2014-01-30

Off-line Multi-Sensor Multi-Source Calibration of Dynamic Traffic Assignment: Simultaneous Demand-Supply Estimation based on Genetic Algorithms in a High-Performance Computer

Omrani, Reza

UNIVERSITY OF CALGARY

Off-line Multi-Sensor Multi-Source Calibration of Dynamic Traffic Assignment: Simultaneous Demand-Supply Estimation based on Genetic Algorithms in a High-Performance Computer

by

Reza Omrani

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF CIVIL ENGINEERING

CALGARY, ALBERTA

JANUARY 2014

Abstract

The complexity of transportation systems often dictates the use of detailed simulation-based dynamic traffic assignment (DTA) models to replicate the complex traffic flow dynamics. Recent advancements in computer technology have led to the development of high-fidelity simulation models; however, in order to be used as reliable tools, the simulation input parameters should be properly calibrated in order to replicate prevailing traffic conditions. Thus, this thesis has focused on the off-line simultaneous calibration of demand and supply parameters of the DTA model in a microscopic context that can capture the interactions between all parameters. The demand parameters include dynamic OD flows, while the route choice and driver behavior model parameters are considered as supply parameters. The calibration process has been formulated as a multi-objective optimization problem that incorporates the traffic data from multiple sources, ranging from traditional loop detector data to traffic data from recent emerging technologies, and allocates relative weights to different terms of the objective function.

A genetic algorithm (GA) is selected as a suitable solution algorithm for the resulting nonlinear stochastic optimization problem. The application of the proposed methodology was implemented in a synthetic case study as well as a complex network in the business district core of downtown Toronto, Ontario, Canada. For this network, the emerging traffic surveillance data from in-vehicle navigation system technology provide an enrich source of disaggregate speed data.

The empirical results from various experiments support the hypothesis that the incorporation of the in-vehicle navigation system speed data can significantly improve the calibration accuracy and minimize the dependency of the calibration process on the historical OD flows. The quality of the solution and convergence speed of a GA is further enhanced by dividing the GA

population into multiple demes and running the GA on a high-performance computer (HPC) cluster with multiple processors (i.e. parallel distributed GA, PDGA).In addition, this research takes a further step towards analyzing the temporal variations of the driving behavior of travelers, especially during different time intervals of peak periods.

Keywords: Off-line simultaneous calibration, dynamic traffic assignment (DTA), high-performance computer (HPC), in-vehicle navigation system technology, driving-behavior parameters.

Acknowledgements

It is still fresh in my memory, the first encounter with her in the office, Dr. Lina Kattan, a very promising, confident looking young professor. At that moment, I had not even imagined that I would become what I am today. I should salute for her extra-ordinary skill to see the hidden talent within a student, help him refine it and make him capable of standing on his own to pave his own path. Not satisfied with the words I use, still a sincere attempt to express my appreciation for her trust, guidance, encouragement, and confidence that she showed upon me to execute this research project. I also highly appreciate her constant encouragement and support for my professional growth and personal life. Lina: let me just say that I am really proud of our joint accomplishments since 2009.

I would like to express my sincere appreciation to my committee members, Prof. John Douglas Hunt, Dr. Alex de Barros, Prof. Naser El-Sheimy, and Prof. Ata M. Khan for reviewing the thesis and their involvement in refining the quality of this dissertation.

My sincere thank goes to Prof. Baher Abdulhai for providing the opportunity and resources of the University of Toronto ITS Lab, our informal discussions about my thesis and for his valuable guidance. I express my deep sense of gratitude to all the ITS lab members, especially Kasra Rezaee and Glareh Amirjamshidi, for being so helpful and friendly and making the stay in the lab a thoroughly enjoyable experience. I also must give special thanks to Dr. Hossam Abdelgawad and Dr. Samah El-Tantawy, for their generous help in answering my questions through the course of this research. I would also thankful to Tamer Abdulazim and Asmus Georgi, for preparing the GA library and setting up the cluster computing facilities that I have used in my research.

I would also like to acknowledge my friends and colleagues in CIMA Consulting Inc., Dr. Pedram Izadpanah, Dr. Alieza Hadayeghi, Brian Malone, Alexandre Nolet, Khaled Hawash, Behzad Rouhieh, Margot Smeenk, Mathew Bilodeau, Sheetal Thukral, and Geoff Knapp. I am indebted to all of them for all the advice, moral support and encouragement. Also my special thanks go to my dear friends from Sharif University of Technology, Hosein Dashtestani, Saeid Sherafatipour, and Elaheh Khademi, for their ever-lasting support and friendship.

I have been very fortunate to receive financial support from the University of Calgary teaching assistantships. In addition, I have also received financial assistance from the Natural Sciences and Engineering Research Council of Canada (NSERC), and Alberta Motor Association (AMA), during the course of my postgraduate studies. I remain truly grateful to the above for their assistance as it helped me a great deal in focusing on my research.

Since the beginning and until this last moment I realized, I was dedicating every word I typed to my family. I'd like to express my deepest gratitude to my dear mother, Shahnaz, for her constant love and unwavering support. Many thanks for my brother, Nima, for his support, encouragement, and being truly a brother when needed.

Finally, to my late father, Rouhollah Omrani, my hard work and efforts put forth in this thesis are merely an extension of all that you have taught me over the years.

In memory of my father

To my mother

With love and eternal appreciation

Table of Contents

Abstract ii

Acknowledgements iv

Dedication vi

Chapter 1: INTRODUCTION 1

1.1 Background 1

1.2 Dynamic Traffic Simulation Models 2

1.3 Need for DTA Model Calibration 3

1.4 Motivation and Problem Statement 4

1.5 Implementation Framework and Contributions 5

1.6 Thesis Outline 8

Chapter 2: LITERATURE REVIEW 9

2.1 The DTA-Based Demand Calibration Studies 9

2.2 The DTA-Based Supply Calibration Studies 11

2.3 Joint Demand and Supply Calibration Efforts 12

2.3.1 Iterative Demand and Supply Calibration 12

2.3.2 Simultaneous Calibration of the Demand and Supply Parameters 22

2.4 Summary 30

Chapter 3: CALIBRATION METHODOLOGY 32

3.1 Overview of the Methodology 32

3.1.1 Input Parameters 34

3.1.2 Output Parameters 34

3.2 Optimization Formulation 35

3.2.1 General Formulation 35

3.2.2 Extension of the General Formulation for the Large-Scale Network 37
3.3 Simulation Environment 39
3.4 Overview of the Solution Algorithm 40
3.5 Calibration Framework 43
3.6 Summary 46
Chapter 4: THE GENETIC BASED SOLUTION APPROACH 47
4.1 Overview of Genetic Algorithm 47
4.1.1 Chromosome Representation 50
4.1.2 Selection 51
4.1.3 Crossover 52
4.1.4 Mutation 53
4.2 Advanced GA Methods for the Large-scale Optimization Problems 53
4.2.1 Parallelization Structure of GA 55
4.2.2 GENOTRANS and GridGain: The PDGA Platform 57
4.3 Integration of High-Performance Computing Cluster with GA 61
4.4 Summary 64
Chapter 5: Case Study I: Synthetic Network 65
5.1 Experimental Design 66
5.1.1 Description of the Test Network 66
5.1.2 Input Parameters for Calibration 67
5.1.3 Synthesized Observed Traffic Data 67
5.2 GA Implementation Details 69
5.2.1 Simple GA Control Parameters 69
5.2.2 Parallel GA Control Parameters 70
5.3 Calibration Results 71

5.3.1 Calibration Results based on Different Traffic Data 71
5.3.2 Comparing Calibration Approaches 79
5.3.3 Comparing the Optimization Engines: Simple GA (SGA) vs. Distributed GA (DGA) 81
5.3.4 Comparing the Simple GA (SGA) and Distributed GA (DGA) to the Parallel Distributed GA (PDGA) 84
5.4 Sensitivity analysis....91
5.4.1 Case I 91
5.4.2 Case II 94
5.5 Summary 95
Chapter 6: Case Study II: Water Front Network 96
6.1 Experimental Design....97
6.2 Input Parameters for Calibration....99
6.3 Historical OD Flows 100
6.4 Observed Traffic Sensor Data....102
6.5 Observed Turning Movement Counts....102
6.6 Speed Data 103
6.6.1 Databases 104
6.6.2 Evaluating Available Data Sources: Selecting the Candidate Technology 106
6.6.3 Speed Data from In-Vehicle Navigation System Technology....108
6.7 Implementation Details: A Note on the Degree of Freedom 115
6.8 GA Control Parameters....116
6.9 Calibration Results....117
6.9.1 Experiment I: Calibration based on Multi Source Traffic Data Using Distributed Computing....117
6.9.2 Experiment II: Calibration based on Parallel Distributed GA 131

6.9.3 Experiment III: Testing the Impact of Link Segmentation 145

6.10 Summary 152

Chapter 7: Conclusions and Recommendations 154

7.1 Context and Scope of Research 154

7.2 Summary of Findings 156

7.3 Research Contributions 157

7.4 Research Challenges 160

7.5 Future Research Directions 161

Bibliography 163

Appendix A: Statistical Validity of Sample Size 175

Appendix B: Comparison between Observed Counts/Speed Data and Their Simulated Counterparts for Different Scenarios (Experiment I) 178

Appendix C: Comparison between Observed Counts/Speed Data and Their Simulated Counterparts for Multi-Source Scenario (Experiment II) 184

Appendix D: Simulated OD Flows for Different Time Intervals Using PDGA (Experiment II) 188

List of Tables

Table 2.1 Summary of iterative calibration research literature results 21

Table 2.2 Summary of simultaneous calibration research literature results 29

Table 3.1 Summary of the classification of optimization algorithms [28, 63, 70, 73, 74, 78, 79, 81] 42

Table 4.1 Summary of GA control parameters [81, 87] 49

Table 5.1 Configuration of control parameters 70

Table 5.2 Parallel GA design elements 71

Table 5.3 Calibration result statistics based on counts 72

Table 5.4 Calibration result statistics based on speed values 72

Table 5.5 Calibrated driver behavior parameters 83

Table 5.6 Comparison between DGA and PDGA calibration result statistics based on counts and speed values 84

Table 5.7 Summary of the experiments 90

Table 5.8 Sensitivity analysis based on NRMSE (Case I) 92

Table 5.9 Sensitivity analysis based on NRMSE (Case II) 94

Table 6.1 Gateways to the Study Area 98

Table 6.2 Sample raw data 109

Table 6.3 Example of performance measures for Gardiner Expressway during AM peak periods of the selected week in the Fall season 113

Table 6.4 Simple and parallel GA control parameters 116

Table 6.5 Calibration result statistics based on NRMSE 118

Table 6.6 Optimal weighting factors, **αi** 125

Table 6.7 Optimal fitness function values (DGA) 126

Table 6.8 Calibrated driver behavior parameters (multi-source scenario) 127

Table 6.9 F-test and t-test results for comparing vehicle driving behavior parameters (multi-source case) 130

Table 6.10 Comparison between DGA and PDGA calibration result statistics based on counts and speed NRMSE values 131

Table 6.11 Optimal weighting factors, **αi** 137

Table 6.12 Optimal fitness function values for multi-source scenario (DGA vs. PDGA) 138

Table 6.13 Calibrated driver behavior parameters 139

Table 6.14 Summary of dynamic OD travel demand during AM peak periods (Fall 2012) 140

Table 6.15 Summary of dynamic OD travel demand during PM peak periods 141

Table 6.16 2006 seed OD matrix for the Water Front network 142

Table 6.17 Optimal fitness function values for two scenarios 149

Table 6.18 Comparing driver behavior parameters of the two scenarios with previous experiment 149

List of Figures

Figure 2.1 Flowchart of iterative demand and supply calibration 13

Figure 2.2 Flowchart of simultaneous demand and supply calibration 24

Figure 3.1 Overview of the DTA model 33

Figure 3.2 Flowchart of OD estimation and calibration of driver behavior parameters 43

Figure 4.1 Examples of islands topologies 56

Figure 4.2 GENOTRANS configuration generator (Simple GA) 58

Figure 4.3 GENOTRANS configuration generator (Parallel GA) 59

Figure 4.4 Interaction flow between GENOTRANS and GridGain (Inspired from [97]) 60

Figure 4.5 Flowchart of the PDGA implemented on the GENOTRANS and GridGain platforms 61

Figure 4.6 High-performance computing facility at University of Toronto [91] 62

Figure 4.7 Master/slave cluster setup in GridGain 63

Figure 5.1 Synthetic network topology 66

Figure 5.2 Interaction between Paramics and Aimsun 68

Figure 5.3 Comparison of different scenarios based on NRMSE values 73

Figure 5.4 Comparison of different scenarios based on GEH values 73

Figure 5.5 Comparison between measures of effectiveness of different cases based on counts (15-minutes intervals) 74

Figure 5.6 Comparison between measures of effectiveness of different cases based on speed values (15-minutes intervals) 75

Figure 5.7 Comparison between simulated and observed counts 78

Figure 5.8 Visual comparison between simultaneous and sequential approaches based on NRMSE 80

Figure 5.9 Visual comparison between simultaneous and sequential approaches based on GEH 80

Figure 5.10 Comparison between measures of effectiveness of DGA and PDGA based on counts 85

Figure 5.11 Comparison between measures of effectiveness of DGA and PDGA based on speed values 86

Figure 5.12 Comparison between simulated and observed counts based on NRMSE (DGA vs. PDGA) 88

Figure 5.13 Effect of parallelization and distributed computing on GA convergence and quality of solutions (PDGA vs. DGA/SGA) 88

Figure 6.1 Paramics network topology 99

Figure 6.2 The Greater Toronto and Hamilton Area [101] 101

Figure 6.3 Loop detector counts along the sections of the study area 102

Figure 6.4 Cumulative travel time and speed for DVP northbound during AM peak period 107

Figure 6.5 Network data coverage 110

Figure 6.6 Route, segment, and link definitions 111

Figure 6.7 Continuous segments at an interchange 111

Figure 6.8 Comparison of different scenarios based on NRMSE values 119

Figure 6.9 Comparison between measures of effectiveness of different cases based on NRMSE for count and speed data 120

Figure 6.10 Comparison between simulated and observed counts (all three scenarios) 123

Figure 6.11 Comparison between simulated and observed speed values (all three scenarios)... 124

Figure 6.12 Temporal variations of the vehicle-driving behavior during peak periods 129

Figure 6.13 Temporal variations of the route choice model parameters during peak periods.... 129

Figure 6.14 Comparison between measures of effectiveness of DGA and PDGA based on NRMSE for count and speed data..... 132

Figure 6.15 Comparison between simulated and observed counts using multi-source data (DGA vs. PDGA)..... 134

Figure 6.16 Comparison between simulated and observed speed using multi-source data (DGA vs. PDGA)..... 135

Figure 6.17 Effect of parallelization and distributed computing on GA convergence and quality of solutions (PDGA vs. DGA/SGA)..... 136

Figure 6.18 Temporal variations of the 15-minutes dynamic OD flows 144

Figure 6.19 Subset of the Water Front network in ArcGIS 147

Figure 6.20 Comparison between NRMSE of speed data for the two scenarios..... 150

Figure 6.21 Comparing the evolution of the fitness function of the two scenarios 152

Chapter 1: INTRODUCTION

1.1 Background

Traffic congestion is one of the major sources of energy consumption and pollution in urban areas, which are characterized by a high concentration of activities that generate high demand for travel, as derived from the complex spatial and temporal interaction of land uses and activity nodes. Due to fiscal, land and environmental constraints, the building of more roads is often not a viable solution.

Intelligent transportation systems (ITS) offer instrumental strategies in achieving sustainable transportation solutions without additional right-of-way costs. ITS rely on control, management and information dissemination to reduce network congestion. Traffic control and management systems optimize the efficiency of traffic networks by responding to the dynamic and random nature of traffic almost instantaneously. ITS solutions have been shown to be capable of reducing traffic congestion and, thus, energy consumption and pollution emission.

ITS technologies make use of recent developments in field sensing and computation to better manage the available roadway infrastructure and obtain traffic information over a wide spatial area at a relatively low cost. Traffic sensing devices, such as inductive loop detectors, Bluetooth technology, video cameras, in-vehicle navigation systems and mobile phone probes with GPS, have made the task of traffic control and management much more tractable.

In order to utilize these advanced traffic control and management systems, traffic control centers (TMCs) collect traffic data, which are fed into a traffic simulation model to be calibrated to provide estimation and prediction of traffic conditions. Traffic simulation models rely on dynamic traffic assignment (DTA), as a decision support system, to develop adaptive control and incident management schemes and to provide guidance, routing policies and traffic information that are capable of achieving system-wide objectives.

DTA is the process of estimating the spatial and temporal progression of traffic flow in a transportation network to reflect the traffic behavior. DTA takes into consideration the

dynamic nature of travel demand and the likely changes in the transportation network to estimate the performance of the network. This latter information is used as a decision tool to develop appropriate advanced traffic management systems (ATMS) and advanced traveler information systems (ATIS) to alleviate congestion. Users may respond to guidance through trip rescheduling, mode change and/or rerouting. ATMS strategies, such as adaptive ramp metering strategies, also require DTA to determine the optimal on-ramp diversion rates for travelers between a given origin and destination. DTA is also needed for the control of traffic signals and optimization of adaptive control systems. In general, DTA systems are viewed as the solution to the problem of accurate traffic estimation and prediction.

1.2 Dynamic Traffic Simulation Models

The complexity of transportation systems often dictates the use of detailed simulation-based DTA models. DTA simulation modeling, as extensively reviewed in the literature, is an increasingly popular and effective tool for analyzing transportation problems that are not amenable to study by mathematical programming, optimal control, and variational inequality formulations [1]. Simulation-based DTA models use a traffic simulator to replicate the complex traffic flow dynamics, such as traffic flow propagation and vehicular movements.

Generally, traffic simulation models can be classified into three categories based on the level of details regarding driver behavior and traffic streams: macroscopic, mesoscopic and microscopic. Macroscopic models characterize the traffic stream as a whole and consider traffic interactions as fluid dynamics. Driver behavior elements, such as route choice and departure times, are not considered in these approaches. Popular macroscopic commercial software packages are VISUM [2], EMME/2 [3], and METANET [4].

At the other end of spectrum are microscopic models, which characterize the behavior of individual vehicles within the traffic stream, or specific pairs of vehicles within the traffic stream. These models consider driver behavior, such as lane changing, car following and merging maneuvers. Among the microscopic software packages, CORSIM [5], PARAMICS [6], AIMSUN [7], MITSIMLab [8, 9], VISSIM [10, 11] and TransModeler [12] are the most recently developed, with NETSIM, TRAF-NETSIM, INTRAS and FRESIM as early developed micro-simulation packages [13].

Mesoscopic models combine driver behavior characteristics from microscopic models with traffic relationships as the interactions between speed, density and flow from macroscopic modes. The most recently developed software packages for mesoscopic modeling are DynaMIT [14, 15], DYNASMART [16, 17], Dynameq [18, 19], and DynusT [20, 21].

1.3 Need for DTA Model Calibration

Recent advancements in computer technology have led to the development of high-fidelity simulation models; however, in order to be used as reliable tools, the simulation models should be properly calibrated to replicate prevailing traffic conditions.

The aim of DTA calibration is the minimization of the discrepancy between the observed and simulated traffic conditions, in order to closely replicate drivers' behaviors. The output of a calibration process is the updated estimates of simulator parameters that are used as inputs for a traffic estimation framework. Although DTA models provide an abstraction of actual demand and supply parameters, their output is highly dependent on the accuracy of the estimated input parameters, which are divided into two types: demand parameters and supply parameters.

Travel behavior modeling and origin-destination (OD) demand estimation are considered in the determination of the demand parameters. Supply parameters simulate traffic dynamics, queue formation, dissipation and spillback and are calibrated in either a microscopic or mesoscopic context. The interaction between the supply and the demand is modeled through the traffic assignment process that simulates the propagation of traffic (demand) on the physical network (supply) taking into consideration driver behavior parameters.

To constantly maintain the internal representation of the traffic network consistent with that of the actual network, DTA calibration has to be a periodic adjustment process that is based on frequently updated traffic data. DTA model calibration is thus performed using surveillance data of the prevailing conditions, such as sensor data, global positioning systems (GPS) tracking devices, and video surveillance.

1.4 Motivation and Problem Statement

There is extensive literature related to calibration of DTA models, on both the estimation of OD flows as well as the calibration of supply parameters. However, most of the studies in this area treated each model's parameters separately and focused on utilization of particular type of sensor data, most commonly link-flow counts. These studies were followed by research efforts to jointly calibrate the DTA model parameters in an iterative sequential fashion and jointly approach. The former calibration framework failed to capture the interactions between the demand and supply parameters and was found to be computationally inefficient. In contrast, the latter approach was highly dependent on historical OD flows, which may lead to misleading solutions. In other words, this technique is mostly applicable when the starting point is close to the optimal one. The readers are referred to Chapter 2 for a detailed overview of the literature related to the previous studies on the calibration of demand and supply parameters of the DTA models.

In the DTA calibration process, several types of information extracted from various sources need to be incorporated to have a reliable estimation framework. In recent years, emerging wireless communication technologies and the widespread use of mobile devices and in-vehicle navigation systems provide the opportunity to automatically obtain traffic information over a wide spatial area at significantly lower cost than using dedicated sensors. These emerging technologies provide a great opportunity to overcome the dependency of the OD estimation on historical OD flows as the starting points. A large set of these technologies can be categorized as in-vehicle navigation systems and automatic vehicle identification (AVI), the data from which form the base database of this study and enrich the accuracy of the calibration process.

The calibration of the DTA model parameters can be formulated as a multi-objective optimization problem, which has the flexibility to accommodate any type of traffic measurements into the calibration framework, ranging from traditional loop detector data to data from in-vehicle navigation systems. The objective function can be represented by a goodness-of-fit function that measures the closeness between the observed (or historical) and simulated traffic data. The objective function is subject to a number of constraints, including the allowable lower and upper bounds of the demand and supply parameters. Moreover, the

fitted (or simulated) traffic measurements are a function of the calibrated parameters and network geometry.

This thesis focuses on the calibration of DTA models using the traditional loop detector count and turning movement data, the speed data from AVI sensors, and data from in-vehicle navigation systems. The general calibration model was formulated as a simultaneous estimation of all the demand and supply variables, including OD flows, route choice model parameters and driver behavior parameters. A genetic algorithm (GA) was selected as the solution tool to jointly estimate OD flows and identify the supply model parameters. The Paramics microscopic DTA system was used to demonstrate the feasibility of the proposed calibration methodology. The application of the proposed methodology was implemented in a synthetic case study as well as a complex real-world network in the business district core of downtown Toronto, Ontario, Canada. The implantation of the calibration framework and the results obtained from the case studies are briefly described in the next section.

1.5 Implementation Framework and Contributions

This thesis is aimed at the development of a GA-based multi-criteria optimization framework for the simultaneous calibration of demand and supply parameters in DTA. During the last decade, there has been growing interest in the application of GAs for data fusion in the large-scale optimization problems. This is mainly due to the fact that GAs are inherently parallel in nature and are able to deal with difficult optimization problems having complex nonlinear and/or non-differentiable objective functions with complicated constraints and non-homogeneous and noisy information.

An important property of GAs is that they can be spatially distributed among multi-deme populations, as opposed to a single deme chromosome pool to improve the efficiency of the algorithm. In addition, GA can be run in parallel processors to speed up the convergence. Therefore, three types of GAs were evaluated in terms of comparative performance: a simple GA (SGA), a distributed GA (DGA) among multiple processors in a high-performance computer (HPC), and a parallel distributed GA (PDGA), which refers to a GA's population structure running in multiple processors. The three algorithms were compared for Paramics'

calibration driver behavior parameters and estimation of dynamic OD flows using a synthetic and a large-scale complex network.

For the synthetic network, a number of calibration experiments were performed to demonstrate the feasibility of the proposed simultaneous calibration process, identify the effect of augmenting the GA operator with parallelization and distrusted computing schemes, and evaluate the effect of adding AVI data into the calibration process. Based on the results obtained from the synthetic network, the proposed methodology was implemented in a complex large-scale network using the traffic data from various sources. Special considerations were given to the enriched speed data from in-vehicle navigation systems to improve the calibration accuracy and minimize the discrepancy between the observed traffic data and their simulated counterparts. Several simulation experiments were performed to achieve the objectives of this chapter.

In summary, this research makes concrete contributions to the state of the art, specifically:

- Development of the simultaneous calibration of complex DTA demand and supply parameters in a microscopic model that considers the complex and nonlinear interactions between demand and supply parameters and minimizes the dependency of the calibration approach on the historical OD flows.

- Development of a generic framework capable of incorporating several types of traffic information derived from different sensors/sources and with different levels of accuracy. The incorporation of the traffic data from multiple sources into the global optimization problem results in a more complete representation of the state-of-the-traffic network, takes advantage of all surveillance data, and thereby reducing the possibility of suboptimal solutions.

- Creation of reliable off-line dynamic OD flows and Paramics model parameters that can be used as a priori estimates for on-line calibration process. In other words, the outputs of the off-line calibration process can be used for real-time OD estimation along key corridors, incident management and reduction of unexpected congestions and, ultimately, provision of real-time traffic data to travelers.

- Incorporation of the weighting factors given to different components of the objective function into the calibration process based on the reliability of different sources, in contrast to the traditional sensitivity analysis (i.e. trial and error) methodology.

- Application of distributed computing with a HPC to expedite the GA calibration process for large-scale complex transportation problems during various time intervals.

- Improvement in the calibration accuracy and convergence speed of the GA by parallelizing the population to multiple slaves (parallel GA, PGA) and distribution within an HPC (i.e. PDGA). While the application of PGA and DGA were studied in the literature, the simultaneous incorporation of the two extensions of the GA (i.e. PDGA) and their joint impact on the quality of the solution and convergence speed have not yet been studied in a large-scale network.

- Significant improvement in the calibration accuracy with the inclusion of the enriched speed data from in-vehicle navigation system technology and minimize the dependency of the calibration accuracy on historical OD flows.

- Sensitivity analysis of the driver behavior and route choice model parameters during peak periods. This research takes a further step towards analyzing the temporal variations of the driving behavior of travelers, especially during different time intervals of peak periods.

- Incorporation of the enriched raw speed data without aggregation for smaller links into the calibration process reduce the dependency of the calibration process on historical OD flows in a medium-sized network and improve the calibration accuracy without having a major impact on the computational time for the subject network. Moreover, the incorporation of the raw speed data into the calibration process can further diminish the dependency of the calibration process on the historical OD flows.

1.6 Thesis Outline

The remainder of this document is organized as follows. Chapter 2 presents a detailed review of the DTA model calibration approaches and identifies the strengths and limitations of recent work in this area. In Chapter 3, the proposed GA-based simultaneous demand and supply calibration framework is formulated. In addition, an extension of the calibration methodology applicable for the real-world network is described.

Chapter 4 presents an overview of the fundamental operational mechanisms of GA and elaborates on the advanced GA-based methods for large-scale optimization problems, based on distribution and parallelization schemes. The calibration results from the synthetic case study are presented in Chapter 5, while Chapter 6 applies the framework to calibrate a large-scale traffic network and illustrates the scalability of the methodology in various experimental studies. Finally, conclusions, contributions and directions for further research are outlined in Chapter 7.

Chapter 2: LITERATURE REVIEW

The literature related to the calibration of DTA models can be categorized based on: 1) calibration of the demand models (or estimation of time-dependent demand), and 2) calibration of the supply models. Travel behavior modeling and origin destination (OD) estimation problems are considered in the determination of demand models. Supply models simulate traffic dynamics, queue formation, dissipation and spillback. The calibration of supply models consists of the estimation of capacities and link performance function for mesoscopic and macroscopic models and the estimation of the parameters of car following, lane changing and driver aggressiveness and of awareness parameters for the microscopic context.

Most of the early research literature on DTA model calibration has treated the various parameters to be calibrated independently. While such a calibration approach is viewed as only a part of the overall problem of DTA model calibration, the experience from such analysis has provided valuable directions for the joint estimation of all relevant DTA model parameters. Earlier calibration efforts were based on an iterative approach, followed by more recent work on simultaneous calibration framework.

The first section of the literature review starts with separate short reviews of the demand estimation approaches and the calibration of the supply parameters. This is followed by the past efforts in DTA model calibration, incorporating both demand and supply parameter estimation. The final section provides a summary of literature discussion in this chapter.

2.1 The DTA-Based Demand Calibration Studies

In general, the DTA-based OD estimation problem is expressed through a fixed-point model [22]. In the literature, the specific case of bi-level programming optimization methods is considered. In this traditional OD structure, the upper level problem estimates the OD matrix, assuming known OD path flows; whereas, the lower level solves an assignment problem assuming that the OD demand matrix, which is obtained from the previous step in the upper level module, is fixed. This bi-level OD structure may result in inconsistency between the link flow proportions and the OD estimation in congested conditions.

This problem of inconsistency between the assignment matrix and the OD estimation problem is addressed by: 1) using the numerical gradient-based methods [23, 24], 2) formulating the OD estimation as a variational inequality [25, 26] or 3) using meta-heuristic approaches, such as evolutionary algorithms (EA) [27, 28], simulated annealing (SA) [29] and simultaneous perturbation stochastic approximation (SPSA) [30, 31].

Most of these research efforts were focused on the offline estimation of the demand; however, Zhou and Mahmassani [32] developed the OD demand consistency checking system to update the model online. The authors proposed predictive and reactive approaches to minimize the deviations between real-world measurements and simulated states by adopting a Kalman Filtering (KF) framework for the development of a prediction-correction methodology in real-time DTA.

The determination of the best configuration of sensor locations (i.e., the Network Sensor Location Problem, NSLP), so that most of the unobserved path flows are captured, would improve the accuracy of the OD estimation problem. Hu et al. [33] attempted to independently solve this problem without assumptions of prior knowledge on model parameters and/or OD demands. A “basic link” method was proposed to determine the locations of vehicle sensors, by using the link path incidence matrix to express the network structure and then identifying its “basis” in the context of matrix algebra. The application of the proposed approach was demonstrated in some synthetic and real networks.

