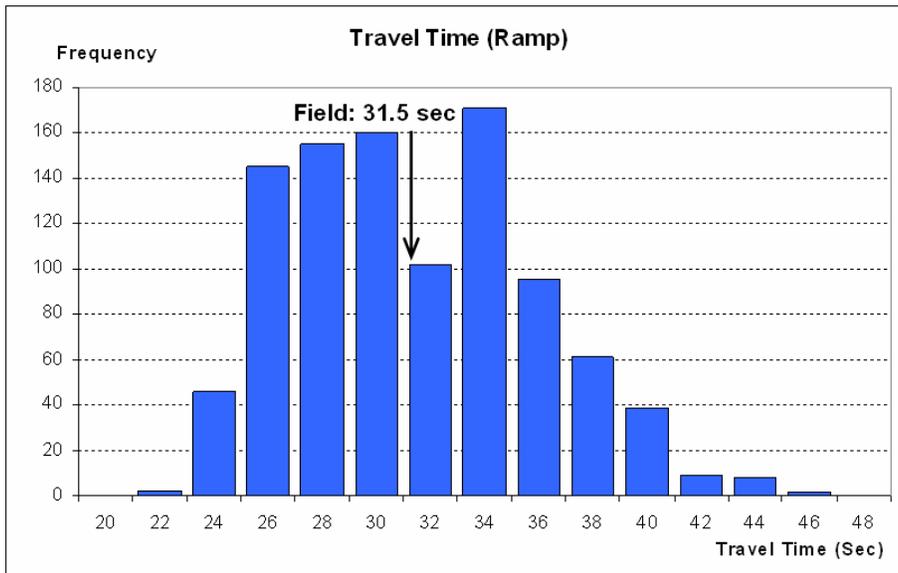
(b) Travel Time—Ramp

Figure 99. Feasibility Test Results for Arlington Network with CORSIM

10.3.5. Parameter Calibration Using Genetic Algorithm

A genetic algorithm (GA) was integrated with the CORSIM simulation model to conduct the calibration procedure. 10 simulation runs were made for each population and 10 populations were made for each generation. GA was able to obtain the best solution when it reached 15th generation and the GA was terminated after 18th generation to ensure that the current optimal solution was the best solution.

$$\frac{\left| \text{Travel Time}_{Field} - \text{Travel Time}_{Simulation} \right|}{\text{Travel Time}_{Field}}$$

142

Convergence of calibration procedure (GA) is shown in Figure 100. The blue and pink lines represent the best and average fitness value of each generation. As shown in the figure, the blue line converges toward the point zero, which means that the parameter set is being calibrated so that the parameter set alters toward the set that can reproduce the field condition.
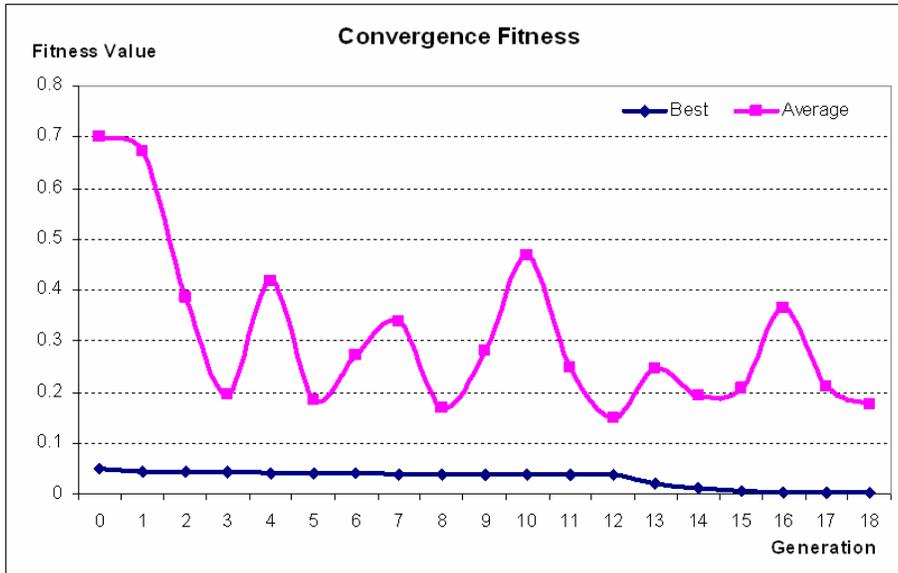
Figure 100. Convergence of Fitness Value with Generation

10.3.6. Evaluation of the Parameter Sets

In order to evaluate the calibrated model, 100 random seeded runs were conducted and travel time data of two sections were recorded to be used as a performance measure. After completing multiple runs, the distributions of travel time outputs from simulation runs were compared with field-collected travel time data. Table 49 shows two parameter sets: default and calibrated with GA optimization program.

As shown in Figure 101, both distributions of travel time outputs could include the field-collected travel time data. Therefore, the parameter set can be considered as a well-calibrated parameter set.

Table 49. Parameter Sets for Arlington Network with CORSIM

CORSIM	Default	Calibrated
Desired free flow speed (Main), mph	65	62
Car following sensitivity, %	100	100
New car-following sensitivity factor, sec	1.25, 1.15, 1.05, 0.95, 0.85, 0.75, 0.65, 0.55, 0.45, 0.35	1.30, 1.20, 1.10, 1.00, 0.90, 0.80, 0.70, 0.60, 0.50, 0.40
New value for Pitt car following constant, ft	10	4
Time to complete a lane change maneuver, sec	2	1.3
Mandatory lane change gap acceptance parameter	3	2
% of drivers desiring to yield right-of-way to lane changing vehicles attempting to merge, %	20	30
Multiplier for desire to make a discretionary lane change	0.5	0.3
Advantage threshold for discretionary lane change	0.4	0.7
Leader's maximum deceleration perceived by follower, 0.1 ft/sec^2	15	18
Free-Flow Speed Adjustment (for each driver type), %	88, 91, 94, 97, 99, 101, 103, 106, 109, 102	82, 86, 90, 94, 98, 102, 106, 110, 114, 118

(a) Travel Time—Mainline

144
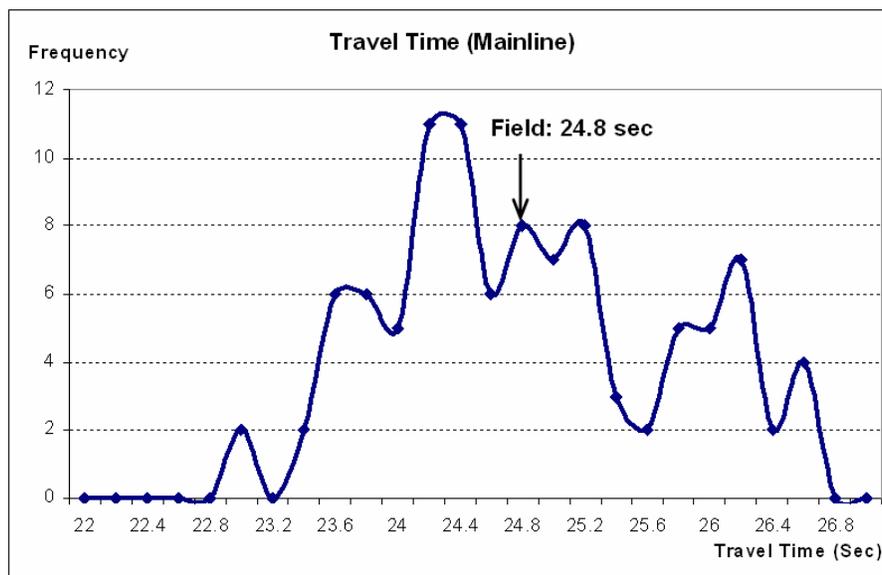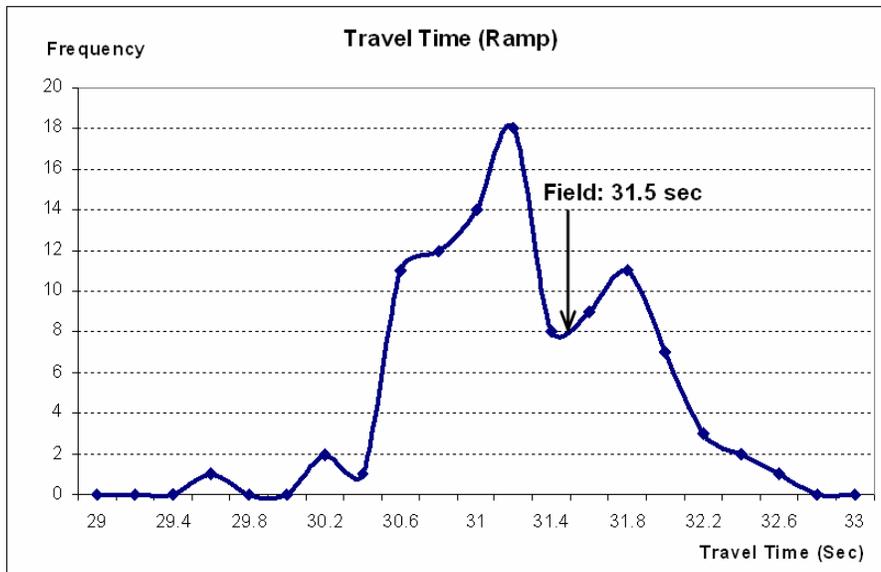
(b) Travel Time—Ramp

Figure 101. Comparison of Covington Network Travel Time with CORSIM

10.3.7. Validation

An additional 100 runs were made with a completely new set of data and compared with field-collected validation data. As shown in Figure 102, both field-collected travel time data fell within the distribution of travel time outputs. Therefore, the validation step of the calibrated model is completed and the model can be used for further analysis purposes with a high confidence level.
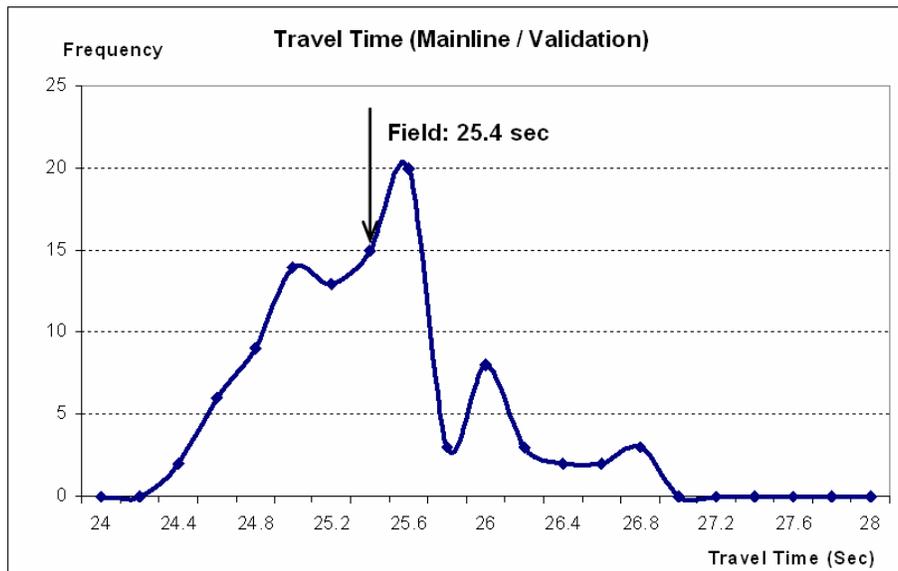
(a) Travel Time—Mainline
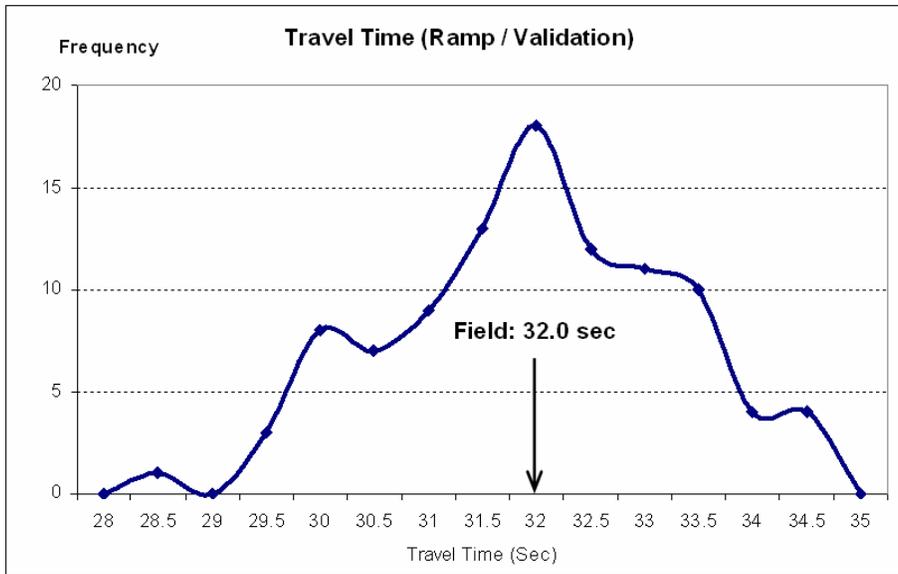
(b) Travel Time—Ramp