In order to address one of the practical aspects of deploying DTA for planning applications, Zhou et al. [34] proposed a two-stage subarea demand estimation procedure to provide time-dependent OD trip information for subarea analysis. In the first stage, path flow patterns in the complete network are generated to calculate the OD demand in the subarea network using DYNASMART-P software. In the second stage, the OD demand information is combined with available real-world traffic observations to update the subarea OD demand matrix.

In a recent work, Verbas et al. [35] addressed another practical aspect of DTA application in large-scale OD calibration. The authors proposed a modified bi-level approach for the OD estimation problem, in order to reduce the number of parameters, as described in [23] and [24], overcome the obstacles for large-scale networks and estimate multiple vehicle classes,

along with an approach to reduce the time and memory requirement of the lower-level problem. The ordinary least squares upper level problem is solved using the reduced gradient / quasi-Newton methods offered in MINOS software and the interior point / conjugate gradient methods in the KNITRO solver package, which are widely used for solving large constrained nonlinear problems.

2.2 The DTA-Based Supply Calibration Studies

The supply calibration component in DTA models has been mostly focused on calibrating only speed-density models, examples of which include the traffic flow models in DYNASMART, DynaMIT and DYNAMEQ software packages. A new mesoscopic modeling concept, the vehicle-based anisotropic mesoscopic simulation (AMS) model, has been proposed [36, 37]. This concept is related to different types of merges/diverges, but departs from the typical link-based queue-server model. With AMS, each vehicle maintains its own prevailing speed. This feature is different from certain previous models using the speed-density model, in which all moving vehicles on the same link travel at the same speed. In the calibration process, vehicle trajectory data from the Next Generation Simulation (NGSIM) research program was incorporated [37]. Two modified Greenshield's models, which are used by the supply simulator model of DynaMIT and DYNASMART, were used. The results satisfactorily validated the AMS model.

Tavana and Mahmassani [38] used the transfer function methods (bivariate time series models) to estimate dynamic speed-density relations from typical detector data. The resulting model is descriptive rather than behavioral in estimating speed and, consequently, predicting its value for future time intervals. Huynh et al. [39] extended this work to incorporate the transfer function model into a DTA simulation-based framework. The estimation of speeds using the transfer function model is implemented as an adaptive process, where the model parameters are updated online, based on the prevailing traffic conditions. A nonlinear least squares optimization procedure is also incorporated into the DTA system to enable the estimation of the transfer function model parameters online. The scope of this study, however, was limited to updating speeds on a single link using synthetic data; therefore, the model was not validated with real data.

Qin and Mahmassani [40] addressed these shortcomings by evaluating the same model with actual sensor data from several links of the Irvine California network. From the numerical results, the performance and robustness of the transfer function model was, in general, found to be superior to the static modified Greenshield's model.

2.3 Joint Demand and Supply Calibration Efforts

The literature for demand-supply calibration for a simulation-based DTA system is rather limited, which can be attributed mainly to the fact that DTA is relatively a new field of research and most studies have focused on developing the theoretical foundations for modeling. In addition, limited traffic data has made it difficult to employ the calibration of the DTA model in a system-level approach. Thus, the primary studies in this context treat calibration of different parameters independently. However, the independent calibration approach is not efficient and optimal, as it ignores the presence of interaction among the various demand and supply parameters.

The joint calibration of DTA models can be categorized into two groups of research efforts. The first group focused on an iterative demand-supply calibration approach, which was followed by a few recent studies that have investigated the joint calibration of demand and supply. In what follows, the iterative calibration of the demand and supply parameters are first reviewed, followed by the simultaneous calibration efforts.

2.3.1 Iterative Demand and Supply Calibration

Figure 2.1 presents a general flowchart for the iterative demand and supply DTA calibration model. As the figure indicates, estimation of the route choice and driver behavior parameters is conducted with available disaggregate data (e.g., vehicle trajectory, surveys), independent of the overall simulation framework. In the second step, aggregate data (e.g., average travel speeds, flow information) are used to fine-tune the parameters and calibrate the general parameters in the simulator. The OD flows and route choice model parameters are calibrated first, followed by calibration of the driver behavior parameters. These two steps are iterated until a convergence criterion is met.

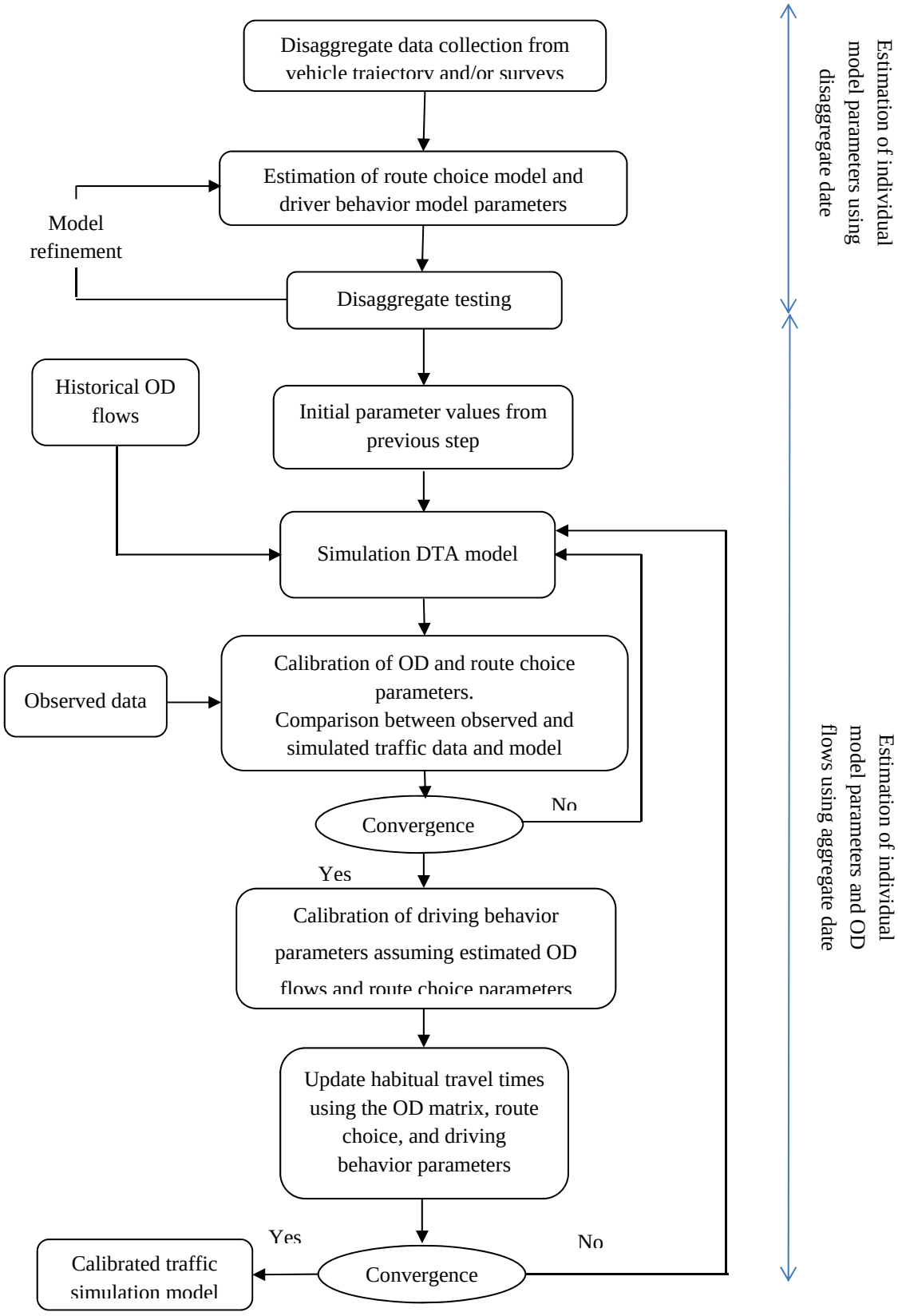


Figure 2.1 Flowchart of iterative demand and supply calibration

Doan et al. [41] conducted one of the earliest studies related to DTA model calibration for a system-level approach. The authors attempted to develop a conceptual framework for a traffic monitoring system that maintains a representation of the traffic state of a network consistent with that of an actual network. To minimize the discrepancy between the actual state and the predicted one, the authors adopted a combination of online and offline adjustment approaches. In real-time DTA, the time frame is divided into several roll periods (i.e., estimation intervals), for which the DTA process is carried out. Every roll period that involves DTA runs is called a rolling horizon real-time DTA (RT-DTA) process. The authors categorized the error sources from the rolling horizon method as associated with the demand estimation, path estimation, traffic propagation, internal traffic model structure, and online data observation. For that purpose, the authors proposed a monitoring system that incorporates both online and offline adjustment modules.

Real-time traffic surveillance, a real-time traffic simulator and a proposed reactive traffic propagation adjustment, formulated as a PID (proportional, integral, derivative) feedback controller, are incorporated into the online module. This module constantly computes the deviations between the actual observed network's traffic measures, such as speed, and the internal simulated measures of the simulator and corrects the errors in the simulator caused by poorly captured traffic propagation. The error sources include OD demand estimation errors, path estimation errors, traffic propagation errors, internal traffic model structure errors, and traffic surveillance and estimation errors. The error sources are adjusted by a heuristic approach using PID feedback control [41].

The reactive traffic propagation adjustment module and the PID tuning were tested on a test network through case simulations using DYNASMART software. The authors concluded that the simulator with the built-in reactive traffic propagation adjustment module performed significantly better than without that module. This study of Doan et al. [41] has made an important contribution in identifying the need for calibrating both demand and supply parameters in an online model refinement process. However, the authors did not incorporate OD flows and supply parameter estimation together in an online calibration framework.

In a similar study, Peeta and Bulusu [42] developed a rolling horizon framework for ensuring operational consistency of online DTA. The authors defined consistency in terms of the difference between the predicted and actual time-dependent path flows. Accordingly, the authors formulated the problem as a constrained least-squares model (CONS), which is expressed as a fixed-point problem, and is solved as a bi-level framework. Simulation experiments based on a small synthetic network were conducted to analyze the effectiveness of the proposed solution algorithm that the authors called generalized singular value decomposition (GSVD). The experiments emphasized the reliability and stability of addressing the online consistency problem, in terms of link and path travel times [42]. Although the proposed solution algorithm showed promising results in using the synthetic data, the feasibility of the approach in a real network using real-time traffic data was not investigated.

He et al. [43] developed an integrated online and offline calibration procedure for an analytical dynamic traffic model by minimizing the discrepancy between the model's output and real-world traffic conditions. In this approach, the authors again attempted to identify error sources in a DTA model. The authors only considered three components in their study: 1) dynamic link travel time functions, 2) route choice, and 3) flow propagation models. It is to be noted that OD estimation is absent from the calibration process. Further, the authors suggested that the calibration of the route choice model should take place offline, due to the large amount of the data and the resulting computational limitations. The calibration process proposed by the authors is an iterative approach that sequentially considers the three components until convergence. The authors noted that the calibration process continued until the model parameters represented the true states "reasonably well". However, the convergence criteria were not clearly defined by the authors.

In a subsequent paper, He and Ran [44] attempted to enhance and expand the approach developed in [43]. The authors developed a process to calibrate and validate the dynamic route choice and flow propagation components in a DTA system [44]. The authors used the maximum likelihood technique to calibrate the DTA model and relate the known factors (time-dependent OD, network real-time link traffic counts), and unknown parameters (dynamic route choice probability). The proposed approach was tested on small and larger

networks. The authors pointed out that it is possible to calibrate a DTA model with measurement errors and insufficient traffic data (i.e., when real-time traffic counts are available on only a few links in the network). However, the maximum likelihood estimation for the route choice model requires disaggregate data, which is a major limitation to the practical application of the proposed approach. Similar to their previous study [43], a priori knowledge of dynamic OD matrices is still an important requirement in their more recent research. Furthermore, this research simplified the demand estimation by enforcing temporal independence of OD flows between all OD pairs.

The above-reviewed research works have made important advancements to the challenging online DTA problem. However, none of these studies attempted to simultaneously calibrate the OD estimation, route choice and supply models.

Hawas presented a framework for the calibration of a dynamic traffic simulation model that can be systematically used to identify and quantify major sources of simulation errors [45]. These calibration errors were identified by allocating a rank (integer index) to each of the model processes based on a heuristic ranking approach that was based on the order of the process execution within the model. In other words, the processes with the least rank and, hence, fewer internal interactions, were calibrated first. The calibration methodology was used to calibrate a small test network simulated in microscopic traffic simulation system for integrated modeling and analysis (MITSSIMA) [46]. The presented approach was successfully tested on a sample network with given dynamic OD flows and two calibrated parameters (speed-density functions).

The author conducted an intensive sensitivity analysis by individually perturbing each variable to study its impact on the model's output. However, the computational overhead would potentially increase further if model outputs from multiple replications must be averaged to account for simulator stochasticity. Moreover, the magnitude of the perturbation may be hypothesized to vary across the parameters, given the nonlinear nature of the objective function. A uniform perturbation for all variables is, therefore, not optimal. The author also suggested that the approach should be further tested on a more realistic real network with large sets of parameters.

Other research efforts related to DTA system calibration have focused on independent calibration of demand and supply parameters based on manual adjustments and prior observations. Chen et al. [47] presented preliminary DYNASMART-P simulation results from a case study in Zwolle, Netherlands. The model estimation method relies on manual adjustments of individual model parameters, based on prior knowledge of the network and its traffic scheme. Thus, the model estimation framework is conducted in an ad hoc fashion.

A similar calibration approach was also proposed by Chu et al. [48] using a systematic, multistage procedure for the calibration and validation of PARAMICS simulation models in southern California. The authors combined heuristics and static approaches to assist in the calibration of driving behavior models, route choice model, OD estimation and fine-tuning of various model parameters.

Mahut et al. [49] presented an application and iterative demand and supply calibration of a DTASQ simulation-based DTA model to a part of the City of Calgary's (Alberta, Canada) network. To establish dynamic user equilibrium travel times on the network, DTASQ iteratively combines network loading model parameters (e.g., gap acceptance and lane-changing model parameters) with a route choice parameter based on volume-delay functions. Hourly OD matrices were simulated for the City of Calgary using EMME/2 software. In order to provide a suitable trip table for the DTA model, main intersection turning movement counts were compared to the output of EMME/2. The posted speed limit and the link capacity were used to determine the volume-delay functions, although in uncongested real-world conditions, drivers travel at a speed higher than the posted speed limit. The authors concluded that traffic parameters, such as gap acceptance, can have a significant impact on the route choice in an iterative equilibration approach to DTA. However, these parameters were manually adjusted for this case study, in order to minimize the objective function. The calibration approach was successful for the Calgary network. However, the transferability of the approach may be questionable since many parameters were adjusted manually.

Mahmassani et al. [50] prepared a technical report on the calibration of DYNASMART-X for a southern California network. The authors again focused on independent calibration of different parameters of a DTA system. Greenshield's model was modified and adopted for

calibrating the speed-density function. The general least squares (GLS) approach, as proposed by Cascetta at al., was used for estimation of the time-dependent OD matrices. The calibration approach for the departure choice, route choice and capacity estimation was not considered in the study [22].

Balakrishna et al. [51] presented an iterative offline mesoscopic calibration approach to calibrate various inputs of a DTA model in the DynaMIT model developed by Ben-Akiva et al. [52]. Demand simulation of DynaMIT consists of the following two components: 1) OD estimation and prediction, and 2) driver behavioral models that estimate and predict route choice [53]. As for the supply simulation, segment speed-density relationships and capacities are considered. The GLS method is used for sequential OD estimation, while queuing phenomena and spillbacks are modeled separately using a modified Greenshield's model and the Highway Capacity Manual (HCM) for segment capacity estimation.

The detailed supply calibration approach was presented earlier in [54]. The calibration problem is solved using an iterative bi-level approach consisting of an estimation of OD flows while fixing model parameters in the higher level, followed by the estimation of the model parameters in the lower level assuming fixed OD flows from the upper level. While it is possible in theory to iterate between the demand and supply calibration steps until convergence, such an approach is likely to be computationally inefficient. In addition, the sequential solution methodology only made use of sensor counts for calibration purposes. This limitation is attributed to the use of linear measurement equations that map OD flows to counts. For estimation of the route choice parameters and habitual travel times, the Box complex algorithm [55] exhibited slow convergence. Hence, its use in large networks is not practical.

In his M.Sc. thesis, Gupta [56] attempted to overcome the dependency of the calibration framework on historical OD flows by developing a methodology to test if it is possible to estimate the same OD matrix in the absence of any a priori OD flows. The author used the DynaMIT model and adopted a sequential approach. A case study involving a practically sized network from Los Angeles with multiple days of data was used to validate the framework. This research has made significant contributions through the application of

observability, which allows the modeler to test if unique OD flows can be estimated from the given sensor configuration. However, further testing and improvements should be carried for different networks with different sensor locations.

While Gupta and Balakrishna et al. focused on the calibration of mesoscopic models, Mahanti [57] studied the calibration of a microscopic DTA model, incorporating both demand and supply parameters. The author integrated the bi-level iterative calibration methodology [51] in the MITSIMLab micro-simulation tool with lane-changing and car-following supply parameters. The proposed iterative approach revealed promising microscopic calibration results in a synthetic case study and also for a real-sized network. The authors suggested further research in this direction, while incorporating scheduled events and incident scenarios in the calibration framework.

In a similar paper, Toledo et al. [58] formulated an iterative bi-level framework for the calibration of microscopic traffic simulation models using aggregate data in MITSIMLab. The problem is formulated assuming a stationary steady-state condition. The assumption is that the observation days are drawn during a period in which steady-state traffic conditions prevail, i.e., while OD flows and experienced travel times may vary for various observation days, these differences are due to random effects and do not represent a change in the underlying distributions of these variables. Furthermore, the authors assumed that driving behavior and route choice parameters are stable over the period of observation.

In a subsequent paper, the same bi-level calibration framework was applied for a test network in Stockholm, Sweden, under congested traffic conditions using sensor data [59]. Later, Darda [60] used the same calibration framework for a different network in Irvine, California. Although the demonstrated approach was successful in replicating real-time traffic conditions, it is computationally burdensome.

Using a similar methodology, Jha et al. [61] highlighted the challenges related to large-scale traffic simulation, such as data collection, computational requirements, conversion of planning OD to simulation OD, and impact of small errors on significant additional efforts in calibration. Each component within the calibration module relies on some convergence

criterion, which was not clarified. The impacts of various levels of aggregation, such as time intervals for habitual travel time and the OD flows, should be considered.

Kim [62] developed an iterative bi-level framework to calibrate both a microscopic traffic simulation model and an OD estimation problem. At the upper level, the best model parameters are identified using a genetic algorithm (GA) optimization tool and disaggregate data (travel time from automatic vehicle identification data); whereas, at the lower level, OD matrices were calibrated using the extended Kalman filter (EKF) algorithm. VISSIM software was used to evaluate the proposed methodology for two test networks: one consisting of an urban arterial, and the other a freeway. The results of the experiments demonstrated the advantages of incorporating disaggregate data to improve the accuracy of joint demand and supply calibration. However, both test networks were simple with only four automatic vehicle identification (AVI) stations. The proposed bi-level calibration approach can be further tested in longer freeways and larger networks. Table 2.1 summarizes the results of several calibration exercises involving separate and iterative demand and supply parameters estimation.

Table 2.1 Summary of iterative calibration research literature results

Category	Offline/ Online Framework	Study	Type of Framework (solution algorithm)	Model	Network Type	Measure of Performance
Mesoscopic	Online consistency problem	Doan et al. [41]	Online reactive traffic propagation adjustment module (PID)	DYNASMART	Arterial	Link congestion
		Peeta and Bulusu [42]	Rolling horizon framework (GSVD)	RH-DTA model	Synthetic freeway/ arterial network	Link and path travel times
	Offline	Hawas [45]	Heuristic approach	MITSSIMA	Synthetic small network	Travel time
		Chen et al. [47], Mahmasani et al. [50]	GLS for OD estimation, Greenshield's model for speed-density relation	DYNASMART-X	Arterial	Traffic counts
		Gupta [56]	GLS for OD estimation, Greenshield's model for speed-density relation, capacity from HCM	DynaMIT	Freeway/ arterial	Traffic counts
		Balakrishna et al. [51]	GLS for OD estimation, modified Greenshield's model for speed-density relation, capacity from HCM	DynaMIT	Freeway/ arterial	Traffic counts/ speed
Microscopic		Chu et al. [48]	Heuristics and static approaches / manual fine-tuning method	PARAMICS	Freeway/ arterial	Travel time/ counts
		Mahut et al. [49]	Iteratively combines network loading model with a route choice parameter based on volume-delay functions	DTASQ	Freeway/ arterial	Traffic counts
		Mahanti [57]	GLS for OD estimation, box-complex algorithm for other parameter	MITSIMLab	Arterial	Traffic counts
		Toledo et al. [58, 59], Darda [60]		MITSIMLab	Freeway/ arterial	Traffic counts/ speed
		Jha et al. [61]		MITSIMLab	Freeway/ arterial	Traffic counts
		Kim [62]	EKF for OD estimation, GA for other parameter	VISSIM	Freeway	Travel time from AVI/ counts

2.3.2 Simultaneous Calibration of the Demand and Supply Parameters

Contrary to sequential calibration approaches, only a few recent studies have focused on simultaneous calibration methodology that is capable of jointly estimating both demand and supply parameters. Figure 2.2 presents a general flowchart for simultaneous demand and supply DTA calibration in offline and online contexts. As the figure indicates, in this approach, the calibration of the DTA model is mainly formulated as an optimization problem that attempts to jointly estimate the OD matrices and the driver behavioral parameters.

The problem is formulated either as an offline and/or online calibration framework in the context of either a mesoscopic or a microscopic calibration. In the context of online calibration, the output of simultaneous offline calibration is used as the input for the online calibration framework. Data used for this purpose is derived from both historical and real-time information extracted from various sources and sensors.

Offline calibration of DTA models as an optimization problem was first proposed in [63] and [64]. The optimization can jointly estimate the demand and supply parameters using archived flow data [63, 64]. Offline calibration results in the creation of a simulated "historical" database. The authors adopted an error minimization framework for the simultaneous calibration of 1) a DTA model's demand (i.e., OD flows and driver route choice behavior), and 2) supply components (i.e., speed-density relationship and segment capacity estimation) in a mesoscopic simulator, DynaMIT. The stochastic optimization framework was solved using various solution methods, namely: the Box complex [55], stable noisy optimization by branch and fit (SNOBFIT) [65], and simultaneous perturbation stochastic approximation (SPSA) [66] algorithms. These different solution algorithms were tested on a simple network with synthetic data, and on a medium-scale simulated network model for Los Angeles, California. Among the various solution methodologies, SPSA was found to be the best performing practical method. The results also indicated that the simultaneous calibration approach outperforms the traditional sequential method.

Balakrishna's work [63, 64] pioneered the calibration framework that incorporates the simultaneous demand and supply parameters estimation. Although the impact of an incident

was demonstrated in a synthetic case study, the capacity reduction factor was not considered in the real test network. A realistic estimate of the reduction in capacity is essential for maintaining the accuracy of the system's calibration and predictions capabilities. Finally, traffic counts and speed were the only available traffic data used in the offline calibration process. However, the authors noted that point-to-point data observations recorded through AVI or Global Positioning System (GPS) technologies may further improve the efficiency of the estimated parameters, especially for demand estimation.

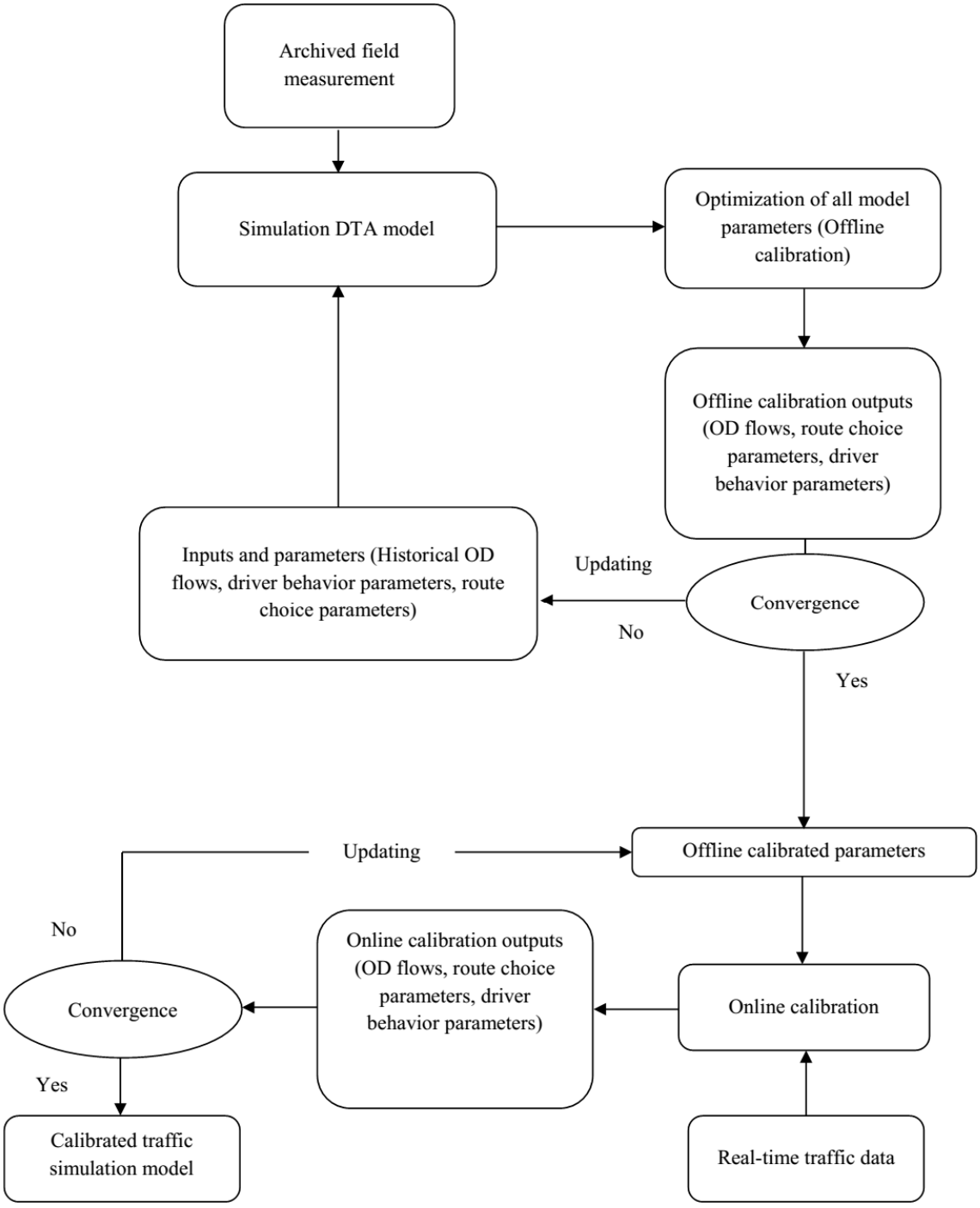


Figure 2.2 Flowchart of simultaneous demand and supply calibration

Balakrishna's study [63, 64] focused on drivers' pre-trip route choice behavior, while capturing their departure time preferences implicitly through the dynamic OD flows. Other extensions as related to commuter's response to en-route information (e.g., variable message

sign, in-vehicle navigation systems, radio, etc.) can also be incorporated in the calibration framework, as proposed by [67] and [68]. Pel et al. [67] introduced a hybrid route choice model, where all travelers have pre-trip route information, but also consider real-time traffic conditions in seeking new routes. This model, however, requires intensive path enumeration; therefore, its application to large-scale network may be problematic. Rather than treating all travelers identically, Zhen and Zhang [68] assumed that some travelers are likely to follow their predetermined routes, while others update their routes en-route in response to real-time information.

As an extension of previous work, Balakrishna et al. [69] used a similar framework for offline simultaneous demand and supply parameter calibration and applied it in a MITSIMLab traffic simulation model. SPSA was chosen as a solution algorithm for solving the stochastic nonlinear optimization problem. Comparing different measures of effectiveness revealed that simultaneous calibration again outperforms the iterative approach. However, the case study examined in this research was a highly instrumented network with relatively high level of sensor coverage, which is uncommon in real-world situations. Thus, the presence of lower observations and sensor coverage may affect the efficiency of the proposed methodology.

Using the same input parameters used by Balakrishna for offline DTA calibration, Antoniou [70] developed an online calibration framework for a mesoscopic model. Online DTA applications require accurate real-time estimation and predictions of traffic conditions, such as the impact of weather, road surface conditions, and incidents. Results from offline calibration are used as a priori information for the online calibration process. These parameters are adjusted in real time, based on updated measurements, to reflect the realistic deviation of traffic conditions from their average values.

Antoniou and Antoniou et al. [70, 71] attempted to formulate this mesoscopic calibration problem as a state-space modeling concept. The authors examined the application of EKF, limiting Kalman filter (LimKF) and unscented Kalman filter (UKF) to solve the optimization problem. Empirical results on 35 kilometers of a simulated freeway network with eight on-ramps, seven off-ramps and 20 OD flows suggested that joint online calibration of demand

and supply parameters can increase the estimation and prediction accuracy of a DTA system compared to its sequential calibration counterpart. Further, this approach was shown to outperform the DTA framework where only OD flows are calibrated online [70, 71, 72]. However, the proposed on-line calibration approach was only demonstrated on a freeway network. Thus, the findings cannot be generalized, as further tests should be conducted to examine the transferability of the approach to more general and larger sized networks. Further, the authors suggested that the variation of traffic due to unexpected events or incidents may also change driving behaviors, which can result in additional parameters to be calibrated as part of the online calibration framework.

Other studies have focused on simultaneous offline calibration of demand and supply parameters using point-to-point AVI disaggregated traffic data in DynaMIT [30, 73]. Following the methodologies proposed in [66] and [70], the calibration problem was formulated as a stochastic optimization and a state-space framework. SPSA and GAs were chosen as the candidate solution algorithms. Further, particle filters (PF) were considered as the solution method for the state-space framework. The authors compared the calibration results using travel time measurements obtained from emerging traffic-sensing technologies with traditional loop detector data. The calibration results from a synthetic case study revealed that SPSA and GA were more effective than PF for solving the multi-objective optimization problem.

The methodology was also applied to a real traffic network to demonstrate its scalability. The calibration results suggested that incorporating AVI data in the optimization problem improved the calibration accuracy. In addition, in accordance with the previous studies, the simultaneous demand and supply calibration problem was found to be superior to the iterative approach. It is to be noted that all previous studies were restricted from incorporating disaggregate AVI data into the OD estimation problem. Thus, Vaze's work [30, 73] is considered to be among the first research efforts that take into account AVI data and travel time information in the multi-objective formulation. This research can be further extended by incorporating additional information from probe vehicle or floating car data. Moreover, the sensitivity of the algorithm to other key parameters can be examined: such

factors may include detectors locations, market penetration and the number of AVI sensors and their location.

In a recent study, different solution algorithms for the online DTA calibration problem in the DynaMIT system were compared [74, 75]. The authors formulated the problem as a state-space (i.e., EKF algorithm) and direct optimization formulation, using the Hooke-Jeeves pattern search (PS) algorithm, conjugate gradient (CG) and gradient descent (GD) algorithms. The authors incorporated various real-time network data from various sources, namely loop detectors, video cameras and toll counters from an intercity highway in Portugal. In fact, the application of the Hooke-Jeeves algorithm [76] was examined earlier by the same authors to the online calibration problem [77]. The results from these studies revealed that real-time data fusion outperformed the offline calibration of DTA without online adjustment. However, the above-examined algorithms were found to be computationally burdensome and became intractable when dealing with a large-scale network.

To overcome these issues, the authors proposed parallel GD (para-GD) and parallel EKF (para-EKF) algorithms. Parallel implementation accelerates the objective function evaluation to reduce the computation time at different stages of the algorithm. The application of parallelization for EKF and GD algorithms revealed promising results, in terms of computational time and calibration accuracy, compared to serial EKF and GD. As two extensions of this research, the authors proposed parallel computation within the DTA jointly with parallel implementation of the calibration algorithms, along with hybrid parallelization. The solution algorithms should be further validated in a larger scale network.

Appiah and Rilett [78] presented a framework for joint OD estimation and calibration of microscopic models using vehicle trajectories from aggregate intersection turning movement counts. The authors developed a methodology for OD estimation problem independent of the presence of a prior historical matrix. As opposed to [62], OD flows were treated as unknown in order to be jointly calibrated with the driver behavior parameters (i.e., car following and lane changing). The problem was formulated as an optimization framework, and GA was adopted as the solution algorithm using a VISSIM microscopic model. The developed

methodology was examined on a small urban network. Although GA has advantages in dealing with non-convexity, locality and complex transportation optimization problems, the final results indicated that the algorithm has a high computational cost, even for a small arterial network. The GAs took approximately 2 months and 18,000 iterations to converge. Some parallelization and hybrid parallelization can be adopted for more complex networks. The final results revealed a strong correlation between the observed and simulated counts. Despite the high computational time, the calibration results and methodology can be considered appropriate, since around 40% of links have GEH values of less than 5, as recommended. However, the methodology that used vehicle trajectories was limited to local and arterial network with high number of intersections turn count data. Table 2.2 summarizes these studies, categorizing them by formulation, methodology, network type, micro-simulation used, and performance measures.

Table 2.2 Summary of simultaneous calibration research literature results

Category	Offline/ Online Framework	Study	Type of Optimization (solution algorithm)	Model	Network Type	Measure of Performance
Mesoscopic	Offline	Balakrishna [63], Balakrishna et al. [64]	Stochastic optimization (Box complex, SNOBFIT, SPSA)	DynaMIT	Freeway/arterial	Traffic count/speed
	Offline	Vaze [73], Vaze et al. [30]	Stochastic optimization (SPSA, GA) and state-space formulation (PF)	DynaMIT	Freeway/arterial	AVI/count
	Online	Antoniou [70], Antoniou et al. [71]	State-space formulation (EKF, LimKF, UKF)	DynaMIT	Freeway	Traffic count/speed
	Online	Huang et al. [77]	Heuristic pattern search algorithm (Hooke-Jeeves)	DynaMIT	Freeway	Traffic count / toll collection / camera counters
	Online	Huang [74], Huang et al. [75]	Direct optimization (GD, CG, para-GD) and state-space formulation (EKF, para-EKF)	DynaMIT	Freeway (same as Huang et al. [77]	Traffic count / toll collection / camera counters / AVI data
Microscopic	Offline	Appiah and Rilett [78]	Stochastic optimization (GA)	VISSIM	Arterial	Traffic count
		Balakrishna et al. [69]	Stochastic optimization (SPSA)	MITSIMLab	Freeway/arterial (same as Vaze, [73])	Traffic count

2.4 Summary

The estimation of the demand parameters from observation data and the calibration of supply parameters of DTA models have been intensively reviewed in the past [79]; however, these research works were mostly conducted separately, as a fixed-point model to adjust the time-dependent OD matrices [23, 24, 25, 26, 33, 34, 35] or calibrating speed-density models (e.g., traffic flow models in DYNASMART, DynaMIT and Dynameq software packages) [38, 39, 40]. The approach of independent calibration of demand and supply parameters is not efficient or optimal, as it does not consider that the supply parameters may affect the estimated demand and vice versa. These early studies were followed by research efforts that attempted to jointly calibrate the demand and supply parameters in an iterative sequential fashion. Iterative calibration is solved using an iterative bi-level approach, consisting of an estimation of OD flows, while fixing model parameters (e.g., driver behavior and route choice mode parameters) in the higher level and followed by the estimation of the model parameters in the lower level, assuming fixed OD flows from the upper level. It was argued that these models treated demand and supply parameters independently, without consideration of the possible interaction between these parameters [48, 49, 50, 51, 56, 57, 58, 59, 60, 61, 62]. Ignoring the complex and nonlinear interactions between demand and supply parameters may lead to suboptimal solutions. Moreover, such an approach is likely to be computationally inefficient. Compared to sequential calibration approaches, a few research efforts have focused on the simultaneous calibration of demand and supply parameters [63, 64, 69, 70, 71, 74, 75, 77]. In this type of approach, the calibration of the DTA model is mainly formulated as an optimization problem that attempts to jointly estimate the OD matrices and the driver behavioral parameters. The problem is formulated either as an offline and/or online calibration framework in a mesoscopic/microscopic context. However, the adopted calibration framework is highly dependent on the use of local search heuristics. Confining the search in the vicinity of the starting point creates a high dependency on the quality of historical OD information [33]. Thus, these techniques are mostly applicable when the starting point is close to the optimal one. In addition, the simultaneous calibration frameworks proposed in the literature ignores the reliability of the different components of the objective function (e.g. historical OD flows, count data from loop detectors) and requires extensive manual adjustments to identify the

optimal weights. Finally, most previous efforts focused on the loop detector data available at the aggregate level, with traffic data collection mostly relying on surveys and vehicle counts that are costly and time-consuming. They have, therefore, been applied infrequently on a small or a medium-sized network. Given the limited range of data acquired from loop detector counts, the calibration process was highly dependent on the historical informations.

In summary, the review of the literature indicates several shortcomings in the state-of-the-art of DTA model calibration. Particularly, there is need to develop a robust calibration methodology that can simultaneously estimate both demand and supply model parameters in a simulation-based DTA system. Chapter 3 presents a rigorous treatment of the DTA calibration problem, and proposes a robust and systematic estimator for its solution.

Chapter 3: CALIBRATION METHODOLOGY

This chapter is focused on the simultaneous calibration framework of the demand and supply parameters of the DTA model. An overview of the methodology is presented first, focusing on the input and output parameters of the DTA model. This is followed by the proposed formulation of the DTA calibration problem. The third section discusses the simulation environment and the selected parameters for the calibration process. The forth section provides a brief introduction of the solution algorithm, followed by the proposed calibration framework. The final section concludes the chapter with the summary of the overall calibration methodology.

3.1 Overview of the Methodology

The set of critical DTA model parameters that must be calibrated for a specific network can be separated into demand and supply side variables. Demand variables are the time-dependent OD flows for the period of interest, while the number and nature of supply variables may vary depending on the level of detail employed while capturing traffic dynamics and queuing phenomena. Microscopic models generally possess a much wider set of models and parameters that operate under different traffic regimes and explain a complex set of individual driver decisions and maneuvers. These include car-following (acceleration, deceleration and desired speed), lane-changing (gap acceptance, merging, yielding and look-ahead), as well as route choice model parameters. These demand and supply models are mutually dependent on each other. In a general DTA model, various approaches model the interaction between the demand and supply models. The interaction between the supply and the demand is modelled through the traffic assignment process that simulates the propagation of traffic (demand) on the physical network (supply) taking into consideration the drivers' behavior parameters. Historically, the estimation of dynamic OD flows and calibration of driving behavior parameters were treated independently or in an iterative sequential approach. However, the independent calibration approach is not efficient and optimal, as it ignores the presence of interaction among the various demand and supply parameters (e.g. the effect of route choice model parameters on the dynamic OD flows). In addition, ignoring the complex and nonlinear interactions between demand and supply parameters may lead to

suboptimal solutions. Therefore, it is essential to capture the non-linear interactions between demand and supply parameters by formulating the calibration process as a multi-objective framework to simultaneously estimate demand and supply parameters. Figure 3.1 presents the general structure of the DTA models, considering the interactions between the two components of the model.

As stated earlier, the objective of the DTA model calibration is to obtain those model parameter values that will minimize the discrepancy between the observed measurements and their simulated counterpart, when these parameters are used as inputs to the models. The inputs and outputs of the off–line calibration module are outlined in Figure 3.1 and discussed in the following paragraphs.

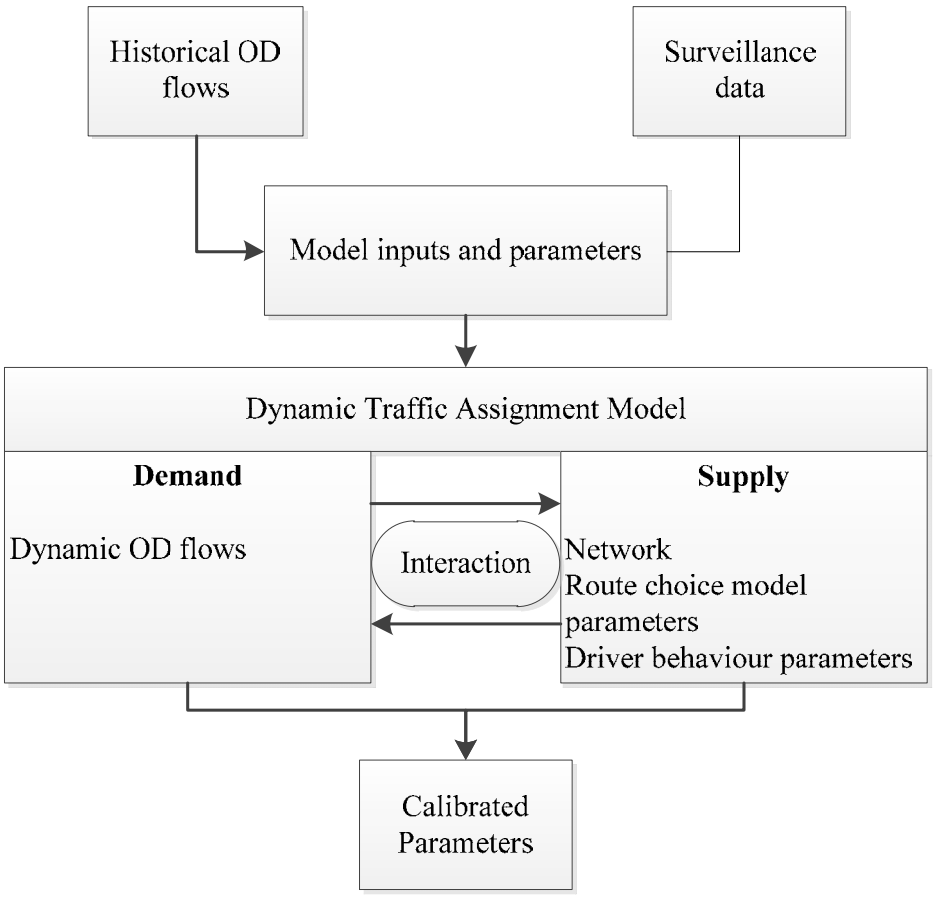


Figure 3.1 Overview of the DTA model

3.1.1 Input Parameters

The DTA calibration component exploits all information that is available within a traffic system, namely historical information describing the transportation system, and surveillance data capturing the prevailing traffic conditions.

Historical information is one of the major components of the DTA models. On the supply side, the historical information can be referred to the geometry of the network, and the traffic control settings. It should be noted that historical data should be periodically updated to capture the dynamics of the network, whenever a new set of information is available.

The available surveillance data from different sources can present the prevailing traffic conditions. As indicated in the literature, the quality and quantity of the available data can directly affect the calibration accuracy, as the system is trying to minimize the discrepancy between these observed data and their simulated counterparts. Recent advances in wireless technologies and use of in-vehicle navigation systems provide the opportunity to improve the quality and quantity of the observed data over a wide spatial area. In addition to the traditional traffic data from loop detectors, this research incorporated the data from in-vehicle navigation system technology into the calibration framework to enhance the calibration process. Chapter 6 of this thesis will provide a more in-depth description of the available surveillance data for the large-scale complex network.

3.1.2 Output Parameters

As stated earlier, the outputs of the DTA calibration problem are the parameter values that minimize the discrepancy between the observed and simulated traffic conditions, when used as input for the traffic estimation framework. On the demand side, the dynamic OD flows that capture the variability of the demand to be loaded onto the network is the output of calibration process. On the other hand, the microscopic driver behavior (car-following and lane-changing) and route choice model parameters of the simulation environment (i.e. Paramics) are among the supply outputs of the calibration.

3.2 Optimization Formulation

3.2.1 General Formulation

As stated earlier, the general calibration problem involves the estimation of OD flows as well as various model parameters using historical OD flows, loop detector counts and turning movements, as well as speed data. The calibration of the DTA model is formulated as a multi-objective optimization problem with the objective of minimizing the discrepancy between the observed and fitted measurement values, as follows:

$$Min\ Z = \alpha_1 f_1(OD^h, OD^{est}) + \alpha_2 f_2\left(Count^{sim}, Count^{obs}\right) + \alpha_3 f_3\left(Turn^{sim}, Turn^{obs}\right) + \alpha_4 f_4(Speed^{sim}, Speed^{obs}) \quad (1)$$

Subject to the following constraints:

$$\left\{Count_t{}^{sim}, Turn_t{}^{sim}, Speed_t{}^{sim}\right\} = Sim\left(OD_1{}^{est}, \dots, OD_t{}^{est}, P_1, \dots, P_t, N\right),$$

$$lb_t^{OD^{est}} < OD_t{}^{est} < ub_t^{OD^{est}},$$

$$lb_t^{P_t} < P_t < ub_t^{P_t}$$

Where:

OD^h, OD^{est}: A priori (historical) OD flows and estimated values, respectively;

$Count^{obs}, Count^{sim}$: Observed link counts and estimated values, respectively;

$Turn^{obs}, Turn^{sim}$: Observed intersection turning counts and simulated values, respectively;

$Speed^{obs}, Speed^{sim}$: Observed travel speed and simulated values, respectively;

f_i: Goodness-of-fit functions (measure of effectiveness);

α_i: Weighting factors;

P_t: Vector of driver behavior and route choice model parameters at time *t*;

lb_t^i: Lower bound of parameter *i* during time interval *t;*

ub_t^i: Upper bound of parameter *i* during time interval *t*; and

N: Network geometry.

In Equation (1), historical OD flows, observed link counts, speed and turning counts on selected intersections are also considered to improve the estimation accuracy. The optimization formulation tries to minimize the discrepancy between the estimated OD flows, OD^{est}, and a priori OD flows, OD^{h}, while trying to incorporate additional information, such as observed traffic counts, $Count^{obs}$, turning counts, $Turn^{Obs}$, and link speeds, $Speed^{Obs}$. The weighting factors, α_i, are determined based on the reliability of observed data (counts, speeds, turns) and historical OD flows. The simulation model, *sim*( ), is a function of the OD flows, the network, *N*, and the vector of driver behavior parameters, *P*. The terms *lb* and *ub* represent the lower and upper bounds, respectively, on the OD flows and model parameters. The f_i functions quantify the discrepancy between the observed and simulated measurements. The goodness-of-fit functions are often described by a sum of squared deviations for sensor measurements, OD flows and model parameters respectively. A good measure of effectiveness, f_i( ), plays a critical role in obtaining good results. In this research, two normalized measures of goodness of fit are used to quantify the relationship between the observed and simulated measurements: the normalized root mean square error (NRMSE), and the Geoffrey E. Havers statistic (GEH). These measures are calculated as follows:

$$NRMSE = \frac{\sqrt{N\sum_{n=1}^{N}(X_n^s - X_n^o)^2}}{\sum_{n=1}^{N} X_n^o} \quad (2)$$

$$GEH = \frac{1}{N}\sum_{n=1}^{N}\sqrt{\frac{2(X_n^s - X_n^o)^2}{X_n^s + X_n^o}} \quad (3)$$

Where

N: Number of observations;

X_n^o: Observed value at time *n*, and

X_n^s: Simulated value at time *n*.

3.2.2 Extension of the General Formulation for the Large-Scale Network

As stated earlier, the weighting factors, α_i, in Equation (1) can be determined based on the reliability of observed data (counts, speeds, and turns) and historical OD flows. Considering the simplicity of this thesis' first case study with synthetic data, the optimal combination of α_i can be estimated by evaluating the objective function with different combinations of weighting factors given to surveillance data (i.e. sensitivity analysis).

In the case of the large-scale network, the reliability of observed data and historical OD flows cannot be easily evaluated. In addition, the sensitivity analysis of the weighting factors cannot produce the optimal solution as the fundamental of the sensitivity analysis is characterized by repeated and varied attempts, which can be computationally intensive. Therefore, the weighting factors were defined as the function of the observed and simulated observed traffic data.

According to the literature, very few studies in the area of multivariate optimization formulations considered different weighting factors for each of the components, and among those studies, the optimal weighting factors estimation was based on the traditional sensitivity analysis [30, 69, 73]. Therefore, to the author's best knowledge, the reliability of different components was not yet attempted as a part of the optimization formulation.

For this purpose, the variance of the Theil's *U* inequality coefficient was used to quantify the reliability of different measurement [69, 1]. This measurement can be quantified as follows:

$$U_j^s = \frac{(\sigma_{ave}^s - \sigma_{ave}^o)^2}{\frac{1}{N}\sum_{i=1}^{N}(X_n^o - X_n^s)^2} \quad (4)$$

Where:

U_j^s: Variance of the Theil's *U* inequality coefficient for measurement *j*,

N: Number of observations,

σ_{ave}^o: Standard deviation of the average observed measurement, and

σ_{ave}^s: Standard deviation of the average simulated measurement,

Generally, this measurement indicates how well the simulation model is able to replicate the variability in the observed data and OD flows. This Theil's *U* inequality coefficient that

incorporated the variance of the observed and simulated data was proven to be a consistent measurement for allocating weights to different measurements in the objective function (i.e. count, speed, turn, OD flows). According to the definition, the variance of the Theil's *U* inequality coefficient should be kept as close to zero (the value is between zero and one). Therefore, the highest weight should be given to the component with the minimum variance of the Theil's *U* inequality coefficient.

In summary, the weighting factors can be mathematically expressed as follows:

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1 \tag{5}$$

Where:

$$\alpha_1 = \frac{1/U_1^s}{\sum_{j=1}^{4}(1/U_j^s)} \tag{6}$$

$$\alpha_2 = \frac{1/U_2^s}{\sum_{j=1}^{4}(1/U_j^s)} \tag{7}$$

$$\alpha_3 = \frac{1/U_3^s}{\sum_{j=1}^{4}(1/U_j^s)} \tag{8}$$

$$\alpha_4 = \frac{1/U_4^s}{\sum_{j=1}^{4}(1/U_j^s)} \tag{9}$$

As is apparent from the above equations, the weighting factors can be expressed as the functions of U_j^s. Therefore, the weighting factors were incorporated into the calibration process and their optimal values were obtained as a part of the calibration process. In summary, the modified multivariate optimization formulation can be expressed as follows:

$$Min\ Z = \alpha_1 f_1(OD^h, OD^{est}) + \alpha_2 f_2\left(Count^{sim}, Count^{obs}\right) + \alpha_3 f_3\left(Turn^{sim}, Turn^{obs}\right) + \alpha_4 f_4(Speed^{sim}, Speed^{obs}) \quad (10)$$

Subject to the following new constrain:

$$\sum_{i=1}^{4}\left[\frac{1/U_i^s}{\sum_{j=1}^{4}(1/U_j^s)}\right] = 1$$

To reduce the computation time in the real-world complex network, the NRMSE was chosen as the goodness-of-fit. It should also be noted that the above objective function can incorporate any type of traffic data into the calibration process without any limitations. The following section provides a brief introduction of the simulation environment and the associated parameter for calibration, *P*.

3.3 Simulation Environment

The methodology described in this research is based on the Paramics micro-simulation package developed by Quadstone in Scotland. Paramics has a set of driver behavior and route choice parameters that need to be adjusted for the specific study network and the intended applications, in order to accurately replicate the field data. In this research, the vector of driver behavior parameters, *P*, includes mean headway, mean reaction time, perturbation, feedback, and network familiarity. The first two supply parameters influence network-wide vehicle-driving behavior, and the other three are route choice model parameters. These five sensitive parameters are defined as follows:

- The mean headway is the average time between the leading edges of successive vehicles (default value: 1 second);

- Mean driver reaction time is the value is associated with the lag in time between a change in speed of the preceding vehicle and the following vehicle's reaction to the change (default value: 1 second);

- The perturbation factor is used to randomize the route-cost perception to affect a stochastic route choice (default value: 5%);

- Feedback is a loop mechanism used to update travel time costs for equipped drivers throughout a simulation period to influence route choice (default value: 300 second); and,

- Familiarity is a factor that affects the route choice and describes the composition of drivers with respect to their different levels of knowledge of the network (default value: 85%).

In this thesis, the search space for driver behavior parameters is, therefore, five dimensional. The searching range of each parameter is decided either by rules of thumb or from the Highway Capacity Manual [80]: mean headway = 0.5 to 1.5 s, mean reaction time = 0.4 to 1.6 s, feedback = 1 to 5 min, perturbation = 1 to 100%, familiarity = 1 to 100%.

3.4 Overview of the Solution Algorithm

The non-linear interactions between demand and supply parameters present a major challenge in solving the optimization problem described above. Similar to the most real life problems, the DTA models are stochastic in nature and cannot easily be represented as an analytical closed-form function of the decision variables. This non-analytical nature of the problem precludes the possibility of differentiation or exact computation of local gradients of objective functions or constraint expressions. Therefore, most of the optimization literatures are not directly suited for the solution of the DTA model calibration problem, and the modeller must turn to simulation optimization methods.

According to the literature, the candidate solution algorithms for the stochastic non-linear problems can be classified into path search, pattern search and random search techniques. Table 3.1 provides a summary of the solution techniques and the candidate algorithms belonging to each of these types. Readers are referred to Balakrishna (2006) and Vaze (2007) for the detailed description of the candidate solution algorithms for the DTA models [63,73].

As is apparent from Table 3.1, a wide variety of simulation optimization algorithms exist in the literature. However, a few of them have been tested on even medium-sized networks. In addition, it was found that path search algorithms became computationally burdensome when dealing with large-scale networks [74, 75]. Recent related studies using GA have shown

advantages in dealing with non-convex and complex transportation optimization problems [81, 82].

This research proposes GA as the basic solution algorithm for the estimation of OD flows and calibration of driver behavior parameters. This GA-based approach is further enhanced by distribution and parallelization schemes to accelerate the convergence of the objective function and reduce computational time. In addition, the GA-based model has a wide variety of selection mechanism, real-coded evolutionary operators, as well as parallel and distributed structure. The detailed descriptions of the GA operators as well as distribution and parallelization schemes are provided in Chapter 4.

Table 3.1 Summary of the classification of optimization algorithms [28, 63, 70, 73, 74, 78, 79, 81]

Type of optimization algorithms	Description	Sample solution algorithms	Discussion
Path search algorithms	• Use an initial point (or a population of several points) in the search step to begin with; • The algorithm keeps moving the current point in a certain direction with the purpose of improving the objective function value; • The gradient of the function is used directly or indirectly to determine the direction of movement.	• Response surface methodology (RSM) • Stable Noisy Optimization by Branch and Fit (SNOBFIT) • Finite Difference Stochastic Approximation (FDSA) • Simultaneous Perturbation Stochastic Approximation (SPSA)	• RMSA is not directly applicable for constrained optimization problem • SPSA was successful in solving for large scale DTA problems; outperforms FDSA and SNOBFIT.
Pattern search algorithms	• Referred to the direct search methods as they do not require any gradient calculations; • Some patterns are used to obtain an improved solution in each iteration	• Hooke and Jeeves method • Downhill simplex method	• Poor convergence to the global optimal and difficulties in handling stochasticity in the DTA models.
• Random search methods	• Direct method as it does not require derivatives to search a continuous domain; • Probabilistic mechanisms to randomly select updated parameter vectors to improve towards an optimal	• Simulated Annealing (SA) • Genetic Algorithm (GA) • Parallel GA	• SA has a poor convergence speed even for small scale problems especially in case of noisy function measurements and continuous variables. • GA has the ability to reach the global optimal efficiently and has been successfully applied to the model calibration in the context of transportation. • Parallel GAs were used to accelerate the convergence of the objective function and reduce computational time.

3.5 Calibration Framework

Upon selection of GA as the solution approach, the simultaneous calibration framework of the demand and supply parameters is created. Figure 3.2 presents the conceptual flowchart of the simultaneous OD estimation and calibration procedure of Paramics' driver behavior parameters utilizing a GA.

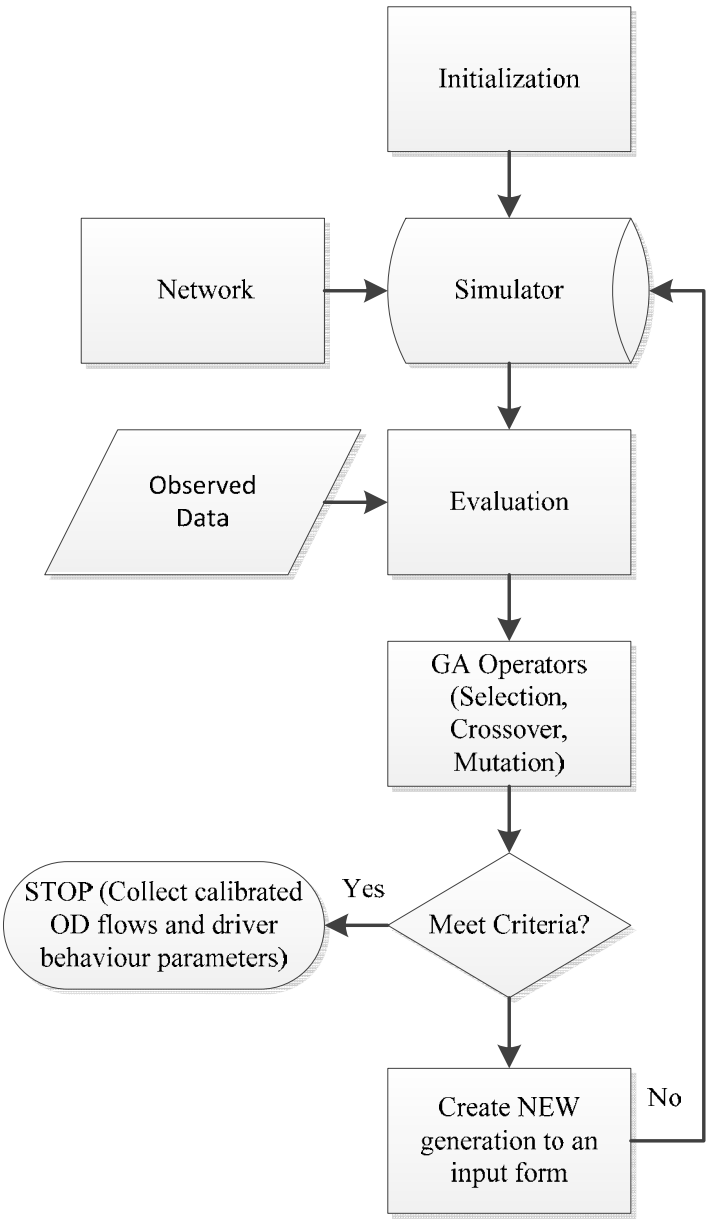


Figure 3.2 Flowchart of OD estimation and calibration of driver behavior parameters

The framework starts from a population of initial solutions, each with a vector of candidate ODs, for several interval estimation steps and the route and driver behavior parameters to be calibrated. To move away from the need for gradient estimation and to better capture the complex relationship between the vector of unknown parameters to be calibrated and the observed data, the framework directly uses the output of a network-loading model in Paramics. Thus, each candidate solution vector of demand and supply parameters and its corresponding assigned outputs are simultaneously evaluated. In other words, the resulting assigned flows, turning counts and simulated speeds for each candidate demand and supply parameters solution are selected for computation of the fitness function. It should be noted that, due to the stochastic nature of Paramics, in all runs conducted in this thesis, a pre-specified number of runs (e.g. 5) were performed, and the average values of outputs are used. The GA then undergoes a series of selection, crossover and mutation processes to generate a new population of new solutions that are again sent to Paramics and for evaluation. The cycle continues until the stopping criteria are met. The final output of the algorithm is the estimated OD flows and calibrated driver and route behavior parameters for Paramics.