Figure 102. Validation of CORSIM Using Arlington Network

10.4. VISSIM Calibration

10.4.1. Evaluation with Default Parameter Set

One hundred random seeded runs were conducted to test the model with a default calibration parameter set and average travel time outputs of two sections were recorded to be used as a performance measure. As shown in Figure 103, both distributions of travel time outputs were unable to include the field-measured travel time, so a further calibration procedure needs to be conducted.
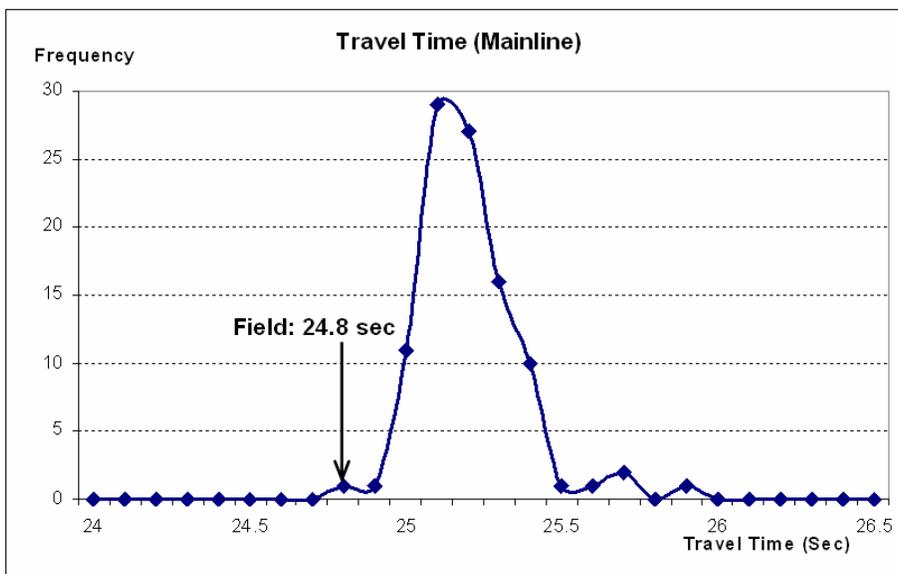
(a) Travel Time (Mainline)
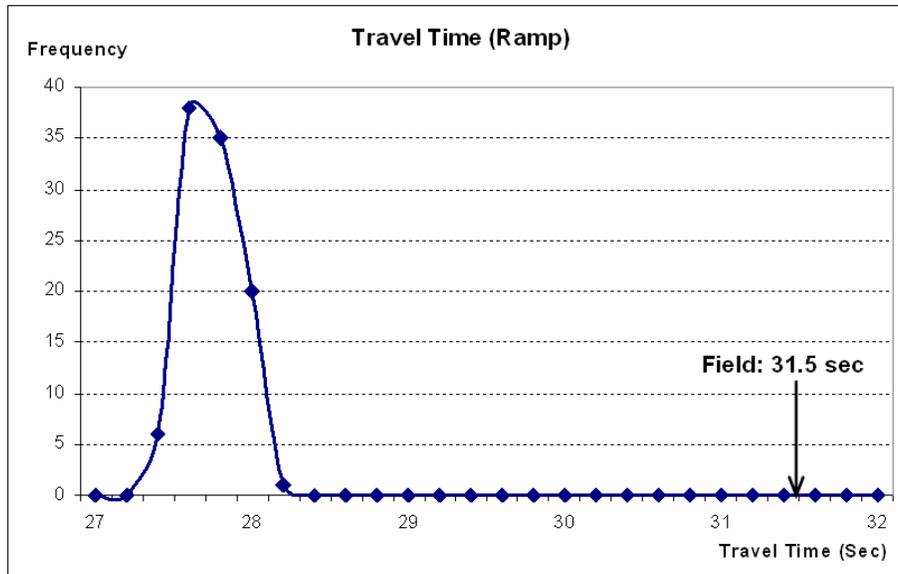
(b) Travel Time (Ramp)

Figure 103. Distribution of Travel Time Output with Default Parameter Set—VISSIM

10.4.2. Identification of Calibration Parameters

As aforementioned, this network is a highway-merging section and, according to the manual, the Wiedemann 99 model should be used for the highway network. Therefore, all calibration parameters for the Wiedemann 99 model were selected as well as lane-changing parameters. In addition to those calibration parameters, three different desired speed distributions were selected as a calibration parameter and the distributions are separated into two parts, minimum value and variance, to test more possible combinations.

The following is the initial set of parameters and acceptable ranges used in the calibration process:
1) Desired Speed Distribution 1 (mph / Mainline / Minimum): 40 ~ 50
2) Desired Speed Distribution 1 (mph / Mainline / Variance): 15 ~ 40
3) Desired Speed Distribution 2 (mph / Ramp / Minimum): 20 ~ 35
4) Desired Speed Distribution 2 (mph / Ramp / Variance): 15 ~ 30
5) Desired Speed Distribution 3 (mph / Ramp Entry / Minimum): 10 ~ 30
6) Desired Speed Distribution 3 (mph / Ramp Entry / Variance): 10 ~ 25
7) Maximum Look Ahead Distance (m): 150 ~ 350
8) Number of Observed Preceding Vehicles (veh): 1 ~ 4
9) CC0 – average standstill distance (meter): 0.5 ~ 7.0
10) CC1 – headway at a certain speed (second): 0.5 ~ 1.5
11) CC2 – longitudinal oscillation (meter): 0.0 ~ 15.0
12) CC3 – start of the deceleration process (second): -30.0 ~ 0
13) CC4 – minimal closing Δv (m/s): -1.0 ~ 0.0
14) CC5 – minimal opening Δv (m/s): 0.0 ~ 1.0
15) CC6 – $\pm dv/dx$ (10^{-4} rad/s): 0.0 ~ 25.0
16) CC7 – car following activities $\pm b$ (m/s2): 0.0 ~ 1.0

147

17) CC8 – acceleration behavior when starting (m/s2): 1.0 ~ 8.0
18) CC9 – acceleration behavior at v ~ 80 km/h (m/s2): 0.5 ~ 3.0
19) Maximum Deceleration (m/s^2): -5.00 ~ -1.00
20) Reduction Rate (meter per 1m/s^2): 20 ~ 80
21) Accepted Deceleration (m/s^2): -3.0 ~ -0.2
22) Waiting Time Before Diffusion (second): 30 ~ 60
23) Minimum Headway (m): 0.1 ~ 1.0

10.4.3. Experimental Design for Calibration

Two hundred combinations of 23 calibration parameters within initially determined parameter ranges presented were generated by using LHD.

Five random seeded runs were conducted in VISSIM for each of the 200 cases, for a total of 1,000 runs. The average travel time was recorded for each of the 1,000 runs to be used as a performance measure.