As noted earlier in this chapter, the calibration of DTA model was formulated as a multi-objective optimization problem (Equation 1), with relative weights given to different terms of the objective function. One of the advantages of such a formulation is that any type of traffic data can be incorporated into the calibration process without any limitations. In this research, the traffic data for the real-world network was available from three different sources: loop detector counts, turning movement counts at signalized intersections, and speed data from in-vehicle navigation system technology. Recently, many jurisdictions and municipalities across Canada became interested in deployment of the Bluetooth receivers along key corridors to obtain the historical and real-time traffic data. The Bluetooth technology along with other emerging sources (e.g. tracking mobile-phones with GPS) provided the opportunity to obtain traffic information over a wide spatial area at relatively low cost. Therefore, the proposed optimization framework can incorporate the traffic data from these new sources (upon availability) and significantly improve the calibration accuracy by minimizing the discrepancy between the observed traffic data and their simulated counterparts.

While adding the new sources of data into the optimization framework can significantly improve the quality of the solution, the search space for finding the global optimal solutions increases. In other words, as the problem tends to grow, the conventional search algorithms that require gradient information of the incidence matrix[1] become computationally intensive. Historically, GA has been successfully applied to both large-scale calibration and general optimization problems by searching the problem domain thoroughly, based on population-search methods rather than on a single solution, and employing heuristics to evolve towards a better solution.

In addition, the quality of the solution and convergence speed of GA can be further enhanced by running GA in multiple processors (i.e. distributed GA, DGA), parallelization of the GA population into multiple demes (i.e. parallel GA, PGA), and a combination of the PGA utilized with a DGA (i.e. PDGA). These flexibility of the GA structure made it more desirable comparing to other solution algorithms, especially for a large-scale network. Therefore, GA was selected as the solution algorithm and a generic framework was developed that can incorporate several types of traffic information derived from different sensors/sources and with different levels of accuracy. It is noted that the general calibration process can be applied to any microscopic DTA simulation environments (e.g. VISSIM).

As stated in Chapter 2, one of the challenges associated with the multi-objective optimization problem was to allocate the proper weights for different components of the objective function. Historically, the optimal values were estimated based on the sensitivity analysis of a limited number of weighting factor combinations. In some other studies, equal weighs (i.e. 1) were allocated for different components of the objective function. The proposed calibration framework incorporated the weighting factors into the calibration process as a function of observed measurements (i.e. OD flows, count, and speed). In other words, a mathematical formulation was defined to quantify the “reliability” of the observed measurements, based on the average and variance of both observed and simulated

[1] The matrix that shows the relationship between the traffic counts of each link segments and OD flows

measurements. The proposed framework is flexible to incorporate any other measure of reliability into the calibration process.

Finally, it should be noted that the outputs of the calibration process are off-line dynamic OD flows and simulation model parameters that can accurately replicate the historical traffic conditions. In other words, the calibration process can generate a strong state of the knowledge of the subject network. In the next step, the off-line calibrated parameters can be used as *a priori* estimates for on-line calibration process. In other words, the outputs of the off-line calibration process can be used for real-time OD estimation along key corridors, incident managements and reduce unexpected congestions, and ultimately provide real-time traffic data to travelers. It is expected that the convergence speed of the on-line calibration process would be significantly less than the off-line process (i.e. in terms of CPU time) as the on-line calibration process would try to update the most recent and reliable off-line OD flows and other model parameters.

3.6 Summary

The simultaneous calibration of the DTA model parameters was formulated as an optimization problem. This formulation is flexible to incorporate any type of traffic data and historical information into the calibration process without any limitations. Since the DTA models are stochastic and non-linear in nature, the optimization problem can be expressed by the simulation methods. According to the literature, a genetic algorithm (GA) is capable of dealing with the non-convex and complex transportation optimization problems with large sets of parameters. Following the selection of GA as the solution algorithm, the simultaneous calibration framework of the demand and supply parameters of the DTA model was developed. Chapter 4 provides a more in-depth review of the basic GA, the parameter configurations, as well as parallelization and distribution of GA for dealing with the large scale complex networks.

Chapter 4: THE GENETIC BASED SOLUTION APPROACH

The previous chapter described the overall optimization framework for formulation of the calibration problem. This focuses on the specific GA methods implemented in this thesis. Section 4.1 presents a brief overview of GA explaining the fundamental operational mechanisms of GA. Section 4.2 and Section 4.3 elaborate on the advanced GA-based methods for large-scale optimization problems, based on distribution and parallelization schemes. Finally, Section 4.4 summarizes this chapter.

4.1 Overview of Genetic Algorithm

Genetic algorithms (GAs) are classified as evolutionary search method, based on the theory of natural selection. The GA was developed by John Holland in the early 1970s at the University of Michigan [83]. During the past four decades, with the growing demand on combinatorial optimization problems, GAs can successfully be applied to large-scale optimization problems [62, 73, 84]. By simulating natural evolutionary processes, a GA can effectively search the problem domain thoroughly, based on population-search methods rather than on a single solution, and employ heuristics to evolve towards a better solution. The capability of restarting the iterative search from a variety of starting points prevents entrapment in local optima, thereby allowing GAs to prevail over conventional search methods, with no need for gradient information.

Recently, there has been growing interest in using GAs for data fusion, as many applications in information fusion can be stated as complex optimization problems [85]. This is mainly due to the fact that GAs are inherently parallel in nature and are able to deal with difficult optimization problems having complex nonlinear and/or non-differentiable objective functions with complicated constraints and with non-homogeneous, and noisy information. In addition, as stated earlier, the GA-based global search technique has the ability to calibrate the supply parameters for traffic micro-simulation models for intelligent transportation system (ITS) applications, such as Paramics [84]. In the microscopic context, the appropriate magnitude set of driver behavior and route choice model parameters can highly affect the

final network performance, as measured by the relevant fitness function. Therefore, the application of GA-based approaches, as a natural evolutionary technique, can efficiently find the optimal solution for parametric optimization problems.

Generally, the evolution starts from a population of randomly selected individuals in generations. After evaluating an initial population (called chromosome), a series of genetic operations, namely selection, recombination, and mutation, work on the population to create a sequence of populations with increasingly enhances solutions. The selection procedure creates a new population for the next generation, ensuring that only good chromosomes are retained. Recombination produces new generations by exchanging genes among the chromosomes. In addition, Mutation is a genetic operator used to maintain genetic diversity from one generation of a population of genetic algorithm chromosomes to the next. This step will provide a mechanism for the algorithm to escape a local optimal [86]. After a new set of chromosomes is created by applying the GA operators, the new population is then used in the next iteration of the algorithm. Commonly, the algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population.

As stated earlier, the basic form of GA involves three types of operators: selection, crossover, and mutation. In addition, the chromosomes in a GA population can take the form of bit strings i.e. sequences of 0 and 1, or real values. Table 4.1 provides a summary of the GA control parameters identified in the literature. It is noted that the selection of the GA operators and the chromosome representation method is based on the extensive review of literature summarized in Kattan (2005) and Mohamed (2007) [81, 87]. In the following, the selected configurations of the genetic operators are briefly discussed.

Table 4.1 Summary of GA control parameters [81, 87]

Control parameter	Types	Description
Chromosome Representation	Binary Coded Genetic Algorithms (BCGAs)	The problem is encoded by representing the elements of search space (phenotype) into the binary alphabet (genotype).
	Real Coded Genetic Algorithms (RCGAs)	Each point in the genotype is represented as a vector of real numbers; the solution chromosome length is similar to the optimization problem dimension
Selection Method	Roulette Wheel Selection	The selection uses a simulated roulette wheel with slots that are sized according to the fitness of each individual. The roulette is rotated once for each individual to be selected.
	Linear Ranking Selection	Chromosomes are ranked in descending order of fitness, with the ranks of n and 1 given to the best and worst chromosomes, respectively.
	Tournament Selection	A sample of q individuals is taken randomly from the population and the fittest chromosome passed to the intermediate population p'.
	Truncation Selection	Individuals whose fitness is above a certain value (the truncation point) are selected as parents for the next generation.
Recombination/Crossover Operators	Arithmetic Crossover	The two offspring are generated from a uniformly distributed number of the parent's chromosomes.
	α - Blend Crossover	The two offspring are formed by generating two uniform numbers from a blended interval.
	Linear Crossover	This crossover produces three offspring. Then, selection mechanism is applied to choose two offspring for the next generation.
Mutation Operators	Random Mutation	Creates a new gene value from its boundary interval. This method is equivalent to a random initialization
	Non-Uniform Mutation	Utilizes a step-size control mechanism by presenting a dynamic mutation operator that changes its behavior over the course of evolution.
	Fixed Gaussian Mutation	The randomly selected chromosome element is mapped to another real number by a random Gaussian number with a mean zero and fixed user defined standard deviation parameter.
	Self-Adaptive Gaussian Mutation	As opposed to Fixed-Gaussian method, the mutation operator adjusts the gene value in the course of evolution and adapts the Gaussian variance.

4.1.1 Chromosome Representation

According to the literature, there are many representations of GAs, although the most common representations that cover any formulation of most optimization problems are binary and real list encodings. The real encoding has a number of advantages over the binary-encoded GAs, namely [81, 87]:

- Real encoding of GA makes it possible to optimize the problems with large dimension domains. The binary coded GA (BCGA) performs poorly because of the enormous number of units needing to be changed to evolve from generation to generation.

- BCGAs have problems with continuous search spaces when a reasonable precision is required in the application.

- BCGAs are facing the *Hamming cliff* is problem, as it is caused when the binary coding for adjacent values differs completely (e.g. the neighbouring values of OD flows and driver behavior parameters).

In addition, the estimation of OD flows and calibration of driver behavior parameters can be in the order of 1000 decision variables, formulated as a non-convex objective function that requires a traffic simulator to solve the problem. Therefore, achieving an efficient and accurate solution for a large-scale and complex application is of primary importance. As a result, the real coded genetic algorithm (RCGA) was chosen over the BCGA, which would be problematic with continuously high dimensional problems.

4.1.2 Selection

In genetic algorithms, the selection mechanism chooses the best candidate solutions with higher fitness values, to pass their genetic information from one generation to the intermediate population for the genetic operators to exchange their traits aiming for better solutions. For the purpose of this study, the proportional selection mechanisms were potentially selected to ensure that individuals with fitness values higher than the average population fitness would contribute more to intermediate population for reproduction to the next generation. Roulette-wheel [83] and linear ranking [88] mechanisms are among the most common selection algorithms of this kind. In the former, chromosomes are mapped onto a roulette-wheel, where each chromosome occupies an area that represents the probability of selecting the chromosome. The selection mechanism copies chromosomes to the intermediate population by spinning the wheel repeatedly until the population is filled [89]. However this method may dominate the selection, where:

- one or a few chromosomes with super fitness functions (compared to other individuals) propagate from one generation to another; or
- the fitness values of the population chromosomes are close, which leads to equal probability of selection.

In either cases, the GA sticks in a local optimal rather than searching for the global optimal solution.

Unlike the roulette-wheel method, the ranked-based selection mechanism calculates the chromosome selection probability based on its ranking, not the fitness value. In this method, chromosomes are ranked in descending order of fitness, with the ranks of n and 1 given to the best and worst chromosomes, respectively. Roulette-wheel selection is then performed with the probability of selecting the individuals based on the associated ranking values. As a result, the linear rank selection avoids getting stuck in the local optimal solution. Therefore, this method was chosen over the roulette-wheel selection mechanism.

4.1.3 Crossover

After the selection step, the crossover operator acts on the pairs of parent chromosomes in the intermediate population by combining their traits to form two new offspring, based on pre-defined crossover probability rate, P_c.

As indicated in Table 4.1, the arithmetic crossover operator produces gene values residing in the two parents' gene intervals. The linear crossover ensures that the resulting offspring genes are out of the exploitation area. Therefore, while population diversity is always achieved, the quality and convergence time may be lower. On the other hand, the α-blend crossover produces offspring values not necessarily limited to the parents' intervals. The effect is to increase the population diversity and decrease the possibility of premature convergence toward local optima.

Consider two parents chromosomes: $C_1 = (C_1^1, C_2^1, \dots, C_l^1)$ and $C_2 = (C_1^2, C_2^2, \dots, C_l^2)$

Where:

C_j^i: j^{th} gene in chromosome i; and

l : Length of each chromosome.

In the α-blend crossover, two offspring are formed by generating two uniform numbers from the following interval:

$$[(1-\alpha)C_{min} - C_{max}\ , (1-\alpha)C_{min} + C_{max}] \tag{11}$$

Where

$C_{min} = \min(C_j^1, C_j^2)$

$C_{max} = \max(C_j^1, C_j^2)$

According to the literature, the $\alpha \leq 0.5$ was recommended to increase the search area from the two parents' boundaries.

4.1.4 Mutation

Mutation is a genetic operator used to maintain genetic diversity by arbitrarily altering one or more genes of a selected chromosome so as to increase the structural variability of the population. The purpose of mutation is to allow the algorithm to avoid sticking in the local optima. Each allele in every chromosome in the population undergoes a random change based on the GA parameter defined by a mutation probability (P_m). Practitioners and researches considered mutation as the secondary operator in order not to lose the fitness potential are in the search space. Therefore the probability of mutation is usually small. For the purpose of this research, the mutation probability of 5% is selected.

There are a large number of mutation operators presented in the literature, summarized in Table 4.1. Among different mutation operators, self-adaptive Gaussian method is used to control the mutation step-size by adapting the mutation control parameters and the genes values concurrently. The method evolves the chromosomes traits and σ as follows [90, 91]:

$$\sigma_i' = \sigma_i + \exp[\tau'.N(0,1) + \tau.N(0,1)] \quad (12)$$

$$C_i' = C_i + \sigma_i'.N(0,1) \quad (13)$$

Where $N(0,1)$ is a random Gaussian number with a mean of zero and standard deviation of 1. Also τ and τ' are the learning rates, calculated as a function of chromosome length, n.

$$\tau = \frac{1}{\sqrt{2\sqrt{n}}} \text{ and } \tau' = \frac{1}{\sqrt{2n}} \quad (14)$$

4.2 Advanced GA Methods for the Large-scale Optimization Problems

As stated earlier, GAs are capable of finding good solutions to practical engineering and science applications in a reasonable amount of time. However, in cases of large-scale evaluations, GAs may require hundreds of expensive fitness evaluations; and, depending on the cost (time) of individuals' fitness evaluations, GAs may take days or even months to find an acceptable solution. Therefore the following established techniques are identified to enhance the basic solution algorithm (i.e. GA):

- *Parallelization*: An important property of GA is the distribution of the populations to multiple slaves (i.e. demes). In the parallel GA (PGA) structure, the GA's population can be spatially distributed among multiple demes[2] in order to improve the nature of population. The detailed descriptions of the available operators are presented in Section 4.2.1.

- *Distributed computing*: Another significant property of GAs is their capability of distribution across multiple processors to speed up the fitness function evaluations and convergence for offline applications. The number of processors in an ideal grid-computing cluster can be in the hundreds, i.e. extensible computing power can be availed as the size of the problem grows, thereby reducing the computation time almost linearly with the number of available processors. In this thesis, this extension of GA is referred to the distributed GA (DGA). In order to avoid confusions, the term *distributed* is used to referring to computation, while the term *parallel* will only refer to the GA's multi-deme population structure.

- *Hybridization*: In addition to distribution and parallelization, it was argued that a parallel GA (e.g., multi-deme) utilized with a distributed GA better mimics the nature of the population than a simple GA with a single population used in a serial GA [28]. Therefore, the application of parallel distributed GA (PDGA) is considered as the enhanced solution approach for a large-scale optimization problem.

The above advanced extensions of the GA methods are implemented into a web-based Java Enterprise Edition (JavaEE) platform called generic parallel genetic algorithms framework for optimizing intelligent transportation systems (GENOTRANS), which is integrated with another Java-based middleware platform called GridGain.

The following section presents the PGAs and the multi-deme control parameters, following by the description of the GENOTRANS, as the encoded GA and PGA library, and its

[2] Also referred to sub-populations or islands

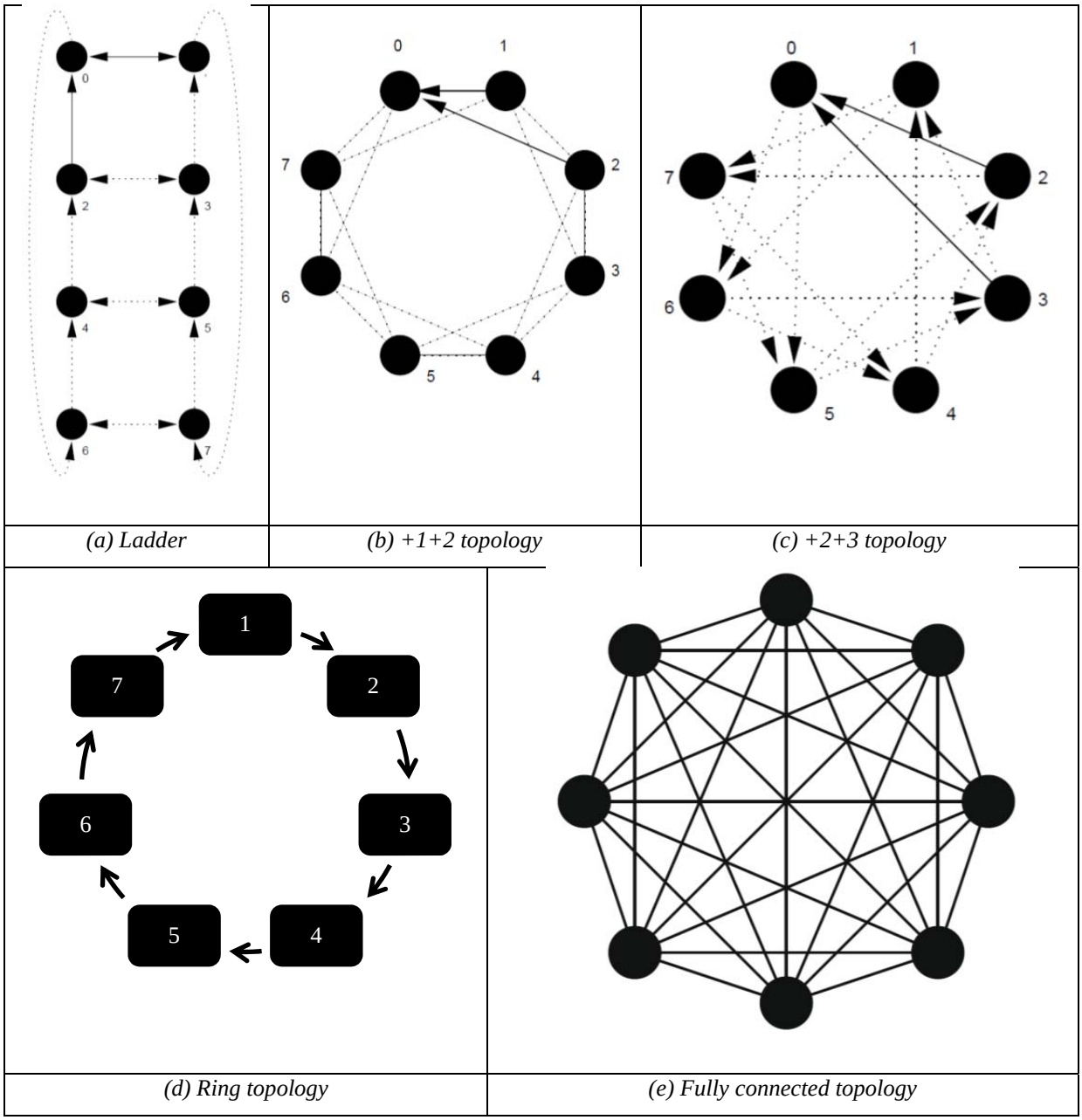
integrations with the GridGain to facilitate the distributed computing of the optimization problem.

4.2.1 Parallelization Structure of GA

As stated earlier, PGAs are shown to yield higher quality solutions, and possess more effective search and convergence properties. Depending on the parallel architecture adopted, PGAs often enhance the quality of the solution, require a smaller number of evaluations of the objective function, and have better chance of obtaining global optimum [81, 92]. In general, three types of PGAs exist: panmictic PGAs, diffusion-style PGAs and island model PGAs [91]. In a panmictic PGA, reproduction can be conducted between any two chromosomes in the population; whereas in the diffusion-style PGA, the chromosomes are spatially distributed (e.g. two-dimensional grid) and only neighbouring chromosomes can be recombined. In the island model PGA, semi-independent subpopulations, *demes*, evolve independently with periodic exchange of some chromosomes through a migration process [93]. The island model PGA exhibits even more correspondence to the evolution theory of species where thousands of subpopulations (or demes) exist and co-evolve in parallel in the same continuous geography.

The island model PGA requires a number of parameters as inputs, namely: the topology that defines the connectivity of demes as simple graph, the size and the number of subpopulations (i.e. demes), the migration policy that determines the selection and replacement schemes, the migration rate that controls the number of individuals to migrate, and the migration frequency (epoch interval) [81]. The importance of these parameters on the quality of the search and on the efficiency of the algorithms has long since been acknowledged [81, 91, 92, 93, 94]. In the following paragraphs, the different types of the above multi-deme control parameters are briefly reviewed.

Topology: The topology that defines the connectivity of the demes is simply a graph. Figure 4.1 shows some of the topologies identified in the literature [92, 93, 94].



(a) Ladder | *(b) +1+2 topology* | *(c) +2+3 topology*

(d) Ring topology | *(e) Fully connected topology*

Figure 4.1 Examples of islands topologies

According to the literature, the fully connected topology performs efficiently by producing a better quality solution and less execution time (converging faster) [92, 93, 94]. Therefore, the fully connected demes topology was the only topology implemented in the present thesis.

Migration Policy: The effect of migration is argued as one of the difficulties faced in the design of PGAs. There are some suggested migration policies, namely: (1) good migrants

replace bad individuals, (2) good migrants replace random individuals, (3) random migrants replace bad individuals, and (4) random migrants replace random individuals. According to the literature, some migration policies, i.e. choosing the migrants or replacements according to their corresponding fitness function, may cause the algorithm to converge at a significantly faster rate [81, 92, 93, 94]. Therefore, migrating the best individuals and replacing the least fit is the policy that most accelerates convergence which is chosen for the purpose of this study.

Other Control Parameters: There are a number of control parameters in the multi-deme structure of GAs, namely the size and the number of demes, the migration rate, and the migration frequency that are not very well presented in the literature. The questions of how big the deme size and how many demes are required resulted in mixed conclusions. However, it is commonly agreed that the smaller demes require more neighbours to succeed [81, 91, 92, 93, 94]. Therefore the size and the number of demes would be dependant of the case-study. The same rules apply for migration rate and frequency. Based on some experimental analysis, Cantu-Paz (1998) showed that the solution quality increases with higher migration rates [95], while some other studies found that the higher the migration rate, the less quality of the solution in the present problem [93, 94]. For the two case studies of this thesis, these control parameters were selected based on the results of the pilot experiments. Readers are referred to Kattan (2005) and Mohamed (2007) for the experimental designs and the impact of augmenting the GA operators with parallelization [81, 87].

4.2.2 GENOTRANS and GridGain: The PDGA Platform

As stated earlier, the simple and parallel GA methods, as well as their advanced features, were implemented into a platform called GENOTRANS, initially developed and updated at the University of Toronto ITS Center [93, 94]. GENOTRANS is a generic ‘Java’ Parallel Genetic Algorithm tool, in which the objective function is evaluated and constraints are satisfied through a simulation model (e.g. Paramics). The updated version of the GENOTRANS contains a full library of actual operators that can model the wide variety of GA options and operators. In addition, the aforementioned simple GA and PGA search techniques can be launched in GENOTRANS. Figure 4.2 and Figure 4.3 present the

snapshots of the simple and parallel GA operators implemented in the GENOTRANS library, respectively.

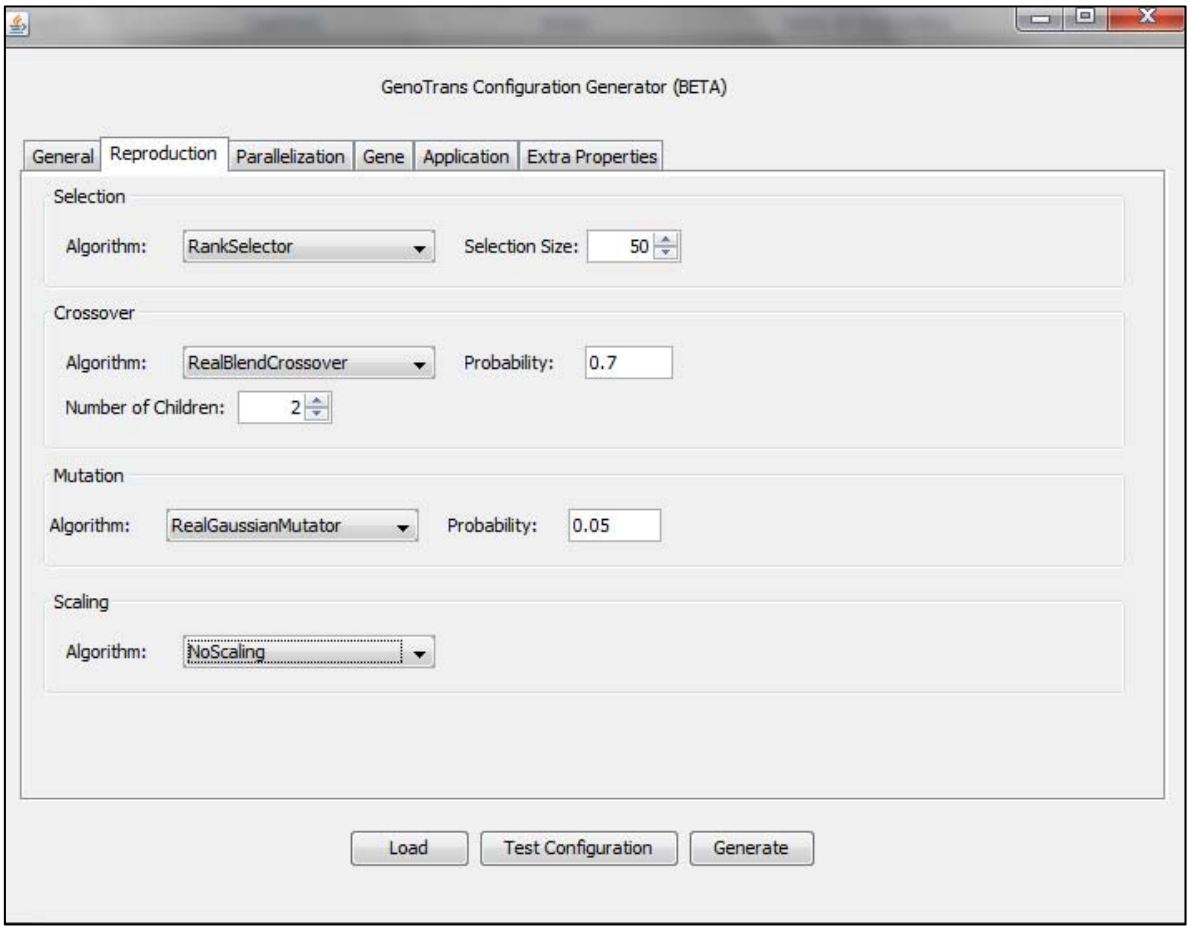


Figure 4.2 GENOTRANS configuration generator (Simple GA)

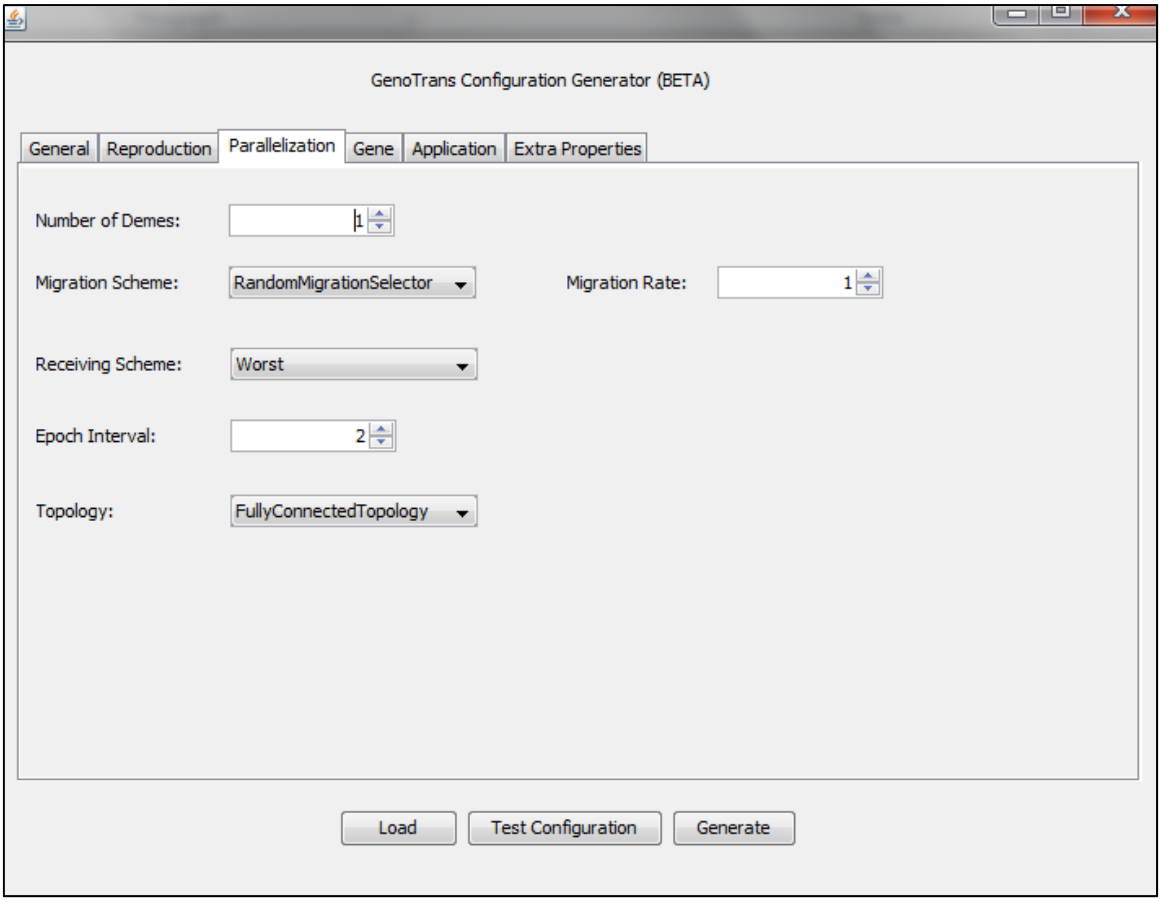


Figure 4.3 GENOTRANS configuration generator (Parallel GA)

The GENOTRANS is built on top of the generic grid-computing platform called GridGain. GridGain is an open-source Java-based platform for the distributed computations [9596, 97, 98]. The GridGain software consists of a master/slave model mediated by a job dispatcher process. A GridGain master process assigns a job to the nodes, which queues it for processing by a GridGain slave process. More specifically, computational grids or defines the method of splitting original compute task into multiple sub-tasks, executing these sub-tasks in a distributed manner on any managed infrastructure and aggregating results back to one final result.

GENOTRANS extends the GridGain master and slave processes by adding generic GA capability. A GENOTRANS slave process is essentially an objective function evaluator, while a GENOTRANS master process manages a given deme (for the simple GAs) or set of

demes (for PGAs). Figure 4.4 illustrates the schematic interaction flow between GENOTRANS and GridGain platforms.

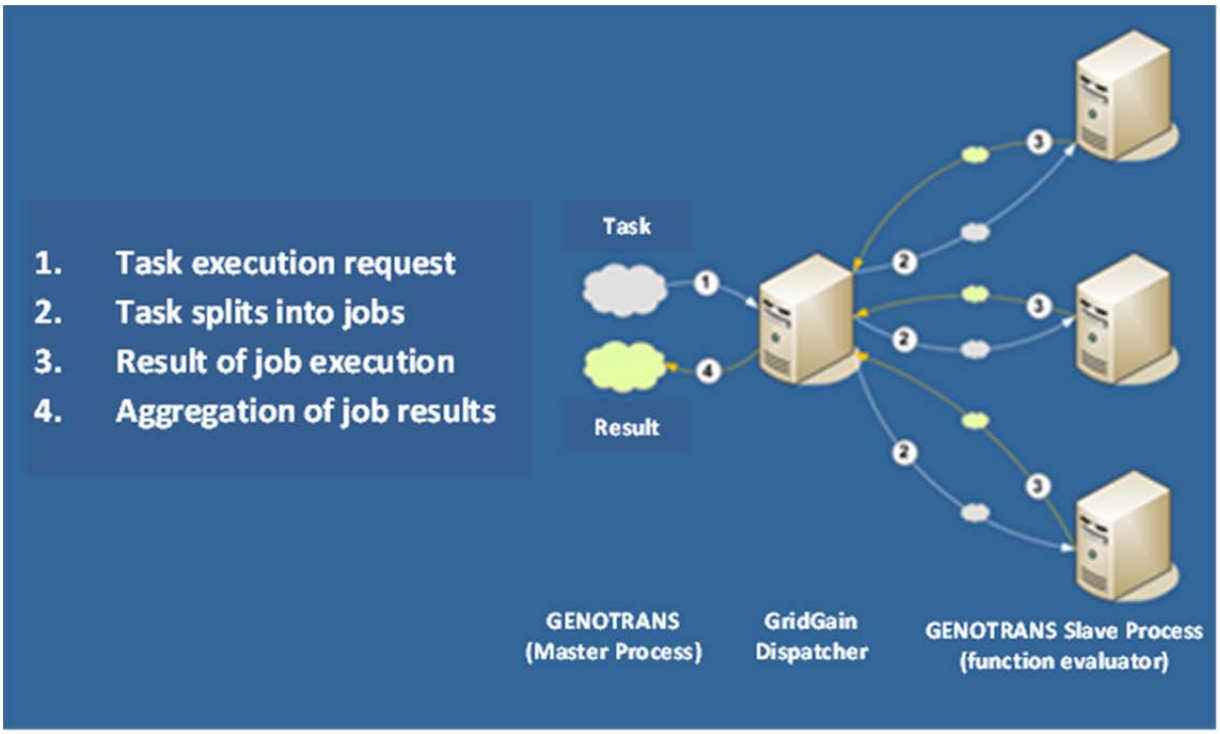


Figure 4.4 Interaction flow between GENOTRANS and GridGain (Inspired from [97])

As is apparent from Figure 4.4, in addition to the parallelized nature of GAs encoded in GENOTRANS, the integration of the two platforms will utilize the distributed evaluation of the objective function across multiple processors in a grid-computing cluster. This distribution would be efficient when dealing with computationally demanding problems. Therefore, a multi-deme distributed GA (i.e. PDGA) is designed to simultaneously calibrate the demand and supply parameters of the DTA model in a large-scale complex network. Figure 4.5 presents the general PDGA structure implemented on the GENOTRANS and GridGain.

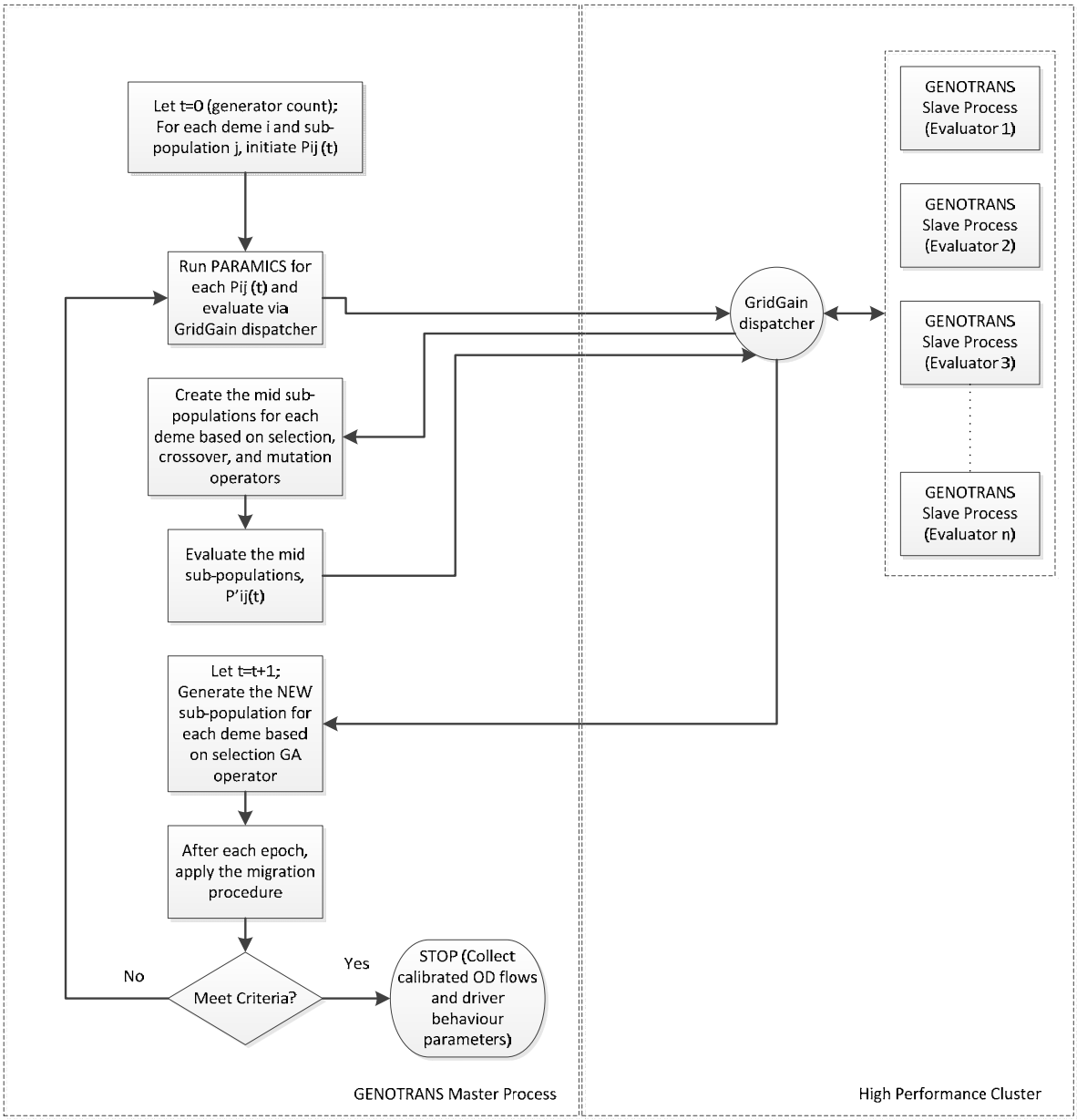


Figure 4.5 Flowchart of the PDGA implemented on the GENOTRANS and GridGain platforms

Upon selection of extended solution algorithms, the next section briefly describes the available computer cluster to efficiently conduct the calibration process for a large-complex transportation network.

4.3 Integration of High-Performance Computing Cluster with GA

This thesis utilizes a high-performance computer (HPC) cluster in the Department of Civil Engineering at the University of Toronto. The cluster has 64 processing nodes, 44 with 4 GB

of memory and 20 with 8 GB of memory, all with two processors, XEON 5150 2.66GHz dual core Woodcrest for a total of 4 processing cores per node (i.e. 265 processors), 36GB 15k rpm SAS hard disk, dual gig Ethernet, one public port and one dedicated cluster port. In addition to a master node, 20 nodes/slaves (i.e. 80 processors) were available to simultaneously distribute the computation across the available processors: the chromosomes are farmed out to multiple processors for evaluation in parallel (Figure 4.6). The cluster is managed through the Compute Cluster Job Manager user interface provided by Microsoft® Windows® Compute Cluster Server (2008). The cluster manager provides an integrated application platform for running, managing, maintaining, monitoring, and developing parallel computing applications [91]. It should be noted that during the course of this research, a number of CPUs in the HPC became unavailable (non-responsive), therefore the simulation models were run in other available CPUs with Paramics license.

Figure 4.6 High-performance computing facility at University of Toronto [91]

GridGain Version 3.6 is used to run the GA based on the GENOTRANS platform. Master/slave architecture is used to control the slave servers, as presented in Figure 4.7. A master node is responsible for sending requests to execute one GA evaluation (e.g., traffic simulation run) for every request. The requests are dispatched to any available node in round-robin fashion, regardless of which machine the application server is running on. Through the

GridGain Administrative Console, a weighting factor is assigned to each node to specify the load that can be assigned to this node, e.g., a weighting factor of 1 indicates that this node can be fully utilized during the execution process.

It should be noted that, in large-scale optimization problems that require considerable reading and writing of associated network and parameter files, conventional disk drive systems are not efficient at handling such massive information transfer across HPC nodes. Therefore, a redundant array of independent disks (RAID), which is a new and efficient technique of improving data availability and transferability using arrays of disks and various data-striping methodologies, is utilized.

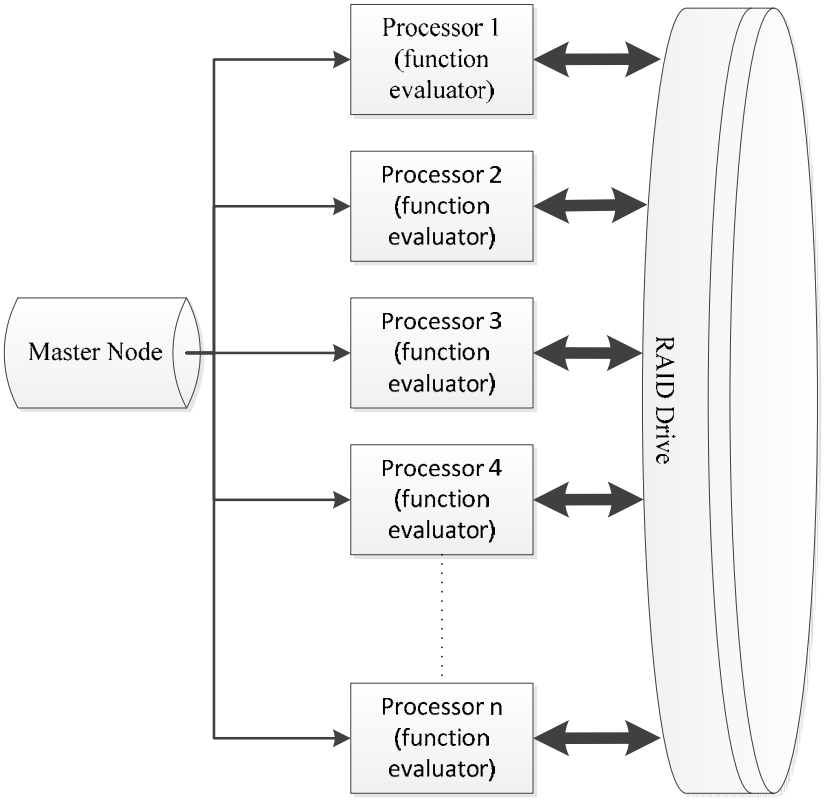


Figure 4.7 Master/slave cluster setup in GridGain

4.4 Summary

This chapter started with an overview of the GA and the various control parameters associated with the algorithm. In the next section, the selection mechanisms of the genetic operators were discussed. While it was argued that the simple GA (SGA) can effectively search the problem domain thoroughly, the convergence speed and the quality of the solutions of the SGA made it less desirable for ITS applications in large-scale networks. Therefore in this research, the performance of the SGA was enhanced by running the GA in multiple processors (i.e. evaluating the objective function in parallel CPUs) as well as distributing the GA's population among multiple demes. The former property of GA enhances the convergence speed while the latter one improves the quality of solution. These two extensions of SGA were referred to a distributed GA (DGA) and parallel GA (PGA), respectively. The aggregation of the two properties of GA resulted in parallel distributed GA (PDGA), which tries to simultaneously improve the convergence speed and quality of the final solutions. Upon selection of the relevant control parameters, the PDGA was implemented on the GENOTRANS/GridGain platform. Finally, for distributed computation of simulation models in Paramics, a high-performance computer (HPC) cluster and a number of other available CPUs at the University of Toronto ITS Lab were selected.

Chapter 5: Case Study I: Synthetic Network

So far, we have described the general simultaneous DTA calibration problem using multi-source traffic data. The problem has been formulated as a stochastic optimization problem. The previous chapter focused on describing the developed framework. Advanced genetic algorithms (GAs), namely distributed GA (DGA), parallel GA (PGA), and parallel distributed GA (PDGA), were selected as the solution methodologies for solving the optimization problem.

This chapter demonstrates the application and the performance of these algorithms for estimating the OD flows and calibrating the relevant driver behavior parameters in a simulation environment, using the applicable traffic data and the AVI sensor data. A small network with synthetic traffic data has been considered for this purpose. The assumed "true" observed sensors data is generated by means of a AimSun mesoscopic traffic simulator [7]. AimSun is assumed to perform as a proxy for real world. Noise is added to the true sensor count values to represent reality more closely. The Paramics microscopic DTA model has been calibrated using the sensor values generated by AimSun.

In summary, the objectives of this chapter are to:

- Demonstrate the feasibility of the proposed simultaneous calibration process in a small network with synthetic data;
- To evaluate the relative effectiveness of simultaneous demand-supply calibration compared with the iterative bi-level calibration approach;
- Identify the effect of augmenting the GA operator with parallelization and distrusted computing schemes;
- Evaluate the effect of adding "synthesized" AVI data into the calibration process and compare the calibration results with the base case where only link counts are available for calibration; and,

- Analyze the sensitivity of the calibration results to the weighting scheme of the objective function.

The remaining of this chapter is organized as follows: the next section describes the detailed design of synthetic network, following by the implementation of the GA. The calibration results are presented in the third section of this chapter, following by the sensitivity analysis of the weighting factors. Finally, Section 5.5 summarizes the outcomes of this chapter.

5.1 Experimental Design

This experiment was conducted on a small synthetic network using the Paramics micro-simulation package. This section explains the details of the experiment design, namely the description of the test network, calibration parameters, measurement of the loop detector sensors and AVI data, and measures of goodness-of-fit.

5.1.1 Description of the Test Network

The proposed methodology is implemented on a small synthetic network, consisting of 9 OD pairs (i.e. 3×3 OD matrix), 10 nodes and 10 links. Out of 10 nodes, 4 nodes are intersections and remaining 6 denote the origins and destinations of drivers in this network. In addition, four out of the nine OD pairs involved a route choice. The simulation period took place between from 7:30 a.m. to 9:00 a.m., with a warm-up period of 7:30 a.m. to 8:00 a.m. The remaining one-hour simulation was divided into four departure intervals of fifteen minutes. Figure 5.1 presents the topology of the small network.

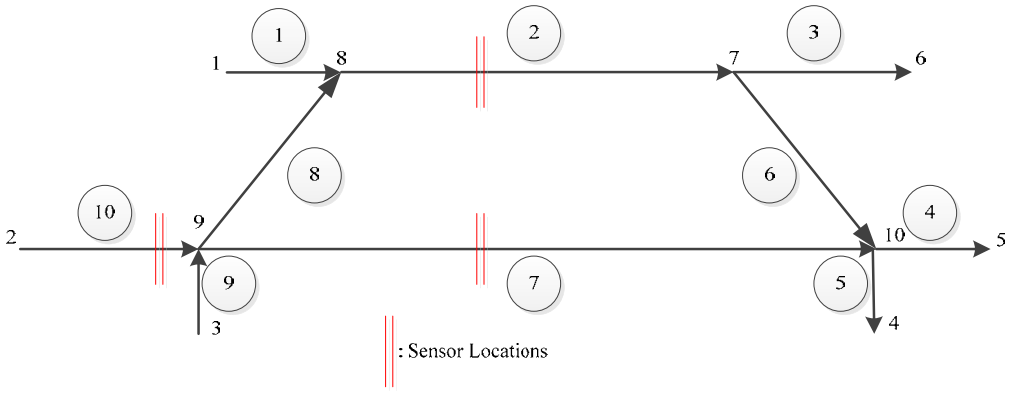


Figure 5.1 Synthetic network topology

5.1.2 Input Parameters for Calibration

As stated earlier, the initial population consisted of randomly perturbed ODs from historical OD flows (if available) and perturbed driver behavior parameters from Paramics' default values. This case study involved synthetic data, as true historical OD flows are not available for a synthesized network. Therefore, the initial OD flows for the first population (i.e., perturbed OD flows) were randomized from a set of user-defined synthetic time-dependent OD flows, which were considered as replacements for the true historical OD flows in the objective function (i.e., OD^h in Equation 1).

It should be noted that the synthetic a priori OD flows were randomly perturbed in a pre-specified range (-50% up to 50%) to obtain dynamic perturbed OD flows as the first starting values for calibration. This user-defined wide perturbation range ensures that the starting OD values do not replicate the synthetic time-dependent OD flows in the minimization problem. Paramics' default driver behavior parameters were randomly perturbed within the acceptable searching range (presented in Section 3.3) and used as the input to the Paramics simulator.

In summary, the total number of parameters for calibration on the demand side includes 9 OD pairs for 4 time intervals (i.e. 36). On the supply side, there are 5 parameters for calibration, namely mean headway and mean driver reaction time as the 2 driver behavior parameters in Paramics, and perturbation factor, feedback and familiarity as the 3 route choice model parameters. It should be noted that in this simple network, these supply parameters are set fixed for all the segments, during the one hour simulation. Therefore, the total number of parameter for calibration equals 36+5=41.

5.1.3 Synthesized Observed Traffic Data

The small network has 3 loop detectors that can provide 3 sets of aggregated counts for the each 15-minutes time interval (on links 2, 7, and 10). Therefore, 12 sensor measurements are available for the calibration process. In addition, there were 3 AVI sensors that can provide 2 sets of point-to-point travel time information between the center of links 10-2 and 10-7 (Figure 5.1) per time interval (i.e. 8 measurements in total). The travel time was then converted to speed values for model calibration. It should be noted that Paramics version 6 was able to provide the point-to-point travel times of the random vehicles that cross an AVI

sensor equipped segment in the network. The microscopic simulator assigns a unique ID to the vehicles that pass by as well as the time stamp associated with the detection instances. Therefore, travel time and average speed of individual vehicles on the road section between two consecutive AVI sensors can be obtained.

Similar to the OD flows, the true (observed) values for the traffic counts, turning movements, and AVI speed data are not available for a synthesized network and thus need to be created. Since the calibration is implemented in the Paramics micro-simulation environment, the traffic data should be generated in a different simulation environment than Paramics; otherwise, the observed traffic data would be highly dependent on the perturbed OD flows in Paramics. Therefore, the synthetic network was simulated with Aimsun software. Aimsun is a mesoscopic model that combines driver behavior characteristics from microscopic models with macroscopic traffic flow to represent the relation among speed, density, and flow. Therefore, the models and parameters would be fundamentally different from Paramics. Figure 5.2 presents the visual interaction between Paramics and Aimsun.

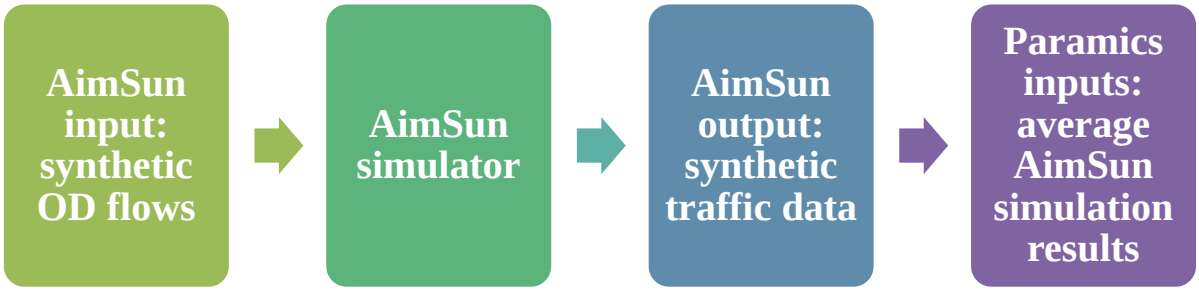


Figure 5.2 Interaction between Paramics and Aimsun

Using the same network and simulation run-time structure as Paramics, the perturbed OD flows were incorporated as the input in the Aimsun model, and the outputs of the simulation were the observed (synthetic) traffic data. As a result, the created output traffic data from Aimsun (e.g., observed counts, and AVI speed data) were utilized as the true observed data for the Paramics calibration purposes. It is noted that, due to the stochastic nature of Aimsun outputs (similar to Paramics), 5 runs of the Aimsun simulator were performed, and the traffic values were averaged for the calibration of Paramics. In addition, the traffic data from the

loop detectors should be fine-tuned, in order to capture the impact of noisy measurement. Therefore, the data obtained from loop detectors in Aimsun were adjusted by a user-defined symmetrical randomizer, ranging from -25% to +25%. These perturbed true traffic count data were used for calibration. The exact output values of Aimsun's AVI and turning movement traffic data were used for calibration. The number of parameters that needed to be calibrated included six OD pairs for four time intervals and five driver behavior parameters. Therefore, the total number of demand and supply parameters to be calibrated was 41.

As stated earlier, the total number of observed data is equal to 20, while the total number of unknown parameters is 41. Therefore, the calibration problem is underspecified with the degree of freedom of 41-20=21.

5.2 GA Implementation Details

5.2.1 Simple GA Control Parameters

The various control parameters in a GA are critical to the solution quality. The population size control parameter is one of the most important parameters that significantly affect the performance of GAs. If the population is too small and sparsely spread, the lack of genetic diversity may lead to quick convergence on local optima before better optima can be visited. However, excessively large populations cause the GA to act like a random search algorithm, and the search may flounder. Moreover, the number of generations has a tremendous effect on the computational performance and convergence rate of a GA.

After some initial trials for the synthetic network, the population size and the number of generations were selected as 80 and 30, respectively. As a summary, Table 5.1 presents the configuration of the simple GA control parameters for the synthetic network. Readers are referred to Section 4.1 for the detailed description of these parameters.

As for the stopping criteria, a combination of two criteria was selected. The algorithm terminates whenever one of the following criterion is first met:

- Reaching the pre-specified number of generations (e.g., 30), and

- No improvement in the value of fitness functions by more than 1% over a specified number of generations (e.g., last 5 generations).

If a termination criterion is satisfied, the GA is then terminated and outputs the best solutions from the last iteration. The final output of the GA is the estimated OD flows and calibrated driver behavior parameters for Paramics.

Table 5.1 Configuration of control parameters

Control Parameter	Selected Value
Selection method	Ranked-based selection mechanism
Cross-over	α-blend crossover (α=0.2)
Crossover rate	P_c= 90%
Mutation method	Self-adaptive Gaussian
Mutation rate	P_m= 5%
Population size	80
Number of generations	30
Number of simulation runs for each chromosome evaluation in Paramics	5

5.2.2 Parallel GA Control Parameters

As stated earlier is Section 4.2.1, there are a number of control parameters in the multi-deme structure of GAs that play an important role in the quality of the final solutions. Based on the literature and pilot experiments, the multi-deme structure is selected for the synthetic network [81, 87]. Table 5.2 summarizes the PGA design elements.

Table 5.2 Parallel GA design elements

Control Parameter	Selected Value
Island topology	Fully connected topology
Migration policy	Good migrants replace bad individuals
Migration rate	10%
Migration frequency (epoch interval)	5
Number of demes	4
Deme size	20

5.3 Calibration Results

With the synthetic data, the methodology presented earlier was applied for the simultaneous calibration of dynamic OD flows and driver behavior parameters. For comparison purposes, separate experiments were performed, in terms of the following comparisons of criteria:

- *Study data:* loop detector counts compared to multi-source traffic data (i.e. loop detector counts, and AVI speed data) based on the simultaneous calibration methodology.
- *Calibration methodology:* simultaneous calibration compared to bi-level iterative calibration of demand and supply parameters.
- *Optimization engines:* non-distributed computing (i.e. SGA) compared to a distributed computing technique (i.e. DGA).
- *Effect of parallelization:* comparing SGA and DGA to the parallel distributed GA (PDGA).

5.3.1 Calibration Results based on Different Traffic Data

Table 5.3 and Table 5.4 present the calibration result statistics based on GEH and NRMSE measures of effectiveness for counts and speed values, using the HPC, respectively. For visual comparison, the MRMSE and GEH values for calibration are summarized in Figure 5.3 and Figure 5.4. In addition, the NRMSE and GEH values of count and speed

values for each 15-minute interval of the analysis period (8:00 a.m. to 9:00 a.m.) are presented in Figure 5.5, and Figure 5.6. The difference between the observed and simulated counts (obtained by loading the seed synthetic OD flows to Paramics without any calibration iteration) are reported as the before calibration case in this table. The base calibration case refers to the calibration process using loop detector count. Furthermore, the base calibration case + AVI data corresponds to the results obtained from the multi-source traffic data. It should be noted that the following weighting factors were considered for the synthetic network (Equation 1):

- Base case: α_{count}=0.6 for counts and α_{OD}=0.4 for OD flows
- Base case + AVI speed data: α_{count}=0.3 for counts, and α_{OD}=0.2 for OD flows, and α_{count}=0.5 for speed values

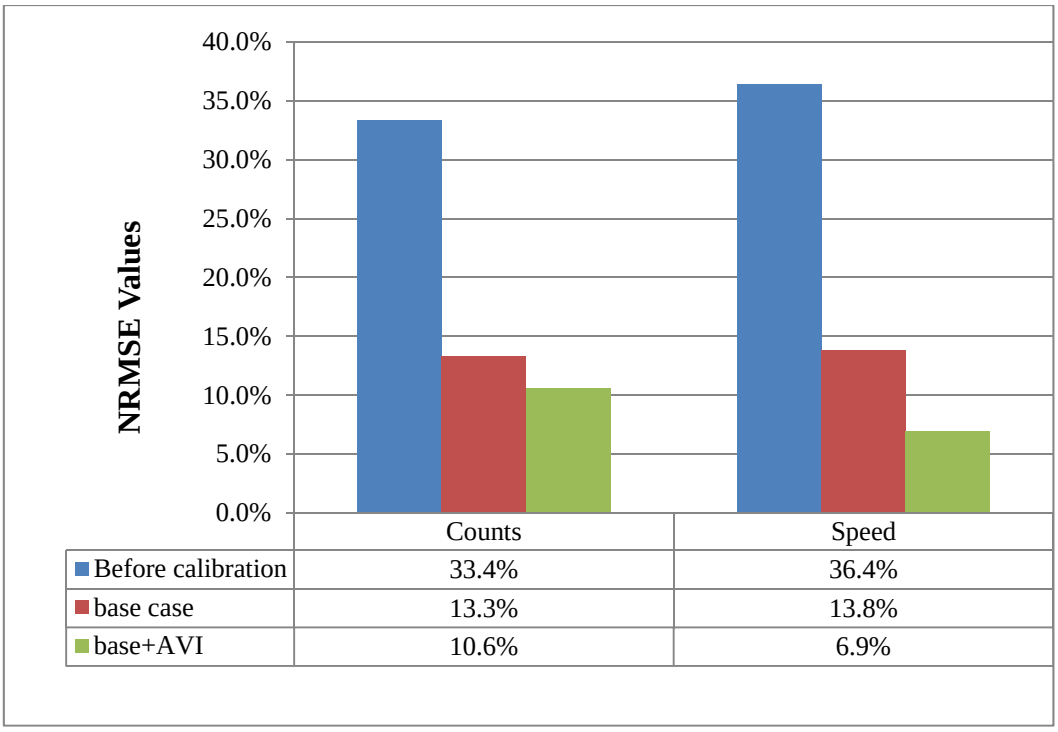
Readers are referred to Section 5.4 for the detailed sensitivity analysis that led to the estimation of the above weighting factors. In addition, for comparison purposes between different scenarios, the same weighting factors were considered for evaluation of the SGA versus PGA.