10.4.4. Feasibility Test

A feasibility test step was conducted to check the validity of initially determined parameter ranges. As shown in Figure 104, field travel time for both sections fall within the distribution of each distribution of simulation travel time output. It indicates that the selected parameters and their ranges were able to reproduce the field condition and can be directly used for further calibration procedure.
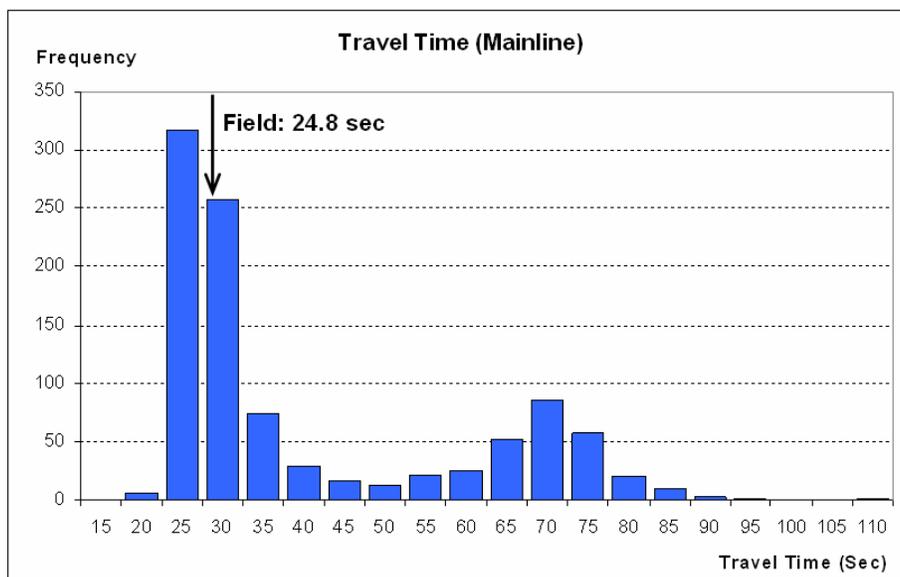
(a) Travel Time (Mainline)
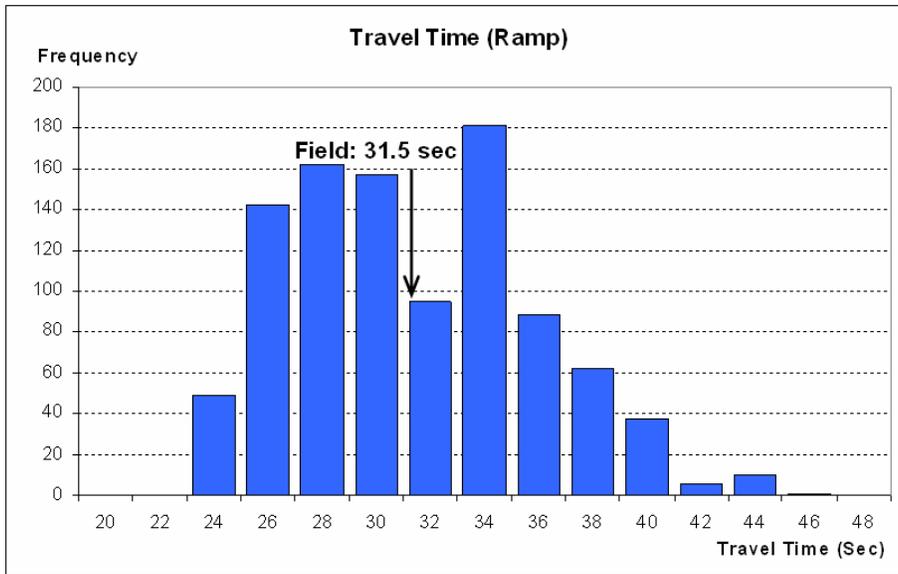
(b) Travel Time (Ramp)

Figure 104. Feasibility Test Results for NOVA Freeway Network with VISSIM

10.4.5. Parameter Calibration Using Genetic Algorithm

With the parameters and acceptable ranges identified in the previous step, Genetic Algorithm (GA) was integrated with VISSIM to calibrate the parameter set for the Arlington network. In this case study, the calibration procedure lasted to generation 24 and 5 replications were made for 10 populations. Figure 105 shows the convergence of the fitness value as the generation number increases.

$$\frac{\left| \text{Travel Time}_{Field} - \text{Travel Time}_{Simulation} \right|}{\text{Travel Time}_{Field}}$$
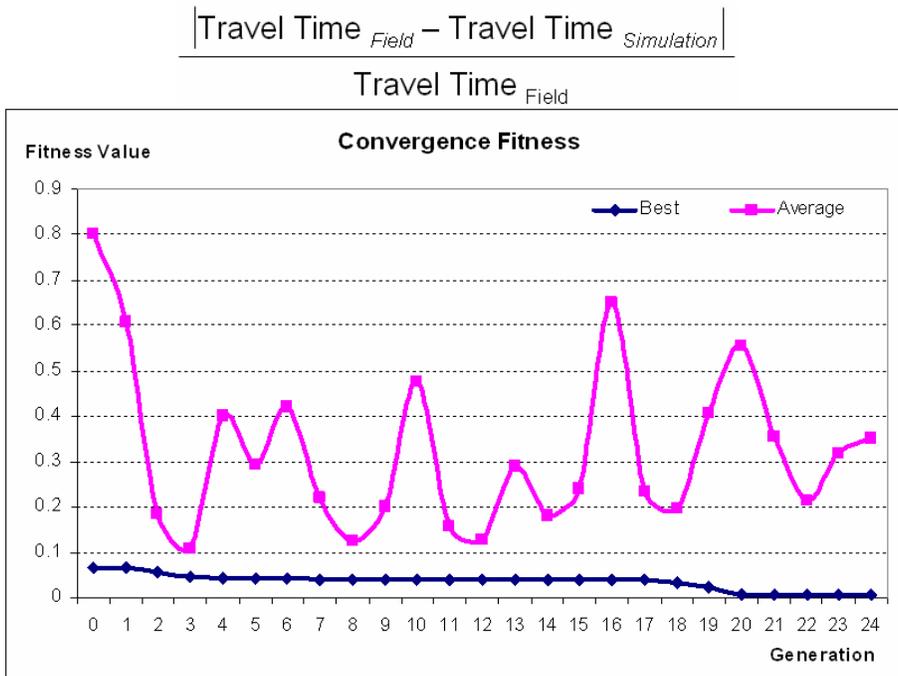
Figure 105. Convergence of Fitness Value with Generation

10.4.6. Evaluation of the Parameter Sets

For evaluation purposes, 100 randomly seeded runs were made for the calibrated model and the travel time data were recorded for two sections. And the distributions of travel time outputs from simulation runs were compared with field-collected travel time data of two sections and presented in Figure 106. As shown in Figure 106, both distributions of travel time outputs could include the field-collected travel time data and considered to be well calibrated. Table 50 shows the default and GA-optimized parameter sets.

Table 50. Parameter Sets for Arlington Network with VISSIM

VISSIM	Default	Calibrated
Desired Speed Distribution (Main/Min), mph	40	45
Desired Speed Distribution (Main/Var), mph	40	35
Desired Speed Distribution (Arterial/Min), mph	30	20
Desired Speed Distribution (Arterial/Var), mph	20	25
Desired Speed Distribution (Ramp Entry/Min), mph	30	20
Desired Speed Distribution (Ramp Entry/Var), mph	20	15
Max. Look Ahead Distance, meter	250.00	253.03
Number of Observed Preceding Vehicles, veh	2	2
CC0, meter	1.50	3.85
CC1, sec	0.90	0.98
CC2, meter	4.00	6.82
CC3, set	-8.00	-9.39
CC4, meter/sec	-0.35	-0.1
CC5, meter/sec	0.35	0.1
CC6, 10^{-4} rad/s	11.44	20.71
CC7, m/s^2	0.25	0.96
CC8, m/s^2	3.50	3.33
CC9, m/s^2	1.50	1.94
Max. Deceleration, m/s^2	-4.00	-3.63
Reduction Rate	200.00	55.15
Accepted Deceleration, m/s^2	-1.00	-1.98
Waiting Time Before Diffusion, sec	60.00	46.06
Min. Headway, meter	0.50	0.27
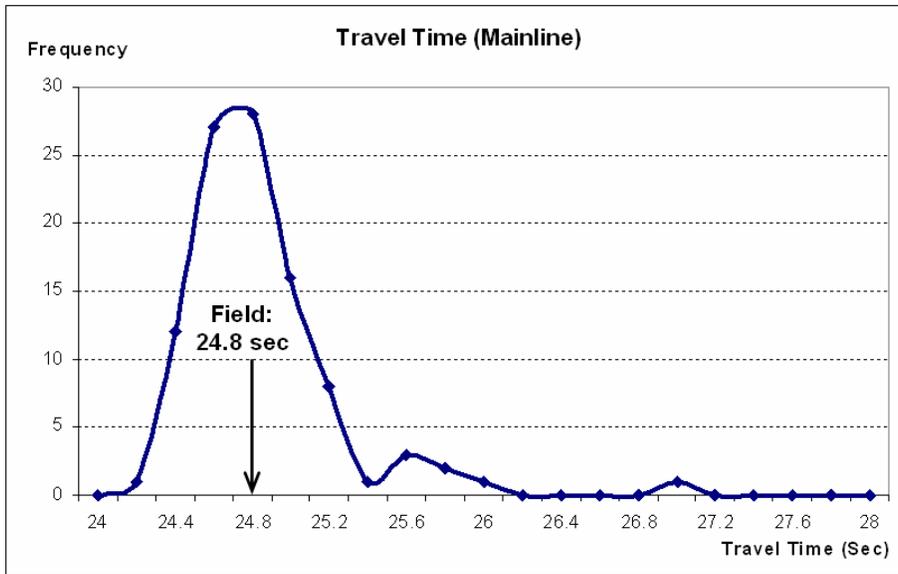
(a) Travel Time—Mainline
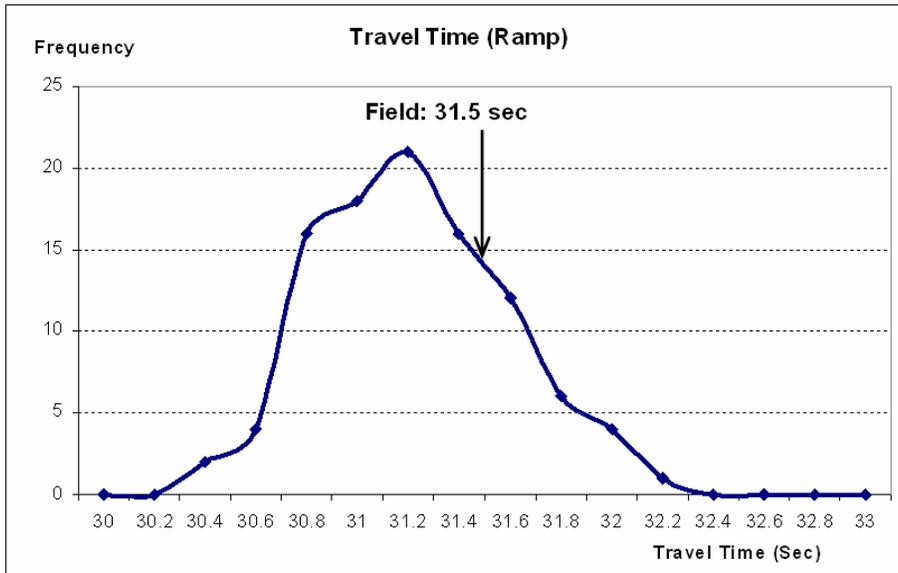
(b) Travel Time—Ramp

Figure 106. Validation Result—Arlington Network with VISSIM

10.4.7. Validation

Traffic data collected on a different day including fundamental traffic data and calibration data were used for the validation of parameter sets obtained from the calibration procedure. As shown in Figure 107, distributions of both travel time outputs were able to include field-collected validation data. Therefore, the result of validation is acceptable and the calibrated model can be used for further analysis.

151
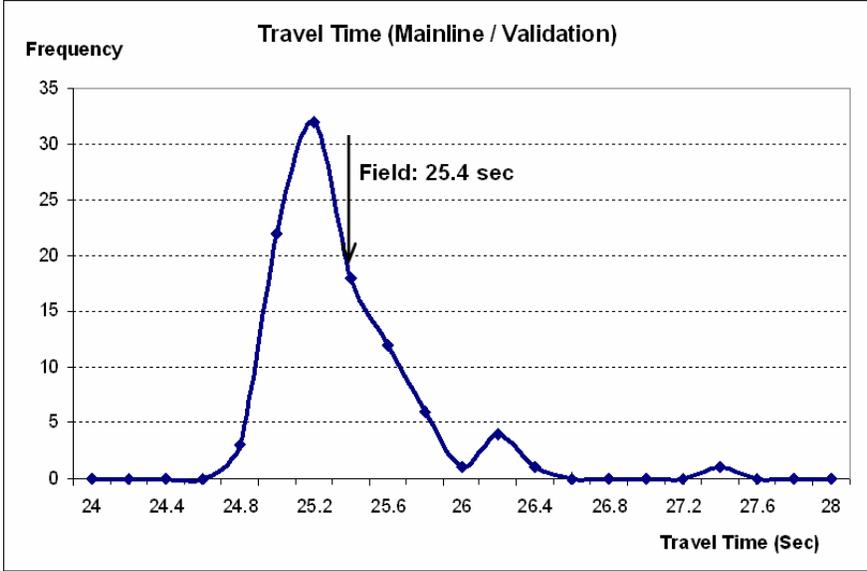
(a) Travel Time—Mainline
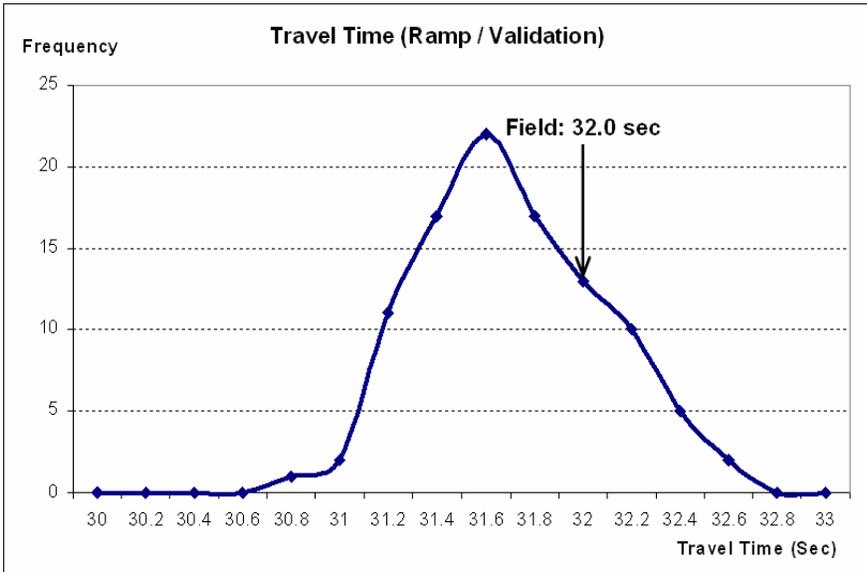
(b) Travel Time—Ramp

Figure 107. Validation of VISSIM Using Arlington Network

References

AIMSUN2, www.tss-bcn.com/products.html, Transportation: Simulation Systems, 2006

Barrett, C., K. Berkbigler, L. Smith, V. Loose, R. Beckman, J. Davis, D. Roberts and M. Williams. *An operational description of TRANSIMS*, LA-UR-95-2393, Los Alamos National Laboratory, Los Alamos, New Mexico, 1995

FHWA, *CORSIM User's Manual*. U.S. Department of Transportation, Office of Safety and Traffic Operations, McLean, VA, 2003

KLD Associates, WATSIM, http://www.kldassociates.com/, 2006

Mitchell, M. *An introduction to Genetic algorithm*, MIT Press, 1996

Park, B., and H. Qi. Development and Evaluation of a Procedure for the Calibration of Simulation Models. *Transportation Research Record: Journal of the Transportation Research Board, No. 1934*, TRB, National Research Council, Washington, D.C., 2005, pp. 208-217.