Table 5.3 Calibration result statistics based on counts

Measures of Effectiveness	Before Calibration	Base Calibration Case	Base Calibration Case+AVI Data	% Change (Base Case vs. Before Calibration)	% Change (Base Case vs. Base Case+AVI)
GEH	12.87	4.55	3.44	64.7%	24.3%
NRMSE	33.4%	13.3%	10.6%	60.1%	20.5%

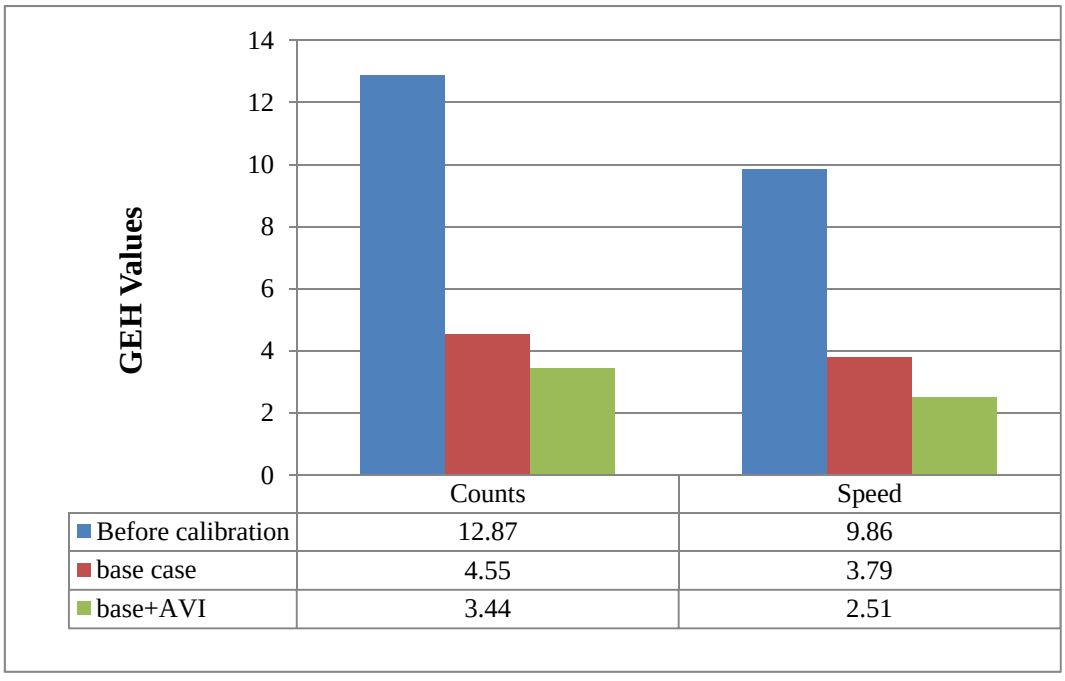
Table 5.4 Calibration result statistics based on speed values

Measures of Effectiveness	Before Calibration	Base Calibration Case	Base Calibration Case+AVI Data	% Change (Base Case vs. Before Calibration)	% Change (Base Case vs. Base Case+AVI)
GEH	9.86	3.79	2.51	61.5%	33.9%
NRMSE	36.4%	13.8%	6.9%	62.0%	49.8%



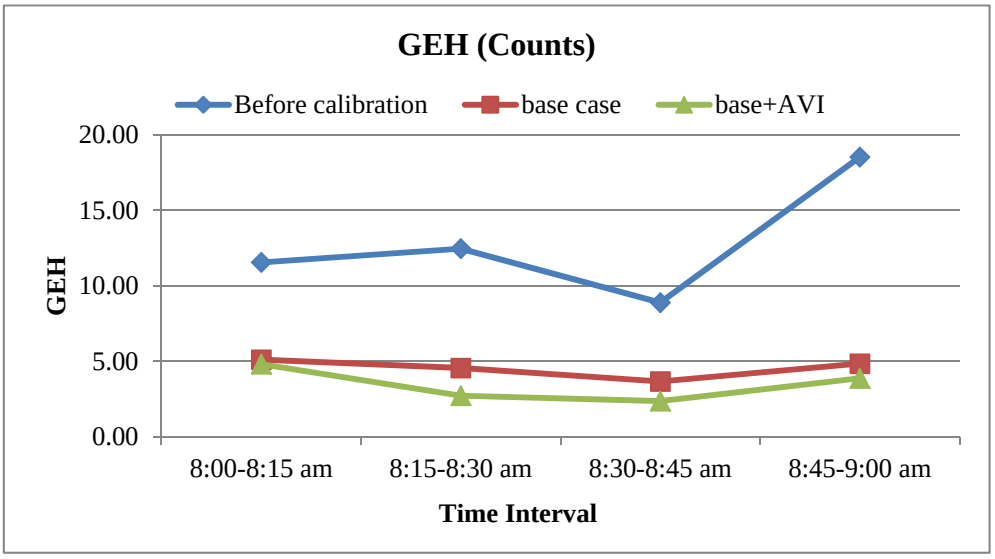
	Counts	Speed
Before calibration	33.4%	36.4%
base case	13.3%	13.8%
base+AVI	10.6%	6.9%

Figure 5.3 Comparison of different scenarios based on NRMSE values

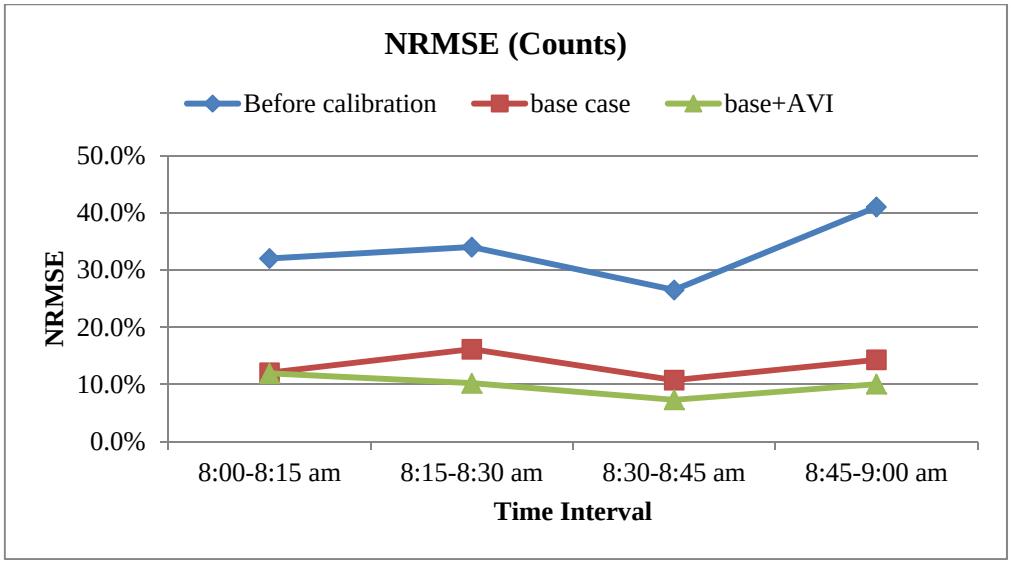


	Counts	Speed
Before calibration	12.87	9.86
base case	4.55	3.79
base+AVI	3.44	2.51

Figure 5.4 Comparison of different scenarios based on GEH values

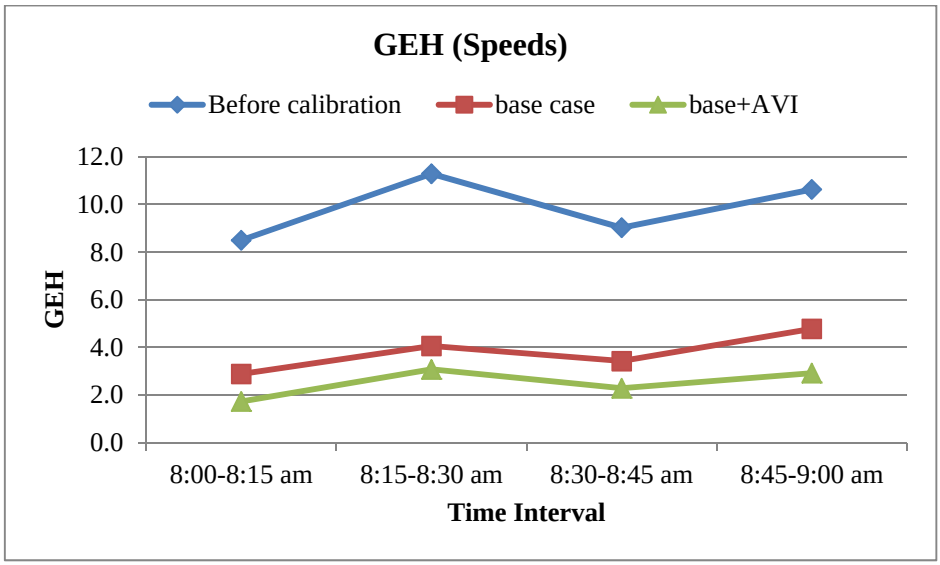


a) GEH

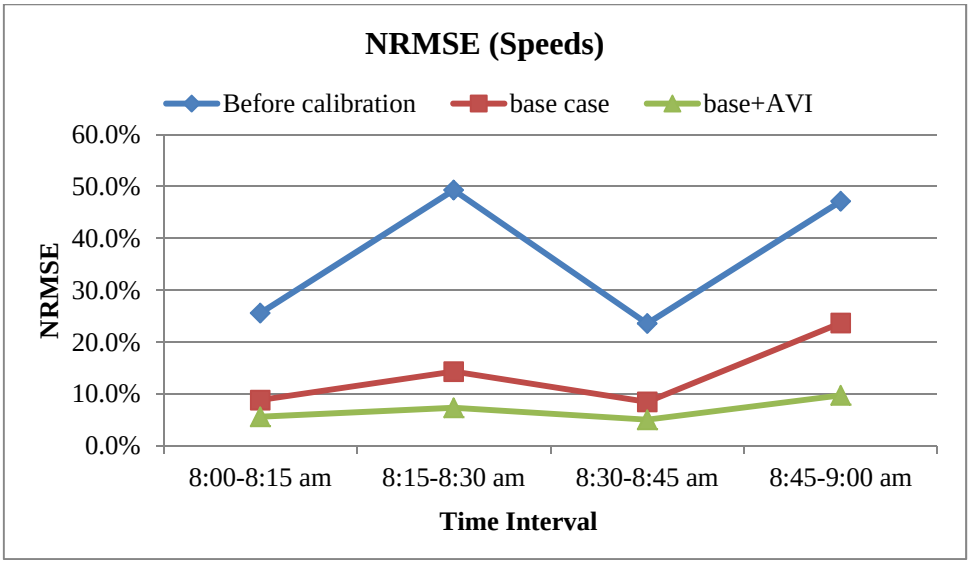


b) NRMSE

Figure 5.5 Comparison between measures of effectiveness of different cases based on counts (15-minutes intervals)



a) GEH



b) NRMSE

Figure 5.6 Comparison between measures of effectiveness of different cases based on speed values (15-minutes intervals)

As is apparent from Table 5.3, the reported percentage changes between the before calibration case and the base calibration case reveals that the simultaneous calibration process using GA performed as expected and had the capability of minimizing the discrepancy between the observed and simulated traffic conditions (i.e., OD flows and loop detector counts). According to Figure 5.5, the percentage change in terms of GEH ranged between 55.8% and 73.8% for various time periods, with an average improvement of 64.7%. The average GEH for the base case was 4.55, which was less that the threshold of 5.0 recommended by the U.S. Federal Highway Administration (FHWA) guideline [99]. The improvement percentage in the objective function using the other measure of effectiveness (i.e., NRMSE) was consistent with the GEH percentage changes, ranging from 52.5 to 65.1%, with an average of 60.1%.

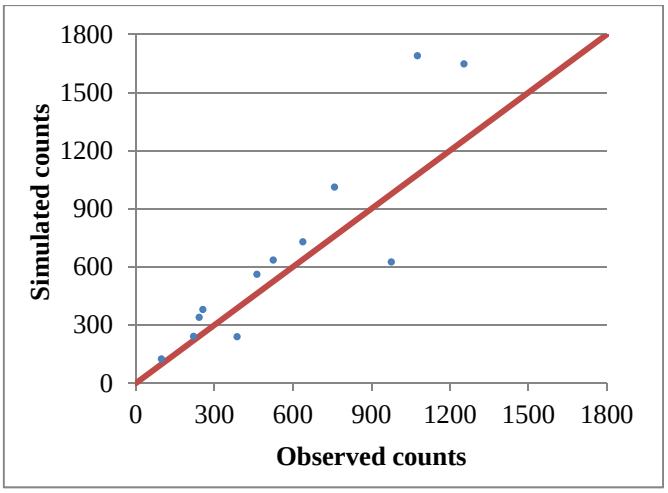
Following the same trend as the count measures of effectiveness, the results of the speed values from different cases are found to be consistent with the count measures of effectiveness. According to the Table 5.4 and Figure 5.6, the following conclusions can be made:

- The percentage change of speed GEH ranged between 55.0% and 66.0% for various time periods, with an average improvement of 61.5%.

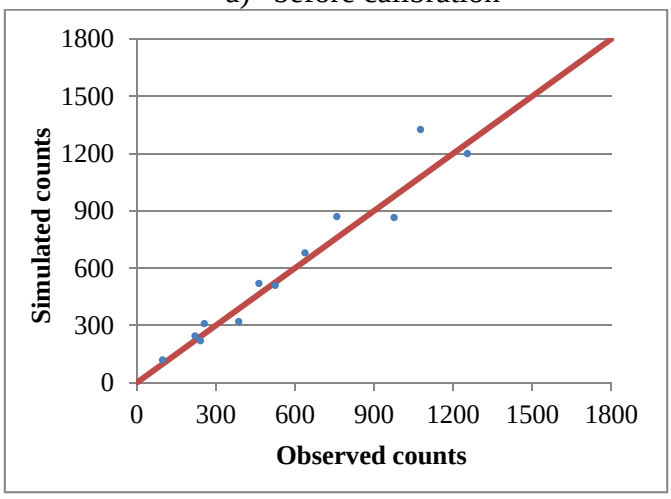
- The improvement percentage in the objective function using the speed NRMSE is consistent with the speed GEH percentage changes, ranging from 49.8% to 71.0%, with an average of 62.0%.

In addition to conducting a before/after calibration study using traditional count data, this research also incorporated AVI traffic data into the calibration process. The percent change between the calibration results obtained from the base calibration case and the combination of AVI and count data revealed that the AVI data significantly improved the calibration accuracy by 24.3% and 20.5% in terms of count GEH and RMSN values, respectively (Table 5.3 and Figure 5.5). This improvement percentage in terms of speed GEH and RMSN values are found to be 33.9% and 49.8%, respectively.

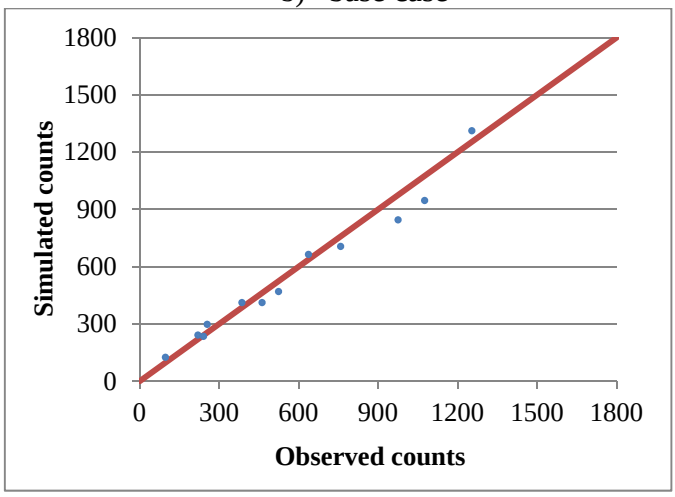
Figure 5.7a, Figure 5.7b, and Figure 5.7c present the plots of the observed counts and simulated entities for the before calibration, base calibration, and multi-source (Base case+AVI) scenarios, respectively. In these plots, a 45-degree line indicates the positions of the observed and simulated loop detector counts. A visual comparison between Figure 5.7a and Figure 5.7b reveals that the calibration process significantly minimized the discrepancy between the observed and simulated traffic counts. In addition, the incorporation of AVI data and sensor counts improved the calibration accuracy compared to the base calibration case (as shown in Table 5.3). As expected, the AVI data improved the calibration accuracy, in terms of travel time or speed data. Furthermore, as demonstrated in Figure 5.7c, that the AVI data also improved the calibration accuracy in terms of sensor count error. A possible reason is that the number of sensors in the network was low; therefore, the AVI data generally improved the calibration performance in terms of counts and speed data. Speed information provides a better indication of the quality of traffic flow. Speed information, in contrast to flow measurements, can clearly distinguish between congested or uncongested conditions. Thus, the incorporation of AVI readings improved the solution quality by decreasing the number of local point solutions.



a) before calibration



b) base case



c) multi-source data: base case + AVI data

Figure 5.7 Comparison between simulated and observed counts

5.3.2 Comparing Calibration Approaches

As stated earlier, several studies have been conducted to jointly calibrate the input parameters of the DTA models in a sequential fashion. This section compares the performance of the proposed simultaneous and the traditional sequential bi-level calibration methodologies. The previous section suggests that the incorporation of the AVI data can improve the calibration accuracy by minimizing the objective function. Therefore, the performance of both approaches was compared based on the same observation data (count and AVI data) in the synthetic network. The iterative bi-level approach consists of 2 optimization steps: the upper level estimates the OD flows while fixing driver behavior parameters; that was followed by the estimation of the driver behavior parameters in the lower level assuming fixed OD flows from the upper level. GA is used as the solution algorithm in both levels to estimate the OD flows and calibrate the driver behavior parameters. Readers are referred to the literature for the detailed bi-level calibration methodology [60, 61, 62, 63]. Figure 5.8 and Figure 5.9 present results of the NRMSE and GEH values of count data and speed values for both the simultaneous and bi-level calibration approaches during the 1-hour simulation period, respectively.

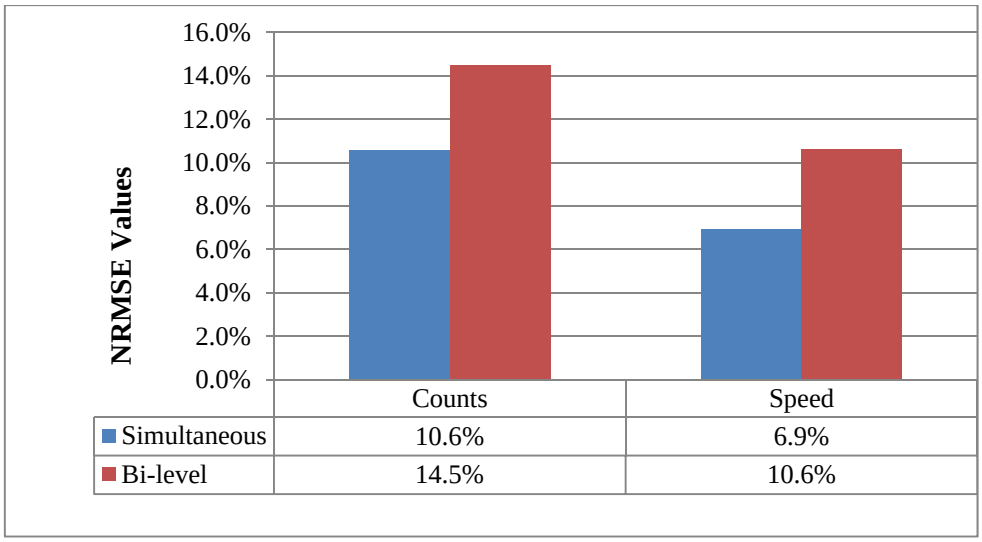


Figure 5.8 Visual comparison between simultaneous and sequential approaches based on NRMSE

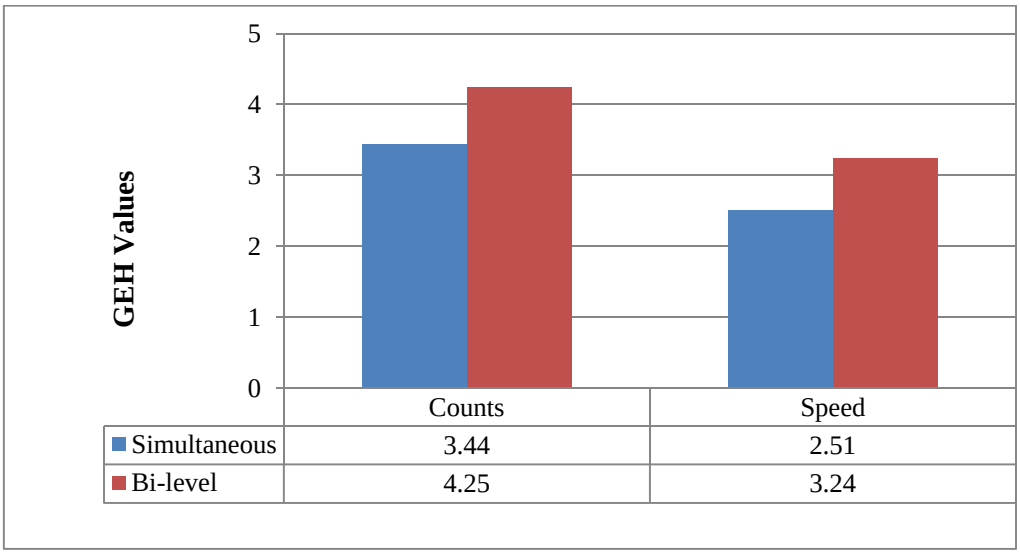


Figure 5.9 Visual comparison between simultaneous and sequential approaches based on GEH

Figure 5.8 and Figure 5.9 indicate that the simultaneous calibration methodology outperforms the sequential bi-level approach by 32.0% and 19.0%, as expressed in NRMSE and GEH for count data, and 34.6% and 22.7% in NRMSE and GEH for speed data, respectively. In other words, the simultaneous calibration of demand and supply parameters had considerably better accuracy as compared to the sequential approach. The results of this experiment are in line with the earlier findings since the complex and nonlinear interactions between demand and supply parameters could not tackled.

5.3.3 Comparing the Optimization Engines: Simple GA (SGA) vs. Distributed GA (DGA)

As stated earlier, this research utilized an HPC to run the GA in parallel computer processing engines (i.e. DGA). To obtain the final calibrated OD flows and driver behavior parameters using the count and AVI traffic data, the synthetic network was simulated a total of 12,000 times, which was the product of the number of generations (30), the number of chromosomes in each generation (80), and the number of DTA iterations in Paramics (5). These configuration control parameters are presented in Table 5.1. Based on the average 30 seconds run time for each hour simulation in a very simple network, this took 100 hours to complete. As a result, it is clear that running the GA in one CPU engine is not practical, even for a synthetic network. Therefore, there have been numerous efforts to make GAs faster and one of the promising techniques is to the use of distributed implementations of GAs. The effect of distributing the GA population to multiple slaves on the performance of the GA is typically examined, in terms of the following criteria [28, 82]:

- Elapsed time, which focuses on the master processor, the number of available processors, and the communication time to create the files necessary for each processor;
- Speed-up, which is the ratio of the execution time of the single processor GA to the elapsed time of the distributed GA; and,
- Efficiency, which is defined as the speed-up divided by the number of processors and represents the utilization of processors.

The analysis of the elapsed time focuses on the master processor and the number of available processors. In a typical GA generation, the master sends a fraction (or all) of the population to each of the slave processors, using communication time T_C. This communication time is exhausted in creating the files/directories necessary for each processor in the available slave list. Although the master consumes sometime in the selection, crossover and mutation processes, this time is typically ignored compared to the communication and execution times. Due to the unprecedented sheer size of the input/output files in this application, the master

remains idle and waits for the results from the available processors. Next, each slave (processor) evaluates a fraction of the population in time $\frac{nT_x}{P}$, where T_x is the execution time of one individual, n is the population size, and P is the number of available processors. Therefore the elapsed time for one generation is given by Equation (15).

$$T_E = PT_C + \frac{nT_x}{P} \tag{15}$$

An important concern when implementing large-scale problems is that the frequent communications between master and slaves may offset the gain in computation time. Therefore, the speedup (S) of the master-slave distributed GA is another measure of the effectiveness of a DGA. The speedup is the ratio of the execution time of the SGA to the elapsed time of the DGA as shown in Equation (16). The greater the ratio of T_x to T_C, the more linear will be the speedup.

$$S = \frac{nT}{PT_C + \frac{nT_x}{P}} \tag{16}$$

Although using more slaves reduces the computation time significantly, the communication time increases. Therefore, a third measure of the effectiveness of the DGA compared to an SGA is the efficiency (E). The efficiency is defined as the speedup S divided by the number of processors P (see Equation 17). The efficiency represents the utilization of processors. Ideally, E would be constantly one and S would be equal to the number of processors used (i.e. linear speedup). However, in reality, the cost of communications prevents this ideal case from happening; therefore, the efficiency is chosen as a measure of the deviation from the ideal case.

$$E = \frac{S}{P} \tag{17}$$

By examining the evolution of the fitness function with each GA generation, it was concluded that the DGA outperformed SGA, in terms of the convergence speed (i.e. CPU time). However, it should be noted that, as the number of processors increases, the elapsed time decreases, speed-up increases, and efficiency decreases. For the synthetic case study network, the GA was tested in the HPC and available Paramics licensed processors in the

University of Toronto ITS Lab. It is noted that out of 80 processors in the HPC, 22 processing engines were made available for the purpose of this research. The calibration process took approximately 9 hours to complete, compared to the 100 hours running the process on a single CPU machine. In other words, the DGA was approximately 10 times faster than the SGA. Considering the limited number of available processors, the total efficiency of the cluster was approximately 49%. Upon the full availability of the cluster processors (i.e. 80), it is expected that the efficiency of the Cluster significantly increases to 90% [91][3]. Therefore, in an ideal situation where all the processing nodes can be assigned to the problem, the DGA can be more than 70 times faster than SGA.

It should be noted that running GA in multiple processors only reduces the computation time, while the quality of the solution would be the same as the SGA (i.e. running GA in one processor). An improvement in the quality of solution was observed by parallelization structure of GA detailed in the next section. The final simulated dynamic OD flows, counts, speed data, and driver behavior parameters were obtained after running 27 generations. Table 5.5 presents the calibrated driver behavior parameters, based on the incorporation of count and speed data into the calibration process.

Table 5.5 Calibrated driver behavior parameters

Parameter Types	Description	Optimal Value
Vehicle-driving behavior	Mean headway	1.15 sec
	Mean reaction time	0.89 sec
Route choice model parameters	Feedback	210 sec
	Familiarity	82%
	Perturbation	11.2%

[3] Abd, H. M. A. E. H. (2010). *Optimization of Multimodal Evacuation of Large-scale Transportation Networks*, Doctoral dissertation, University of Toronto

5.3.4 Comparing the Simple GA (SGA) and Distributed GA (DGA) to the Parallel Distributed GA (PDGA)

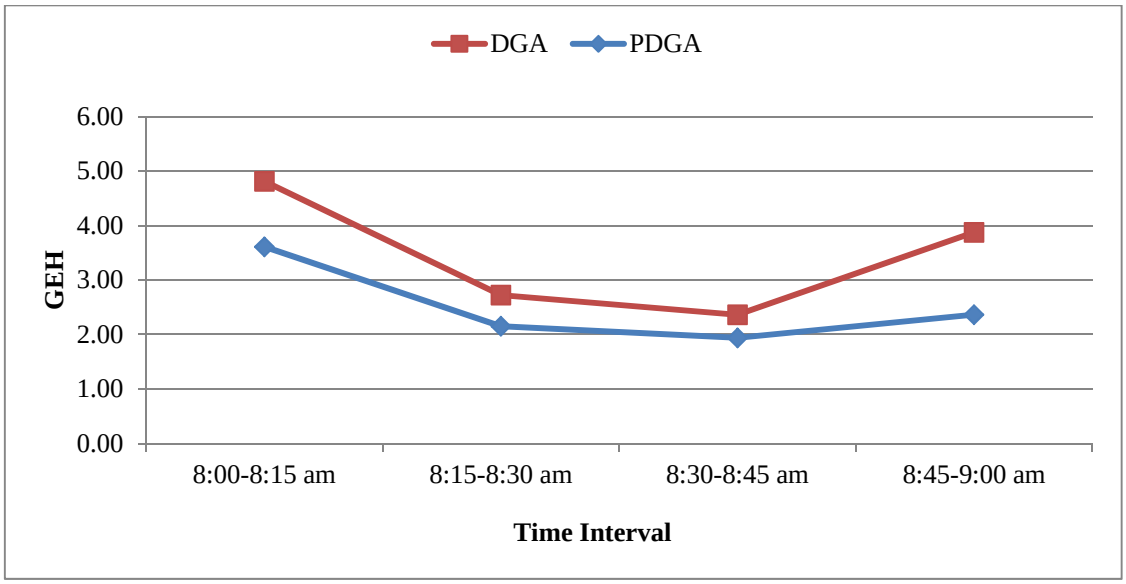
Parallel computing distributes the computation across multiple processors simultaneously, in which chromosomes are farmed out to multiple processors for evaluation. As stated in Section 5.3.3, the DGA improved the convergence speed of the solution algorithm comparing to SGA. On the other hand, the parallel GA (i.e. multi-deme) better mimics the nature of the population than an SGA with a single population and improves the quality of the solution. Therefore in this section, the performance of the SGA and DGA is compared with the PDGA, in terms of the convergence speed and calibration results obtained from the multi-source traffic data. It should be noted that the weighting factors, α_i, were estimated from the sensitivity analysis described in Section 5.4.