Park, B., and H. Qi, *Development and Evaluation of a Calibration and Validation Procedure for Microscopic Simulation Models.* VTRC 05-CR1. Virginia Transportation Research Council, Charlottesville, 2004

Quadstone Limited. *PARAMICS Modeler Version 3.0 User Guide and Reference Manual.* Edinburgh, U.K., 2002.

Trafficware, *SimTraffic 6, Traffic Simulation Software User Guide*, Albany, CA, 2003

Van Aerde, M., H. Rakha, and H. Paramahamsan, Transportation Research Record: Journal of the Transportation Research Board, No. 1831, TRB, National Research Council, Washington, D.C., 2003, pp. 122~130.

Van Aerde, M., Hellinga, B., Baker, M., and Rakha, H. *INTEGRATION: An overview of traffic simulation features*, Technical report, Department of Civil Engineering, Queen's University, Kingston, Canada, 1996

VISSIM Version 4.1 Manual, PTV Planug Transport Verkehr AG. Innovative Transportation Concepts LLC, Karlsruhe, Germany, December 2005.

Yang, Q., *A Simulation Laboratory for Evaluation of Dynamic Traffic Management Systems*. Ph.D. Thesis, Center for Transportation Studies, Massachusetts Institute of Technology, USA 1997

Appendix: Prototype Program User's Guide

System Requirements
In order to run this program your computer must have the following specifications and programs installed.
- Windows XP Professional
- CORSIM 5.1
- VISSIM 4.1
- Microsoft Excel

1. Selecting a Simulation Model

Choose to run the CORSIM simulation model by clicking on CORSIM.
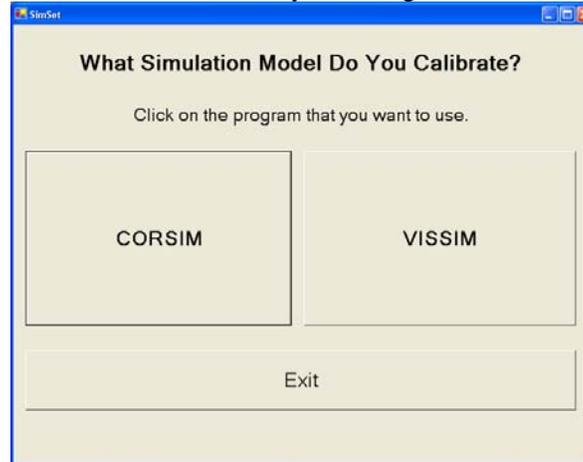
2. Running the Default Procedure

The DEFAULT option enables an evaluation of the "Default" Model, whereas the CALIBRATION option conducts the Calibration Procedure. Choose to start with the RUN DEFAULT procedure.
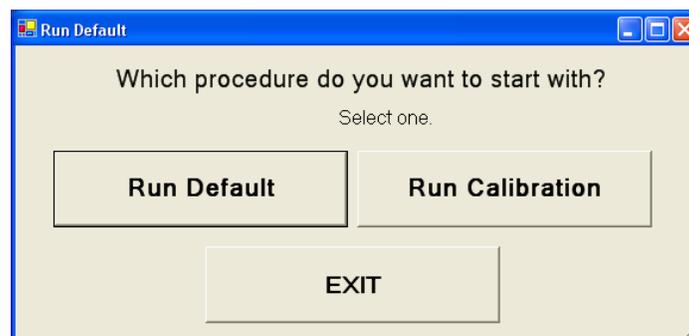
3. Number of Simulations, Location of Files

For the Parameter Range Selection, first choose the number of multiple runs you want to make with the default model. Enter this number in the first blank space. Next, define the location of the simulation input file. For example, CORSIM: *.trf and VISSIM: *.inp. Third, define the location of the executable simulation program in the last blank space.
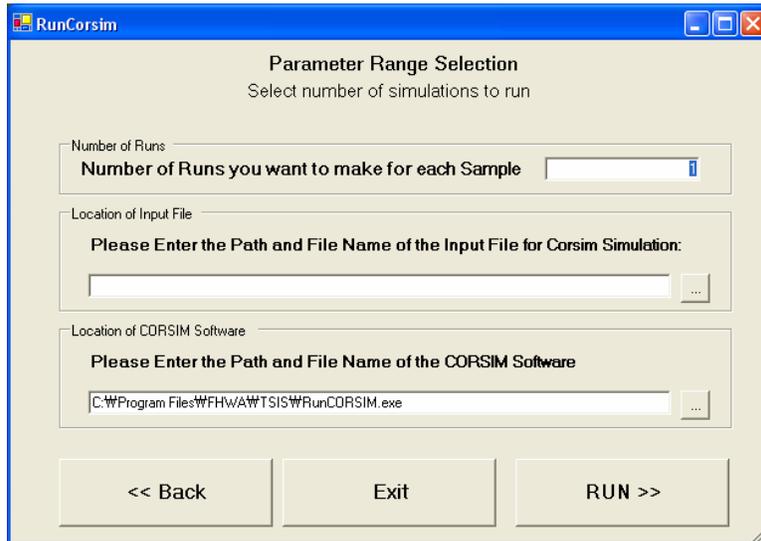

4. CORSIM Calibration

Continue with the Parameter Range Selection by using the pull down menu to choose the link of interest. The travel time output value for each of the runs can be seen in the center white space. The VIEW HISTOGRAM button will display the output distribution of the travel time. After viewing the histogram, click on START CALIBRATION to calibrate the model.
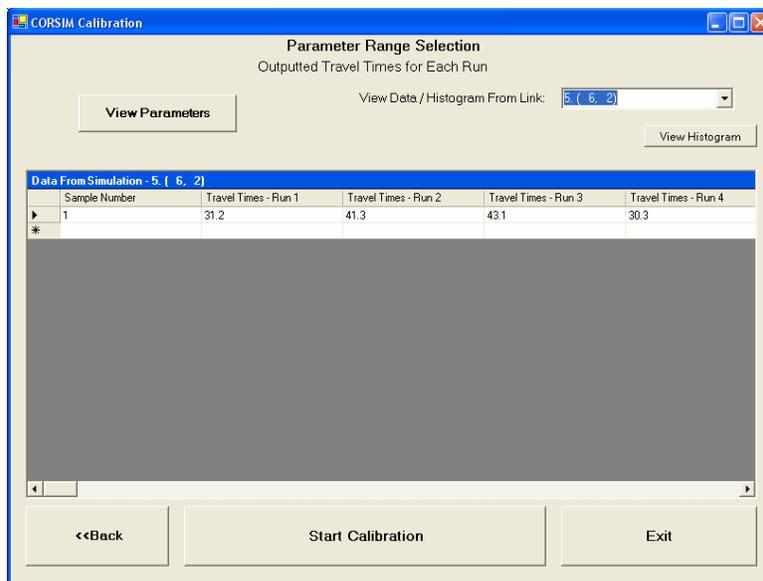
5. Drawing Histogram
The histogram for 100 Travel Time outputs will be displayed. Check "yes", type the field-measured travel time value, and click the button on the right side. You can also use the BACK and EXIT buttons on top for navigating back through the program or exiting completely.
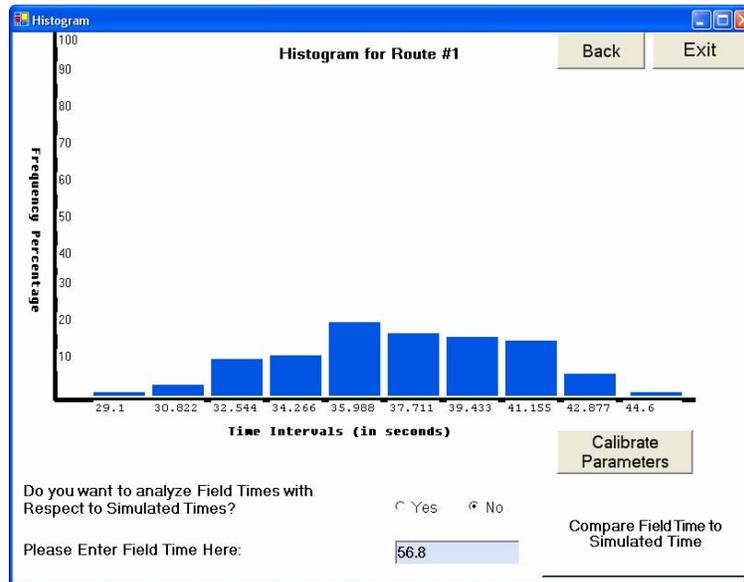
6. Parameters to Test
Select all the parameters that need to be calibrated by marking the box next to the appropriate parameter. When finished, click the OK button to continue to the next step.

7. Parameter Range Selection

The "activated" or white boxes indicate that the parameter was selected on the previous screen; the others remain gray and "inactive". Enter in the determined minimum and maximum value for each of the selected parameters. For Record Type 145 and 149, unless you select any specific type, it uses the same value for the two parameters. Use the value for driver type 1 for the parameters that take a distribution format. Next, indicate the number of samples you wish to test. Lastly, click OK when you are satisfied with everything.

8. Number of Simulations, Location of Files

For the Parameter Range Selection, first choose the number of multiple runs you want to make with the default model. Enter this number in the first blank space. Next, define the location of the simulation input file. For example, CORSIM: *.trf and VISSIM: *.inp. Third, define the location of the executable simulation program in the last blank space.

157

9. Travel Times

Use the pull down menu to select the desired link. The data from the selected link will be displayed in a table. Click on VIEW HISTOGRAM to see the distribution of the travel time output.
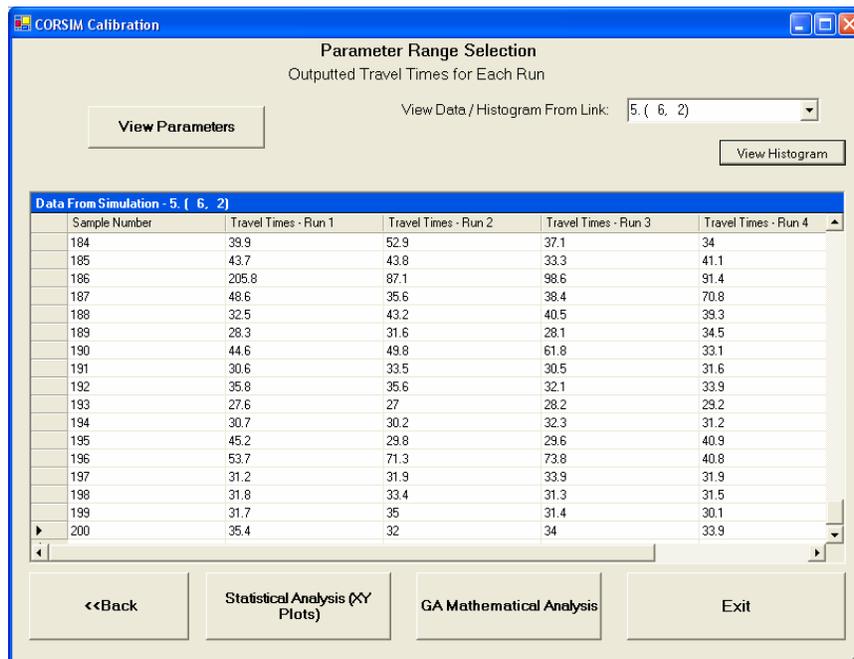
10. Drawing Histogram II

The BACK and EXIT buttons help you to go back to previous options or exit the program. A histogram of 200 average travel time outputs is shown. Choose YES to analyze field times with respect to simulated times and then enter in the desired field-measured travel time value in the space provided. When finished, click the COMPARE FIELD TIME TO SIMULATED TIME button to continue.
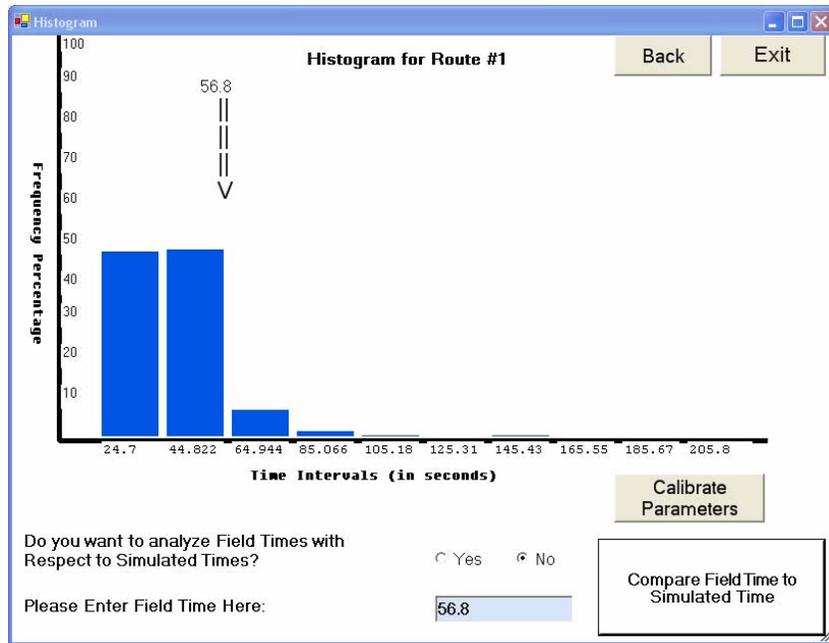
11. Travel Times

Is the set of ranges acceptable? If no, click on the STATISTICAL ANALYSIS button to adjust the ranges. If yes, select the GA MATHEMATICAL ANALYSIS button to start the calibration.