Following the methodology set forth in Section 4.2 and the parallel GA control parameters in Table 5.2, the final simulated OD flows, counts, AVI data, and driver behavior parameters were obtained. Table 5.6 presents the calibration result statistics for counts and speed values. In addition, the calibration result statistics for DGA were adopted from Table 5.3 and Table 5.4 and incorporated in Table 5.6 for the direct comparison between DGA and PDGA.

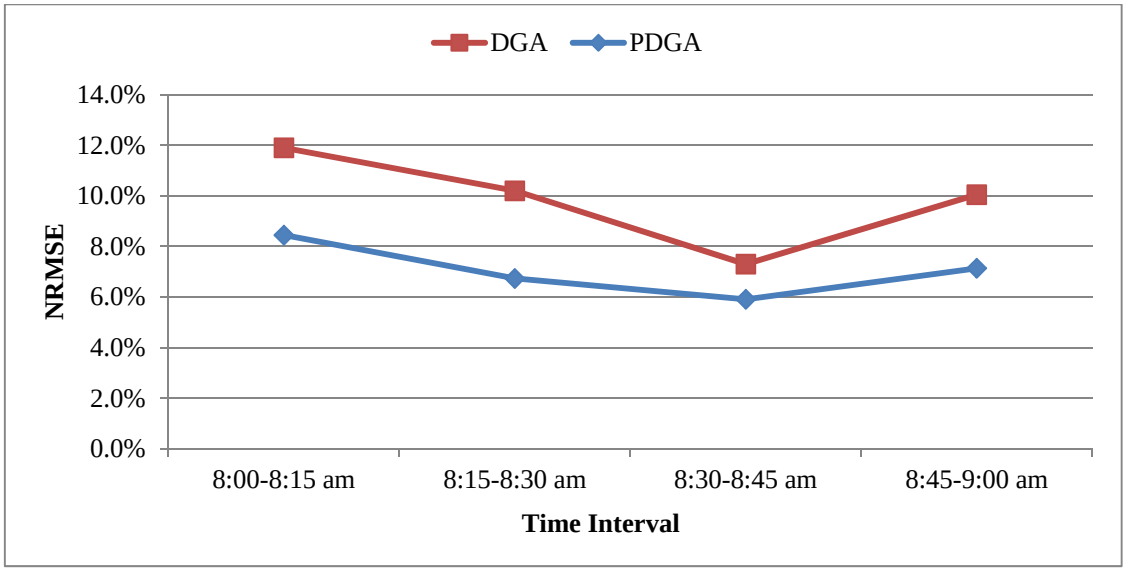
For the visual comparison, the count and speed NRMSE and GEH values for each 15-minute interval of the analysis period (8:00 a.m. to 9:00 a.m.) are presented in Figure 5.10 and Figure 5.11, respectively.

Table 5.6 Comparison between DGA and PDGA calibration result statistics based on counts and speed values

Category	Measures of effectiveness	DGA	PDGA	Improvement percentage
Count	GEH	3.44	2.52	25.8%
	NRMSE	10.6%	7.5%	27.8%
Speed	GEH	2.51	1.73	30.2%
	NRMSE	6.9%	4.9%	28.1%

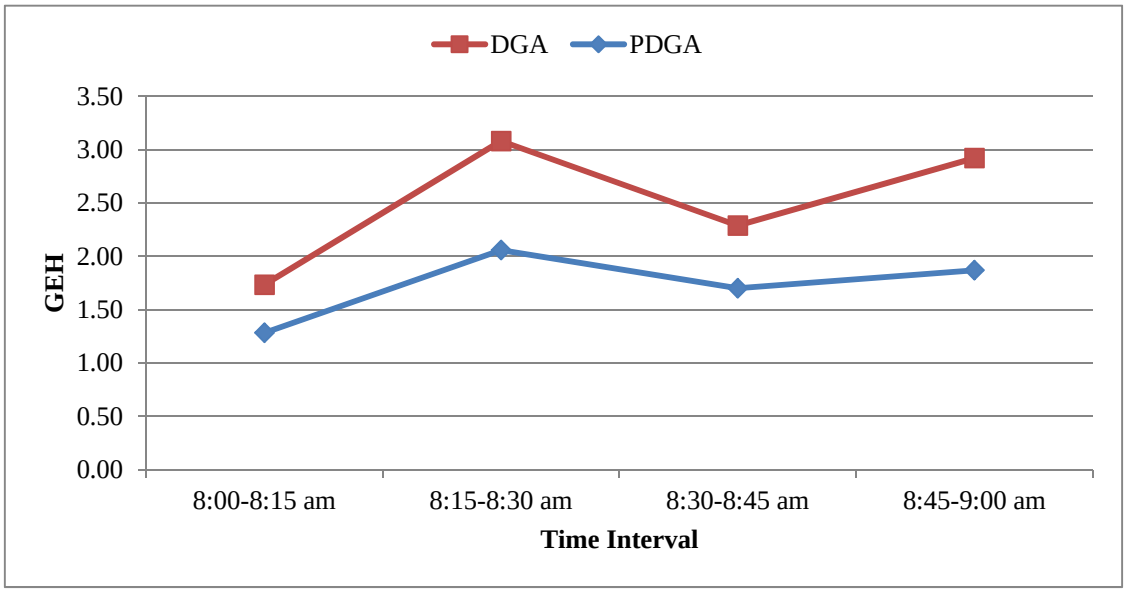


a) GEH

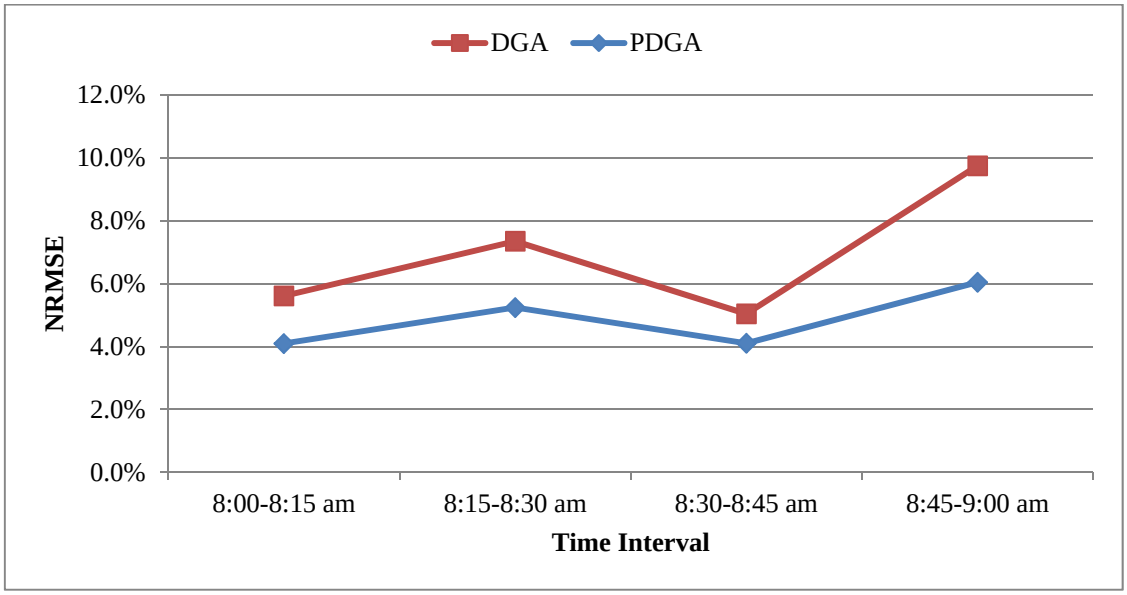


b) NRMSE

Figure 5.10 Comparison between measures of effectiveness of DGA and PDGA based on counts



a) GEH



b) NRMSE

Figure 5.11 Comparison between measures of effectiveness of DGA and PDGA based on speed values

As is apparent from Table 5.6, the reported percentage change between the calibration results based on the DGA and PDGA reveals that the parallel GA leads to further reduction in fitness function results, as compared to the single GA running in the HPC. According to Figure 5.10, the improvement in terms of GEH values varies depending on the simulation

periods ranging from 18.1% and 39.2%, as compared to the single population GA with an average improvement of 25.8%. In addition, the improvement percentage in the objective function using the count NRMSE is consistent with the GEH percentage changes, ranging from 19.1 to 34.0%, with an average of 27.8%.

Following the same trend as the count performance measures, Table 5.6 and Figure 5.11 shows that parallelization of GA result in an overall reduction of speed performance measures as compared to the simple GA runs in the HPC. The percentage change of speed GEH ranged between 25.7% and 36% for various time periods, with an average improvement of 30.2%. In addition, the improvement percentage in the objective function using the speed NRMSE is consistent with the speed GEH percentage changes, ranging from 18.6% to 38.0%, with an average of 28.1%. It should be noted that the above reduction in the fitness functions from DGA to PDGA were found to be statistically significant based on the 95% confidence interval. Figure 5.12 presents the plots of the observed counts and simulated entities for the PDGA and DGA calibration scenarios, using multi-source traffic data. A visual comparison of the two figures reveals that the calibration process based on parallelization of the GA structure significantly minimized the discrepancy between the observed and simulated traffic counts, comparing to the simple GA running in the distributed processors. In addition, Figure 5.13 presents the evolution of the fitness function with each GA generations.

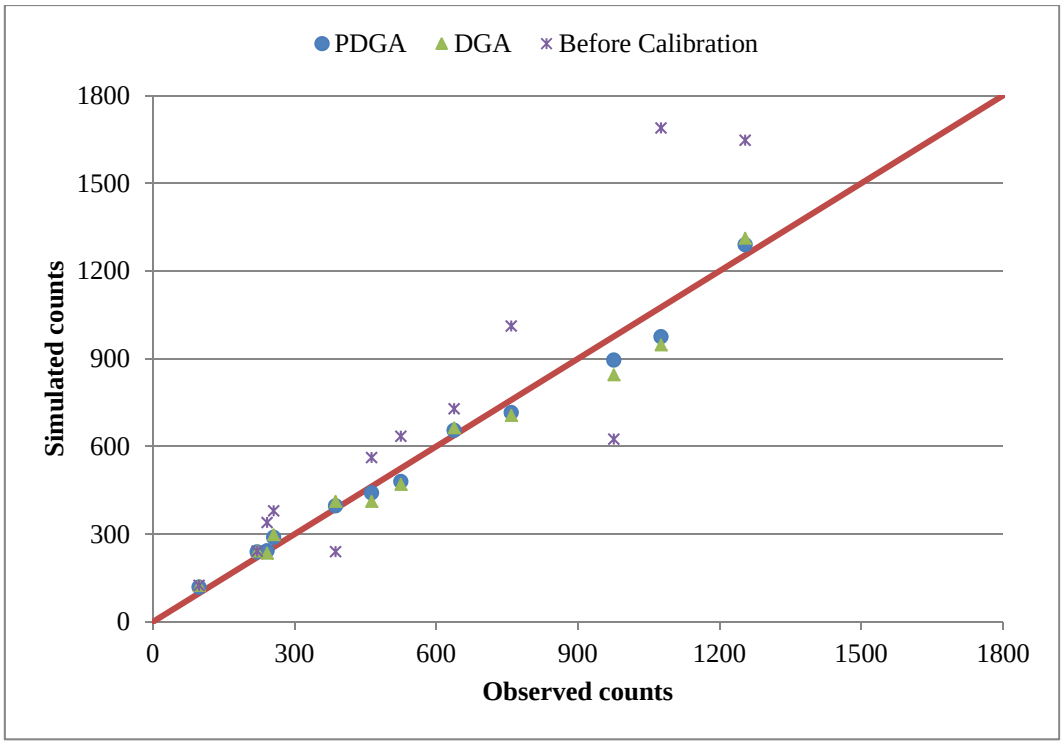


Figure 5.12 Comparison between simulated and observed counts based on NRMSE (DGA vs. PDGA)

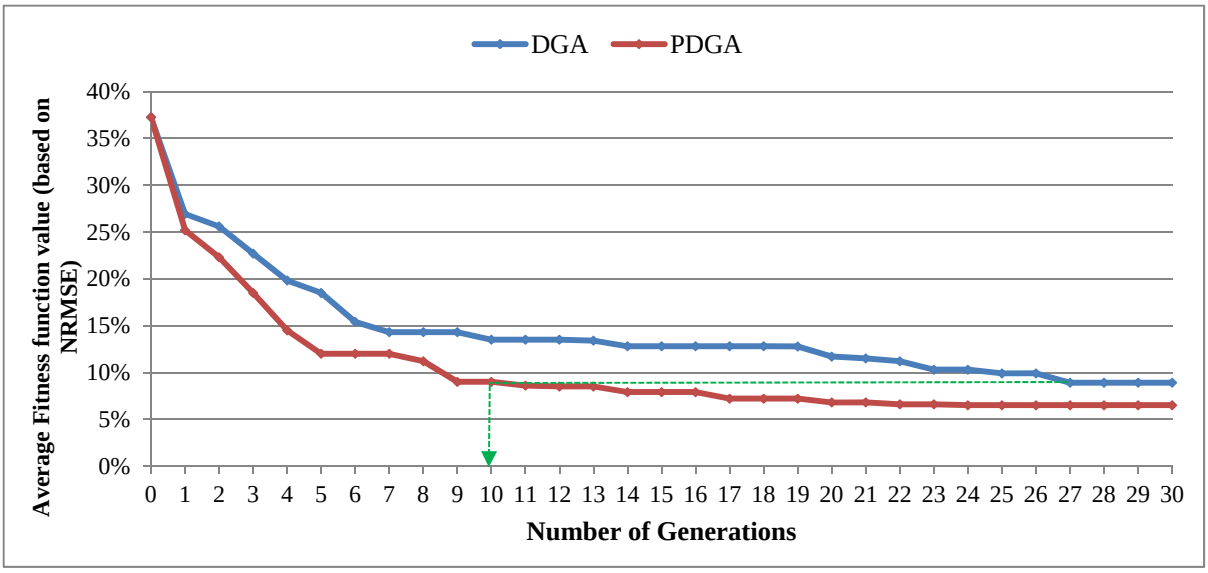


Figure 5.13 Effect of parallelization and distributed computing on GA convergence and quality of solutions (PDGA vs. DGA/SGA)

In summary, the results of the experiments demonstrated the advantages of the PDGA techniques over the simple GA distributed among parallel processors (i.e. DGA) with regards to achieving a higher quality solution for off-line applications.

By examining the evolution of the fitness function with each GA generations in Figure 5.13, the following observations can be made:

- *Termination point and quality of solution:* As is apparent from Figure 5.13, PDGA outperforms the DGA in terms of fitness function value and convergence speed (i.e. number of generations required to reach a certain fitness function value). First, the PDGA results in a better fitness function value at the termination point of the GA (i.e. generation 24), with the fitness function of 6.5% vs. 8.9% in terms of NRMSE values (i.e. 26.9% reduction). This means that for the same number of generations, it is found that the PDGA provides higher quality solutions. Secondly, the PDGA results in faster convergence when compared to the DGA. This is clearly shown in Figure 5.13 by looking up how many generations of the PDGA will result in the same fitness value as the corresponding DGA: it is found that the PDGA can produce the same quality as the DGA (fitness of 8.9%) in 10 of the number of generations comparing to 27 for the DGA (see dotted line), i.e. the PDGA required approximately 1/3 number of generations to find the optimal fitness function values, comparing to DGA.

- *Convergence speed*: As stated earlier in Section 5.3.3, the DGA outperforms the SGA in terms of convergence speed (i.e. CPU time). However, the multi-deme structure of GA populations did not have any significant on the convergence speed of the algorithm in terms of total CPU time (i.e. 9 and 11 hours for DGA and PDGA, respectively). While PDGA requires more communication time between the core and slave processors for creating the necessary files/directories, fewer number of generation was required for convergence of this algorithm.

Table 5.7 provides the results of the above experiments with regards to the various traffic data, solution algorithms, convergence speed, and the quality of the final calibrated parameters.

Table 5.7 Summary of the experiments

Experiment No.	Purpose	Traffic data	Computing structure	GA structure	Outcome
1	Compare the calibration results based on different traffic data	Comparing Traffic counts from loop detectors vs. multi-source traffic data (count and speed values from AVI data)	Distributed computing in HPC	Simple GA (SGA)	AVI data improves the quality of solution
2	Comparing simultaneous vs. bi-level calibration approach	Count and AVI data	Distributed computing in HPC	Simple GA (SGA)	Simultaneous calibration methodology outperforms the sequential bi-level approach
3	Compare the optimization engines	Count and AVI data	1 processor vs. Distributed computing in HPC	Simple GA (SGA)	The distributed GA is approximately 10 times faster than the simple GA running in one processor (i.e. CPU time).
4	Comparing the distributed GA (DGA) to the parallel distributed GA (PDGA)	Count and AVI data	Distributed computing in HPC	Simple GA (SGA) vs. parallel GA (PGA)	• PDGA can produce the same quality as the DGA in 1/3 of the number of generations. • PDGA outperforms the DGA in terms of quality of solution and convergence speed (i.e. required number of generations).In terms of CPU time, there is no significant difference between the PDGA and DGA.

5.4 Sensitivity analysis

In the previous section of this thesis, the calibration results of different scenarios of the synthetic network, based on different scenarios of the available traffic data (i.e. count, speed, OD flows), were presented. The purpose of this section is to describe the methodology for estimating the weighting values in the aforementioned objective functions.

A sensitivity analysis was performed to test the sensitivity of the calibration accuracy in terms of NRMSE to the weight given to sensor counts, speed data, and OD flows' deviations in the stochastic objective function (Equation 1). According to the Section 5.3.1, the base calibration case was defined as the calibration process using loop detector count, while the base calibration case + AVI data represented the results obtained from the multi-source traffic data. The calibration results in Section 5.3 were based on the optimal weights given to different components of the objective function. In general, the sensitivity analysis was carried out on two cases:

- Case I) Estimate the relative weight given to speed, count, and OD flows (i.e. base case+AVI)
- Case II) Estimate the relative weight given to count, and OD flows (i.e. base case)

As stated earlier, for comparison purposes between different scenarios, the final optimal weighting schemes were used for both simple GA and parallel GA structures.

5.4.1 Case I

In order to evaluate the effect of weighting factors on the final results, different combination of weight values were considered, ranging from 0% to 100% for all parameters. Given the synthetic structure of the network and traffic data, the following weighting schemes were considered for the analysis. It should be noted that for the large-scale network, the optimal values of the weighting functions are estimated as a part of the calibration framework. Readers are referred to Section 3.2.2 for the detailed methodology.

Table 5.8 summarizes the NRMSE value for sensor counts, AVI data, and OD flows, respectively.

Table 5.8 Sensitivity analysis based on NRMSE (Case I)

Scenario number	Relative weight			NRMSE values			Fitness Function Value (Z)
	OD flows (α_1)	counts (α_2)	speed values (α_3)	OD flows	Counts	Speed	
1	0.1	0.45	0.45	10.2%	9.7%	5.7%	8.0%
2	0.1	0.4	0.6	10.8%	10.2%	4.4%	7.8%
3	0.1	0.6	0.4	10.6%	8.4%	6.1%	8.5%
4	0.2	0.3	0.5	9.5%	7.1%	4.9%	6.5%
5	0.2	0.4	0.4	9.8%	8.9%	6.2%	8.0%
6	0.33	0.33	0.33	10.2%	7.9%	7.4%	8.4%
7	0.3	0.3	0.4	9.8%	7.6%	6.4%	7.8%
8	0.4	0.2	0.4	9.0%	7.7%	6.8%	7.9%
9	0.4	0.3	0.3	8.5%	8.2%	7.6%	8.1%
10	0.5	0.2	0.3	7.9%	9.0%	8.3%	8.2%
11	0.5	0.3	0.2	7.9%	9.4%	9.7%	8.7%

Based on the results of the sensitivity analysis, the following observations were made:

- First scenario: initial trial with the weighting factor combination of 0.1, 0.45, and 0.45 for OD flows, counts, and speed values, respectively.

- In the 2nd and 3rd scenarios, the weight for the OD was set to be the same as the first scenario (i.e. 0.1). It was observed that as the weight of speed data increases, the speed/travel time calibration accuracy increases (i.e. lower speed NRMSE). However, that was achieved at the cost of slightly decreased accuracy for the count data. Therefore there is a trade-off between the weights given to all parameters. In general,

the calibration framework was found to be more reliant on the speed/travel time data which resulted in the increase of the objective function in the 3rd scenario comparing to the 2nd scenario (i.e. worsen results). From the traffic management viewpoint, this result is as expected since the speed/travel time data are the direct measures of the link performances and highly affect the routing decision, compared to the traffic counts from inductive loop detectors.

- As the larger weight is given to the OD flows and speed values in the 4th scenario comparing to the 3rd scenario, the NRMSE of all three components have decreased comparing to the previous scenario. However, the slight change in the weighting schemes of count and speed data in the 5th scenario with the same weight for the OD flows (i.e. 0.2) resulted in an increase of the total NRMSE (i.e. 8.0% in 5th scenario vs. 6.5% in 4th scenario). This finding is in line with the previous observation as the calibration framework is more reliant on the speed/travel time data.

- An equal weighting scheme in the 6th scenario gives the second least desirable results so far (i.e. 8.4%), while increasing the weight of the speed values in the 7th scenario slightly reduces the total NRMSE (i.e. from 8.4% to 7.8%).

- An important general observation can be made from the remaining scenarios. With a larger weight given to the OD matrix, the NRMSE of the estimated demand gets closer to those of the prior demand. This means that a large weight on prior error term would strap the solution closer to the prior demand and possibly prevents discovering the true matrix. In addition, the NRMSE of the count and speed values significantly increased, resulted in an increase of the total NRMSE (up to 8.7% in the last scenario).

In summary, the best fitness results correspond to the 4th scenario with the weighting scheme of 0.2, 0.3, and 0.5 for the OD matrix, traffic counts, and speed values, respectively. This weighting scheme indicates that the apriori demand gives a good starting search space by building the initial GA population; however it does not have to be necessary accurate in the objective function. In addition, it was found that more weight can be given to the speed/travel time data from AVI sensors as the direct measures of the link performance, which can affect

the routing decision, comparing to the traffic counts from inductive loop detectors. These findings from the synthetic case study are further analyzed in the next chapter with a more advanced weighting scheme.

5.4.2 Case II

Following the same procedure of Case I, and given that the calibration involves only two components (i.e. count and OD flows), the following weighting schemes can be considered for the analysis:

Table 5.9 Sensitivity analysis based on NRMSE (Case II)

Scenario number	Relative weight		NRMSE values		Fitness Function Value (Z)
	OD flows (α_1)	counts (α_2)	OD flows (α_1)	counts (α_2)	
1	0.1	0.9	28.9%	9.5%	11.4%
2	0.2	0.8	18.9%	9.3%	11.2%
3	0.3	0.7	14.2%	9.8%	11.1%
4	0.4	0.6	12.3%	10.2%	11.0%
5	0.5	0.5	11.2%	12.9%	12.1%
6	0.6	0.4	10.4%	14.6%	12.1%
7	0.7	0.3	9.8%	15.6%	11.5%
8	0.8	0.2	9.6%	22.3%	12.1%
9	0.9	0.1	8.9%	32.2%	11.2%

In summary, the best fitness results correspond to the 4th scenario with the weighting scheme of 0.4, and 0.6 for the OD matrix, and traffic counts, respectively.

5.5 Summary

In this chapter, a small network with synthetic data was used to demonstrate the feasibility of the proposed simultaneous calibration process, identify the effect of augmenting the GA operator with parallelization and distrusted computing schemes, and evaluates the effect of adding the AVI data into the calibration process. Several simulation experiments were performed to achieve the objectives of this chapter. In summary, the following observations were made:

- AVI speed data can improve the quality of solution,
- Simultaneous calibration methodology outperforms the sequential bi-level approach,
- The distributed GA (DGA) is approximately is 10 times faster than the Simple GA running in one processor (i.e. SGA),
- PDGA can produce the same quality as the DGA in 1/3 of the number of generations,
- PDGA outperforms the DGA in terms of quality of solution and convergence speed (i.e. required number of generations), and
- PDGA improved the quality of solution without having any significant impact on the computational time for a small synthetic network.

Based on the results obtained from the synthetic network, the proposed methodology is implemented in a complex large-scale network. The implementation details are described in the next chapter.

Chapter 6: Case Study II: Water Front Network

The previous chapter demonstrated the application of the simultaneous calibration approach for calibrating the demand as well as supply parameters of a DTA system and its advantages over the traditional sequential approach. It also established the importance of incorporating the AVI data into the calibration process for improved solution quality. Finally, it was found that the parallelization and distribution of GA can significantly enhance the calibration accuracy and convergence speed of the algorithm.

So far, the performance of the proposed DTA calibration framework was tested on the synthetic network with 41 numbers of unknown parameters (i.e. 9 OD flows in 4 departure intervals and 5 driver behavior and route choice model parameters). In larger networks, the numbers of unknown parameters increases with the larger number of OD pairs and departure intervals, additionally the path choice becomes more complicated. Thus, the aim of this experiment is to test the performance of the proposed approach for a more general real-size network where drivers have several route choice alternatives to reach their destination. Therefore, the number of parameters to be calibrated is much higher than the previous small network. The Water Front area in the downtown Toronto has been used for calibration purposes. The traffic data were incorporated from different sources to enrich the accuracy of the calibration process.

In summary, the objectives of this chapter are to:

- Demonstrate the application of the simultaneous calibration process for a realistic large-scale network;
- Evaluate the effect of adding the enriched in-vehicle navigation system data from private data provider into the calibration process;
- Evaluate the dependency of the calibration results to the historical OD flows; and
- Application of parallelization of GA control parameters with an HPC to expedite the GA calibration speed and accuracy for the large-scale transportation problems.

The remaining of this chapter is organized as follows: the next section describes the Water Front network, following by the input parameters for calibration. The observed traffic data, including historical OD flows, observed traffic sensor data, observed turning movement counts, and speed data are presented in the remaining sections of this chapter. Section 7 describes the implementation details of the case study, following by the GA control parameters. Calibration results are presented in the ninth section of this chapter. After the calibration process, the validation results for the Water Front network are presented in section 10. Finally, the last section summarizes this chapter.

6.1 Experimental Design

As a case study, the developed methodology is implemented on a very busy part of Toronto, the financial district in the downtown area. The network covers two major highways, namely Don Valley Parkway and Gardiner Expressway, arterials, minor roads, signalized intersection, and uncontrolled intersections. The network is bounded by Queens Quay corridor (south), Front Street (north), Don Valley Parkway (east), and Bathurst Street (west).

The network was originally coded in Paramics Version 5 in a project conducted for the Toronto Waterfront Revitalization Corporation [100]. Within that project, efforts were invested into building the correct geometry, defining the roadway attributes (speeds, and land configurations) and coding signal timing. For signalized intersections, actuation algorithms were developed to best represent the SCOOT traffic signal control system in the Waterfront area. The reader is referred to [100] for a detailed information on the Waterfront network coding effort. Figure 6.1 shows the coded Water Front network in Paramics, which contains 53 OD pairs, 1,483 segments, 121 junctions, 563 nodes, and 293 km of roadway. Each major roadway crossing the study area boundary is considered to be a gateway zone to the study area. A total of 26 gateways were included in the simulation model. The remaining 27 traffic zones were considered as internal study area zone. The gateways to the study area are presented in Table 6.1 and are shown in Figure 6.1.

Table 6.1 Gateways to the Study Area

Gateway zone ID	Roadway	Location
1	Lakeshore Blvd. West	East of Parklawn Rd.
2	S. Kingsway Ramp	Just south of The Queensway
3	Gardiner Expressway	East of Parklawn Rd.
6	Bathurst St.	North of Front St.
7	Portland St.	North of Front St.
8	Wellington St. West	West of Spadina Ave.
9	Spadina Ave.	North of Wellington St. West
10	Peter St.	North of Wellington St. West
11	John St.	North of Wellington St. West
12	Simcoe St.	North of Wellington St. West
13	University Ave.	North of Wellington St. West
14	York St.	North of Wellington St. West
15	Bay St.	North of Wellington St. West
16	Yong St.	North of Wellington St. West
17	Church St.	North of Wellington St. West
18	Lower Jarvis St.	North of Front St.
19	Sherbourne St.	North of Front St.
20	Parliament St.	North of Front St.
21	Eastern Ave.	North of Front St.
22	Sumach St.	North of Front St.
23	Don Valley Parkway (DVP)	North of Front St.
24	Lakeshore Blvd. East	East of DVP
25	Don Roadway	South of Lakeshore Blvd. East
26	Cherry St.	South of Gardiner Exp.
27	Harbour St.	West of Yong St.
28	Eireann Quay (Downtown Airport Rd.)	South of Queens Quay West

Figure 6.1 Paramics network topology

The simulation period took place between 6:30 to 9:30 for the AM peak period, and 15:30 to 18:30, for the PM peak period, with a warm-up period of 6:00 to 6:30, for a typical week (excluding weekends) in the Fall season of 2012 (15^{th}-19^{th} of October, 2012, aggregated for each 15 minutes interval during the peak periods). It should be noted that the duration and occurrence of the peak periods for Greater Toronto Area (GTA) were obtained from the 2006 Transportation Tomorrow Survey (TTS) and bi-annual Travel Time Studies conducted by the Ministry of Transportation Ontario (MTO) [101]. The six-hour peak simulation (i.e. 3 hours AM peak and 3 hours PM peak) was divided into 24 departure intervals of 15 minutes each.