12. CORSIM Calibration

Copy the location information that appears in the space provided and then click on the STATISTICAL ANALYSIS button to start parameter range adjustment.
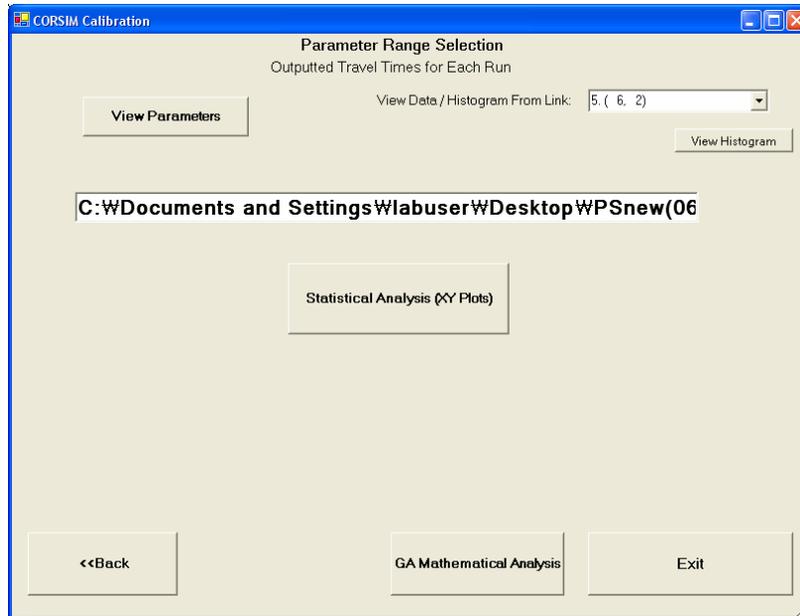
13. Statistical Analysis – Excel

Select the START button. Note: if you cannot load the Excel program, please check your security level again and make sure it is "LOW". To do this, follow the instructions presented here.
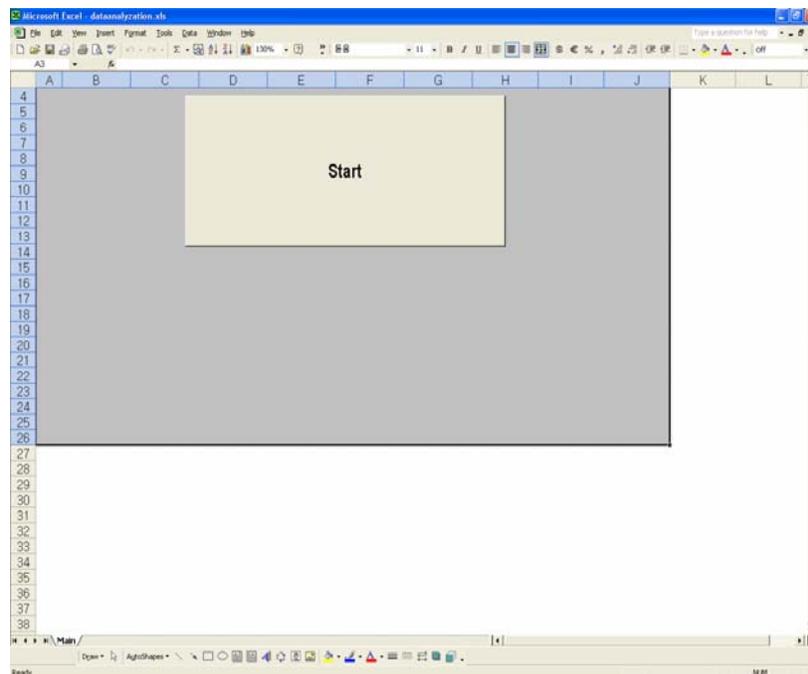
14. Excel Security Adjustment

To set the security level to the necessary LOW level, open the TOOLS menu, select MACRO, and then select SECURITY.
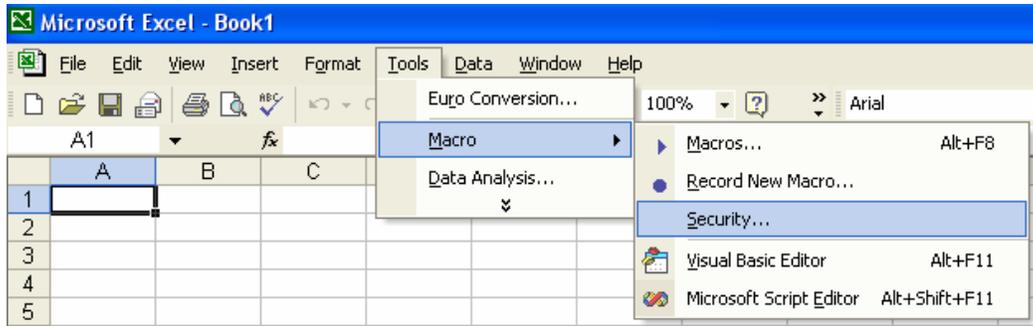
15. Security Levels

You will now have a screen similar to this. Choose the LOW setting under the SECURITY LEVEL tab. Once LOW is chosen, click OK for the changes to take effect.
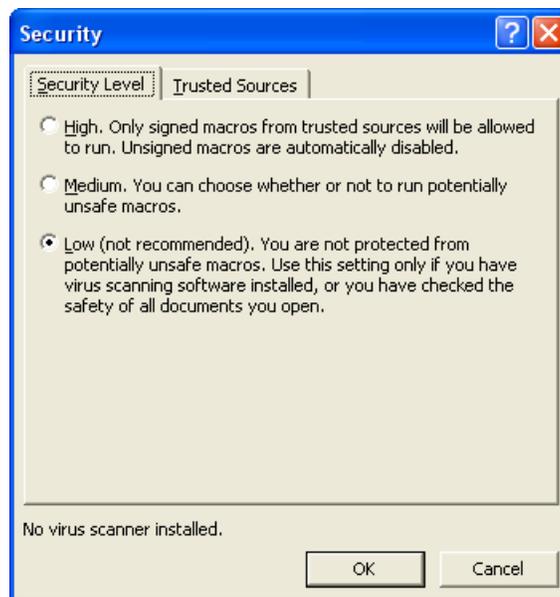
16. Find File Folder

Use the white space to paste the location information that was copied in a previous window. Then choose the simulation model name that you are using, either VISSIM or CORSIM. When done, click the RUN button.
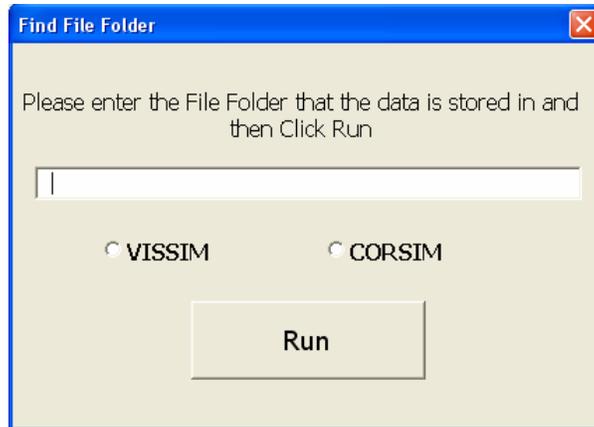
17. Histograms, Times, and Plots

Use the three buttons to do what you want. VIEW HISTOGRAMS will show the histograms of travel time data. ENTER FIELD TIMES will present field travel time value on the histogram. MAKE X-Y PLOTS will create X-Y plots of travel time data and each parameter value.

18. Parameter Selection for CORSIM

Select the calibration parameters that you want to generate an X-Y plot from. For example, the X value could be Travel Time and the Y value may be the Parameter Value. When finished, click the RUN button to generate the X-Y plots.

19. Starting the Calibration Procedure

In order to start the calibration procedure, you need to specify four things. First, the number of generations; second, the field travel time; third, the link identification number; and fourth, the number of populations. When asked for the information, type it in the black window.

Index