6.2 Input Parameters for Calibration

As stated earlier, the initial population consisted of randomly perturbed ODs from historical OD flows and perturbed driver behavior parameters from Paramics' default values. On the demand side, the Water Front network consists of 2809 OD pairs (53×53) for 24 time intervals (i.e. 6 hours of 15-minutes interval during the AM and PM peak periods). On the supply side, there are 5 parameters for calibration, namely mean headway and mean driver reaction time as the 2 driver behavior parameters in Paramics, and perturbation factor, feedback and familiarity as the 3 route choice model parameters. These supply parameters were separated for the 24 time intervals to investigate the temporal variations of the Paramics model parameters in different time intervals. Therefore, the total number of demand and

supply parameters to be calibrated equals $(2{,}809 \times 24) + (24 \times 5) = 67{,}536$. It should be noted that this number is the *theoretical* number of unknown parameters. Since many of the OD pairs are not feasible (i.e. zero counts), the real number of unknown parameters will significantly decrease. This is further explained in Section 6.7.

6.3 Historical OD Flows

The historical OD matrix was extracted from the Toronto Tomorrow Travel Survey in 2006 via traffic assignment using EMME/2, and used as input for the microscopic traffic simulation model. The TTS is the largest and most comprehensive travel survey in Canada and is conducted once every five years. The TTS covers 5% of all households in the Greater Toronto Area (GTA) and surrounding areas [91]. The data used in this application are the TTS records collected for the year 2006. The 2011 survey data were still undergoing final refinements at the time of conducting this research and hence were not used.

The demand estimation model includes the entire Greater Toronto and Hamilton Area (GTHA) which is divided into six regions; namely, Toronto, Durham, York, Peel, Halton and Hamilton (Figure 6.2). The demand was calibrated at the GTHA's level to reflect traffic counts at cordons across the City, and was further calibrated for the Toronto Waterfront Area using OD matrix updating. Comparing to 2001 OD flows, the estimated 2006 static OD demand for the Water Front Network exhibits no increase in all trips entering exiting or travelling through the study area. Internal trips and outbound trips are estimated to have increased by 13% and 19%, respectively. Inbound trips are estimated to have decreased by 5.9%, which largely reflects a decrease in inbound trips from the west side as observed in the count data. In addition, through trips decreased by 0.6%. The steps for this stage and the data sources used to develop the model are described in detail in [102].

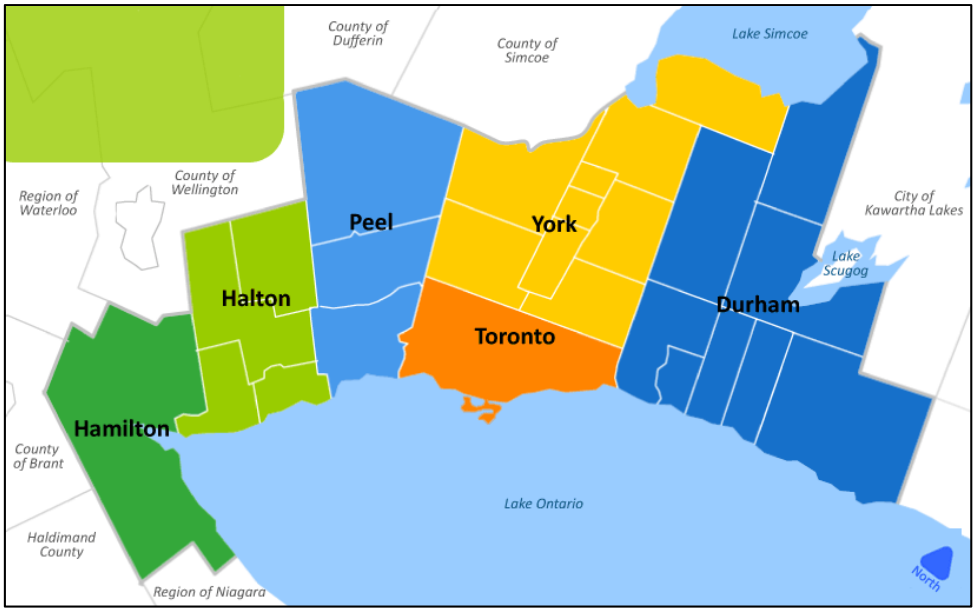


Figure 6.2 The Greater Toronto and Hamilton Area [101]

The 2006 static OD auto matrices were developed for the AM peak and PM peak. To obtain the dynamic OD matrices, the static passenger planning matrices were distributed across the 24 time intervals within the study period. The fraction of the flows assigned to a particular hourly interval was proportional to the AM and PM peak hour factor (i.e. trips in the peak hour / trips in the 3-hour peak period). The appropriate AM and PM peak hour factor were found to be 41.2% and 37.6%, respectively [102]. According to the 2006 TTS, the AM and PM peak hours (i.e. trips in the peak hour compared to trips in the 3-hour peak period), were found to occur from 7:30-8:29 AM and 4:30-5:29 PM, respectively [102]. In the next step, the normally-distributed hourly OD matrices were further broken down into the 15-minutes intervals based on the normal distribution and the same AM and PM peak hour factors (i.e. 41.2% and 37.6%).

Finally, the approximated a priori OD flows for each interval were randomly perturbed in a pre-specified range (-50% up to 50%) to obtain dynamic perturbed OD flows as the first starting values for calibration. This user-defined wide perturbation range ensures that the starting OD values do not replicate the time-dependent OD flows in the minimization problem.

6.4 Observed Traffic Sensor Data

The loop detector traffic counts for the highway segments of the network were obtained from the ONE-ITS platform associated with the study time periods. The network contains 1483 links, and out of which, 67 links[4] along the Gardiner Expressway and Don Valley Parkway (DVP) were equipped with loop detectors. For the purpose of this research, the loop detector measurements were aggregated into 15 minutes intervals [103]. Hence they provide 67 sets of link flow counts for each interval. Over all the time intervals, they provide $67 \times 24 = 1{,}608$ sensor measurements. Figure 6.3 presents the spatial distribution of the available loop detector measurements in the Water Front network.

Figure 6.3 Loop detector counts along the sections of the study area

As is apparent from Figure 6.3, the loop detector data were only available for calibration of the network along the highway corridors.

6.5 Observed Turning Movement Counts

Given the lack of the observed loop detector counts, this research incorporated the traffic data from turning counts at selected intersection. For the purpose of this research, the turning counts at 60 key signalized intersections were obtained from the City of Toronto database for

[4] At the time of obtaining the data

the typical weekdays in the Fall season of 2012. It should be noted that the provided traffic data were only disaggregated to the hourly mean counts for AM and PM peak hours. Given the available aggregated peak hour (i.e. 7:00-8:00 and 4:30-5:30) turning movement for 356 links from these intersections, the total observed count measurements from this data source would be 356×2=712.

6.6 Speed Data

As stated earlier, this research incorporated traffic data from different sources to enrich the accuracy of the calibration process. Given the appropriate sensor configuration throughout the network, it is argued that the calibration accuracy and estimation of OD flows can be significantly improved [32, 56]. Therefore, the purpose of incorporating the enriched speed/travel time data into the calibration process is to improve the calibration accuracy and minimize the dependency of OD estimation to the historical OD flows as starting point.

For the purposes of this research, a number of technologies from different data providers were available which are able to provide speed/travel time information, including:

- GPS-equipped probe vehicle technology;
- Mobile phone probes with GPS technologies;
- In-vehicle navigation system based technologies; and
- Bluetooth technology.

Among the above-noted technologies, the GPS-equipped probe vehicle technology was considered as the primary source of data for many experimental and research studies. However, the high cost of field data collection and limited study time periods made it less desirable to obtain the travel time information on a large-scale network. On the other hand, very few large-scale independent assessment and comparative analyses were conducted to evaluate other emerging technologies with the traditional GPS-equipped probe technology. Therefore, as a part of this thesis, a research project funded by Ministry of Transportation Ontario (MTO) was conducted to evaluate the recent emerging technologies with the

traditional GPS-equipped probe technology and select a more efficient data collection methodology for future travel time studies within the Greater Toronto Area (GTA). One of the challenges with these technologies is that they have different levels of accuracy and the different market penetration of the underlying technologies or devices. Therefore, it was essential to compare the various technologies and data sources in terms of travel time accuracy, reliability, and sample size requirements. The following sub-section briefly describes the nature of these data sources and the evaluation methodology for selecting the candidate technology.

6.6.1 Databases

Bluetooth technology

Bluetooth is a telecommunications industry specification that defines the protocol by which mobile phones, computers, personal digital assistants, car radios, and other digital devices can be interconnected using short-range wireless communications. Every Bluetooth device has a unique 48-bit address referred to as ID. Bluetooth transceivers that are powered on and are set in the "discover" mode continuously transmit their ID for the purpose of identifying a device to communicate with; and to establish a link with the “responding devices”. If receiver units are deployed on the side of roadways, they can register the ID associated with the Bluetooth enabled devices in vehicles that pass by as well as the time stamp associated with the detection instances. Therefore, travel time and average speed of individual vehicles on the road section between two consecutive Bluetooth receivers can be obtained.

Bluetooth receivers were installed on a sub-section of the study area along a few arterial corridors and freeway to freeway ramps. Once Bluetooth enabled devices are in range of one of the receivers, the Media Access Control ID (MACID) of the Bluetooth device and the timestamp associated with this event were recorded. Then the MACIDs are matched between two consecutive Bluetooth receivers to calculate the travel time of individual vehicles.

In-Vehicle Navigation Systems

The data from in-vehicle navigation systems were purchased from a data provider, which has millions of navigation devices in use around the world with a comprehensive historical database of traffic information. The service provider has developed a service which provides historical traffic information (e.g. travel time, speed, standard deviation of travel time, etc.) about transportation networks to potential customers in various geographical areas in the world.

The data obtained from this provider included two components: (1) network data and (2) traffic data. The network data were obtained in the form of a GIS map that were then geo-referenced to the following traffic data: average travel time, standard deviation of travel time, average speed, number of observations, and percentiles of speed from 5% to 95% in increments of 5% (i.e. 5^{th} percentile speed, 10^{th} percentile speed, etc.).

The data provider was only able to provide the aggregated traffic data for every 15 minutes and for every week in the study period (e.g. every Monday through Friday from 7:15 a.m. to 7:30 a.m.). As a result, the data provided neither contained travel times of *individual vehicles* which travelled each roadway segment during the study period nor travel time for each *individual day* (e.g. October 3^{rd}, 2012 from 7:15 a.m. to 7:30 a.m.). The primary source of traffic information was passenger cars and the technology was similar in nature to traditional travel time studies conducted by MTO.

Mobile Phone Probes with GPS

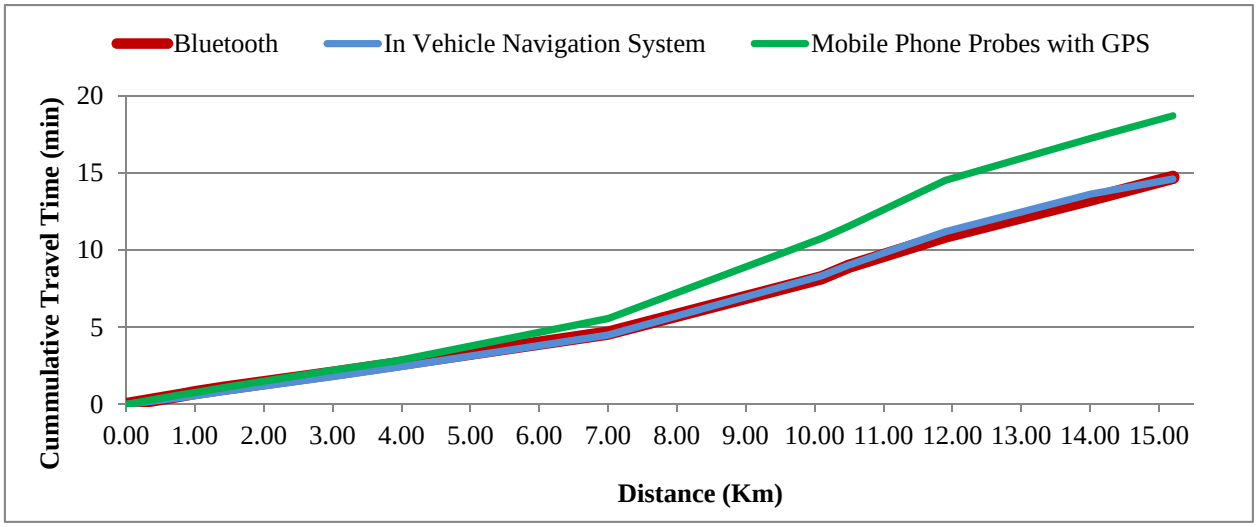
These data were also purchased from a data provider that uses a proprietary data fusion engine to process various sources of data and generate traffic information. The data sources include mobile phones, GPS navigation systems, and other sources of data, covering overs 5,419 centerline km and 21,963 centerline km in the GTA and Ontario respectively. The coverage area includes freeways, urban arterials, rural arterials, and side streets. The data is in a very similar format (i.e., network data in a GIS format and traffic data) to that of the previous data provider with the following exception:

- Average travel time and standard deviation of travel time were not explicitly provided. Therefore the average travel time and standard deviation of travel time (or variance of travel time) were estimated from average speed and variance of speed [105].

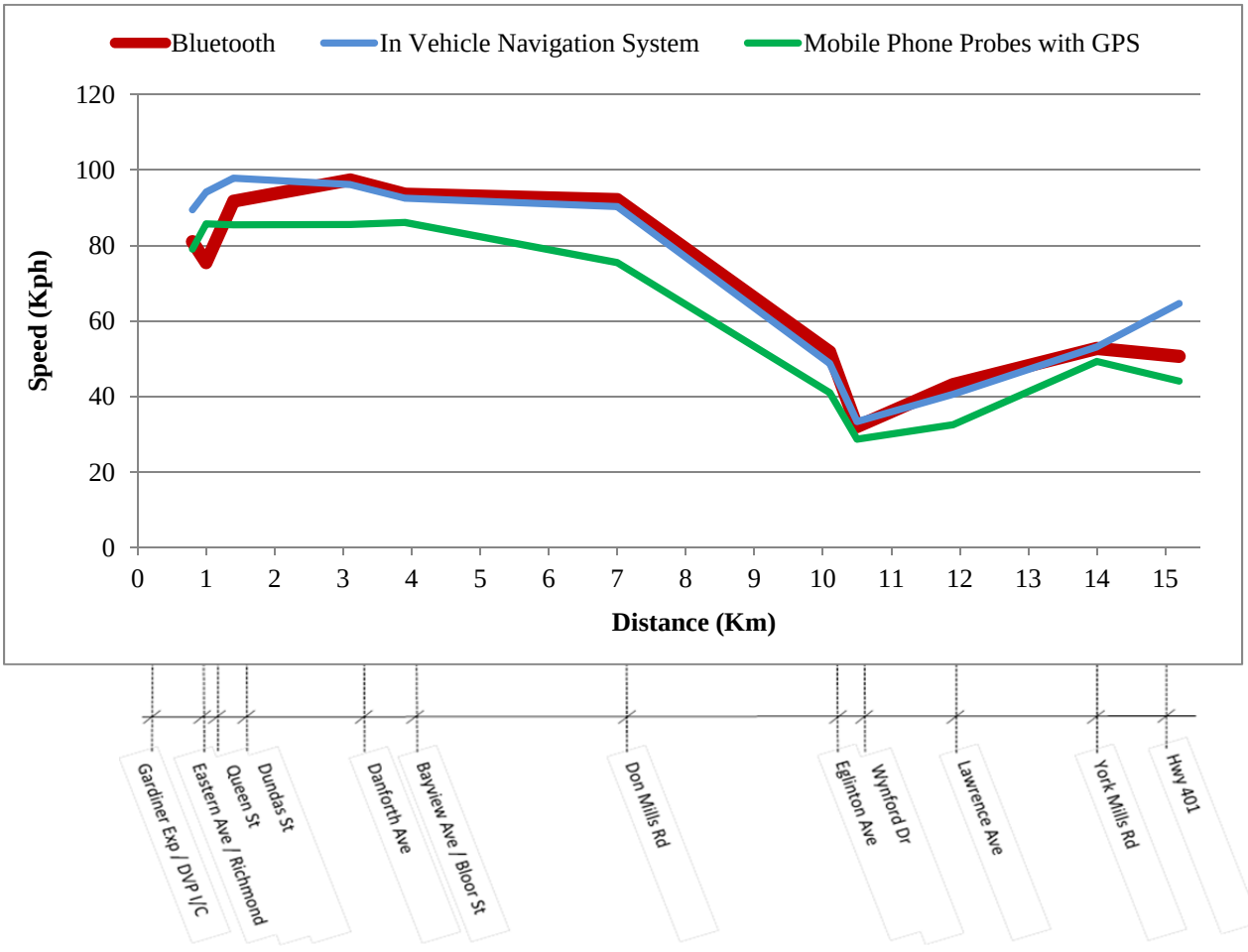
6.6.2 Evaluating Available Data Sources: Selecting the Candidate Technology

The traffic data gathered from various data providers were validated in order to maintain the database consistency with the benchmark (i.e. GPS-equipped probe vehicle data) and prepare the database for comparison between each data source. A multi-criteria methodology was developed to evaluate data from each data provider based on accuracy, coverage, number of observations, and ability to provide data for special facilities such as High Occupancy Vehicle (HOV) lanes.

In terms of accuracy, it was found that the Bluetooth technology is superior to the GPS-equipped probe vehicle technology and can serve as the "ground truth" to evaluate the data purchased from other vendors. Therefore, the accuracy of the data provided from the in-vehicle navigation system and mobile phone probes with GPS technologies were evaluated against the Bluetooth data. Figure 6.4a and Figure 6.4b represent visual comparisons of cumulative travel time and speed profile for an arterial route (which consists of multiple road segments) between the Bluetooth, in-vehicle navigation system, and mobile phone probes with GPS, during the study period respectively. As can be seen in the figures, the data obtained from the in-vehicle navigation systems is closer to the Bluetooth data, as the ground truth. However, as the average travel time and standard deviation of travel time were estimated for the mobile phone probes with GPS technology, this comparison might be biased.



a) Cumulative travel time



b) Speed

Figure 6.4 Cumulative travel time and speed for DVP northbound during AM peak period

In terms of accuracy, the results of Figure 6.4 suggest that the Bluetooth technology can be replaced with the in-vehicle navigation system technology, as the wide-area deployment of the Bluetooth receivers in a large-scale network, such as the Water Front network, could to be costly.

In terms of number of observations, in-vehicle navigation system had the highest number of observations per road section per peak period with more than 2300 observations for arterial roads and 3700 for freeways. The Bluetooth technology ranked second and mobile phone probes with GPS technology ranked third. It is noteworthy that the number of observations for the GPS-equipped probe vehicle (which is traditionally collected by road agencies) was as low as approximately 10 observations per link per AM and PM time periods. Each of the three data sources evaluated in this research provided significantly more observations. It should also be noted that the real time traffic data can be directly collected from Bluetooth receivers deployed onto the road; however there might be difficulties for collecting real-time traffic data from mobile phone probes such as encouraging smart phone owners use the application for tracking the device. This might highly affect the penetration rate of the mobile phone probes data.

In-vehicle navigation system and mobile phone probes with GPS technology were able to provide traffic data for collector and express facilities, as well as separate data for HOV lanes and GPL. The Bluetooth technology was found generally incapable of providing data separately for GPL and HOV or express and collector lanes.

In summary, the data provided by the Bluetooth technology was found the closest to the truth and the most reliable data source. However, the challenges involved with implementation of the Bluetooth receivers in a wide-area network and processing the data made it less desirable for MTO as the primary data source for the future Travel Time Studies. Therefore, it was recommended to acquire the traffic data from the private vendor associated with the in-vehicle navigation technology.

6.6.3 Speed Data from In-Vehicle Navigation System Technology

As stated earlier, the available data from in-vehicle navigation systems were obtained from a private data provider, which included two components: (1) network data as shown in

Figure 6.5 and (2) traffic data. The network data were obtained in the form of a GIS map in which each link had a unique link ID. The traffic data was received in text files including link ID, date, time, average travel time, standard deviation of travel time, average speed, number of observations, and percentiles of speed from 5% to 95% (i.e. 5^{th} percentile speed, 10^{th} percentile speed, etc.). Table 6.2 presents a sample raw data obtained from the data provider.

Table 6.2 Sample raw data

TIME BIN	HITS	AVG TT (SEC)	AVG SPEED (KPH)	STD DEV TT	P5_TT	P10_TT	P15_TT	…	P95_TT
Weekdays AM peak	37	280.2	32.3	36.8	144.6	152.9	160.1	…	683.9
Weekdays PM peak	80	308.7	29.3	53.5	147.7	157	164	…	724
7:15-7:30	12	259.3	34.9	16.2	229.6	229.6	229.6	…	291.4
9:15-9:30	26	286.7	31.6	35.5	158.2	158.2	169.8	…	564.5
17:00-17:15	74	289.9	31.2	30.2	186.7	186.7	187.6	…	451.4

For the purpose of this thesis, the following temporal and spatial distributions of the traffic data were available:

- *Temporal distribution*: Selected weekday traffic data associated with the Fall season of 2012, averaged for both peak periods (e.g. 15:30-18:30 for PM peak period) and every 15-minutes (e.g. 15:30-15:45 averaged for the selected days).

- *Spatial distribution*: As stated earlier, the Water Front network contains of 1,483 segments, and among those, the traffic data was available for 438 segments in the study area. In other words, the speed/travel time data coverage is approximately 30% of the road network.

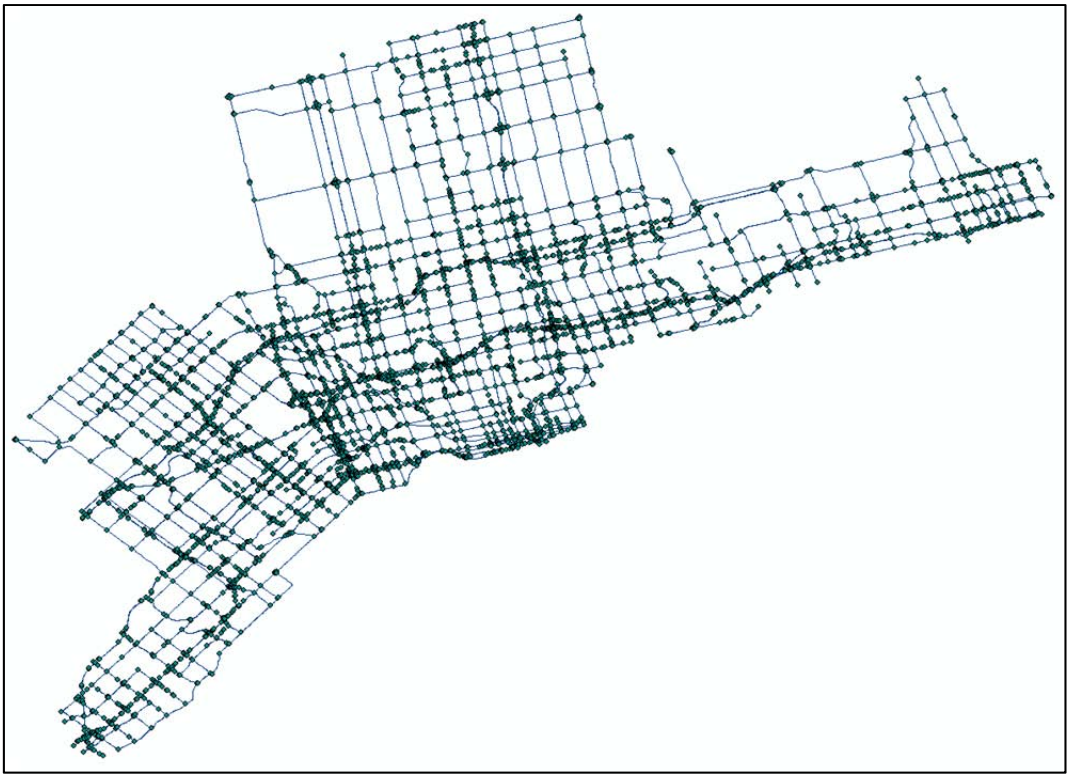


Figure 6.5 Network data coverage

In the next step, the aggregated 15-minutes data obtained from data provider were spatially matched with the study area segments. In this research, a road segment was defined as the section between two consecutive interchanges for freeways and two consecutive major intersections for arterials. This is consistent with the MTO definition of road segments applied in biannual Travel Time Studies.

Travel time information including average travel time, variance of travel time, and average speed were obtained for the road segments. It should be noted that each of the interchange-to-interchange or intersection-to-intersection routes may consists of multiple links in the GIS of the vendor. Therefore, the GIS map of the vendor should be spatially matched with the segmentations of this research. For this purpose, Network Analyst Tool, as an extension of the ArcGIS engine, was used for the network-based routing analysis, to find the sequence of links between two consecutive boundaries for each segment, based on start-point and end-point coordination. The output of ArcGIS was then mapped with the associated traffic data. This process is called the Routing Process. Figure 6.6 visually illustrate the definition of the routes, segments, and links.

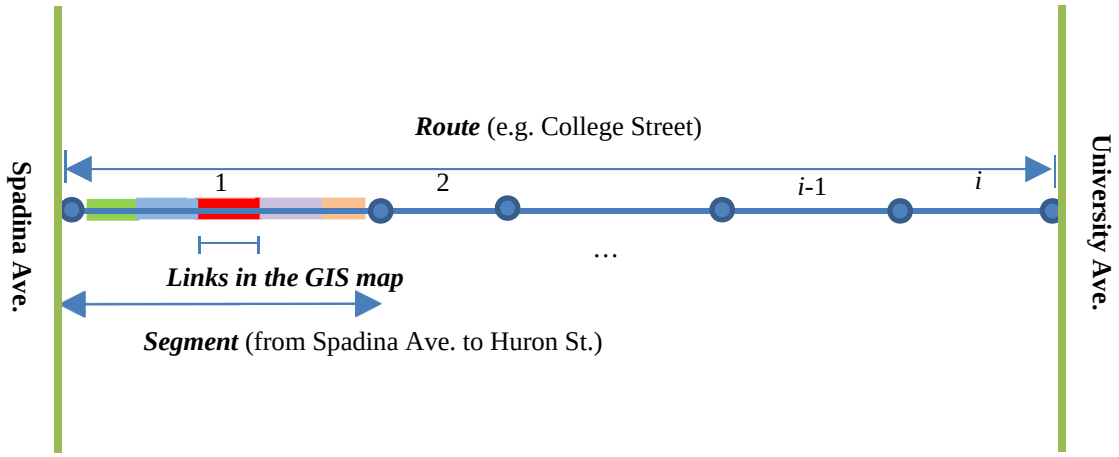
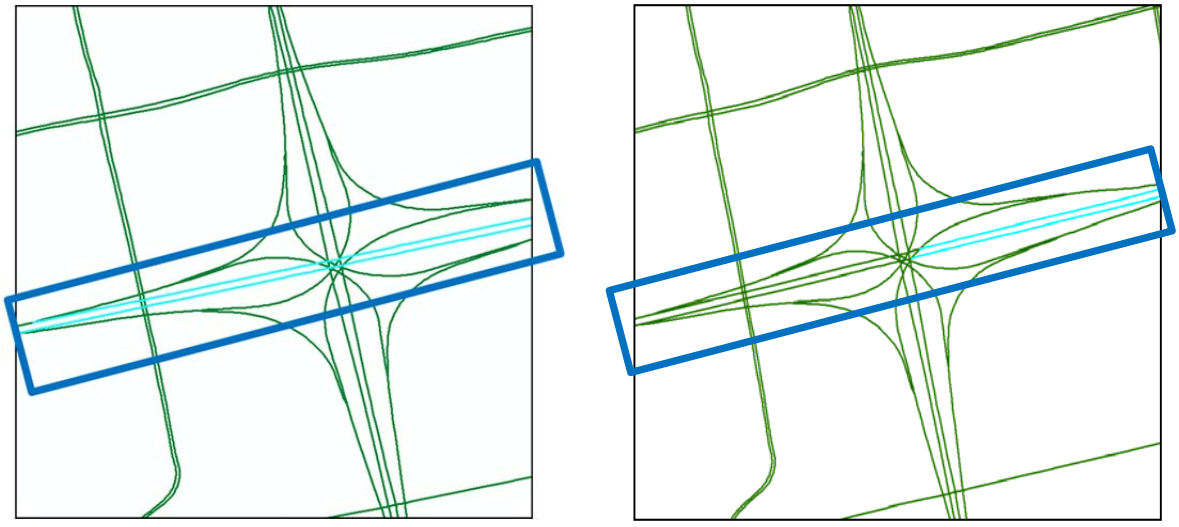


Figure 6.6 Route, segment, and link definitions

The main challenge associated with the spatial matching was that the GIS maps provided by the data vendors did not necessarily match pre-defined the road segmentations. For example, there was no node at the middle of interchanges and intersections. In other words, links were continuous through the interchanges and intersections. It was necessary to add nodes at interchanges and intersection because such nodes define the beginning or end of a given segment. Figure 6.7a illustrates a continuous segment at an interchange and Figure 4.4b shows the same interchange at the middle of which the links were broken. An extension tool in ArcGIS, called “Planarize lines”, was used to automatically break segments at the middle of interchanges and interchanges.



a) Continous links through an interchnage b) Broken links at the interchange

Figure 6.7 Continuous segments at an interchange

As stated earlier, for each road segment between two consecutive interchanges or intersections (i.e. segments), the following performance measures were provided:

- Average travel time;
- Average speed;
- Variance of travel time; and
- Percentiles of speed from 5% to 95%

The variance of speed for each road segment was calculated based on average travel time and variance of travel time as follows [105]:

$$V_s^{u,p} = \frac{V_s^{T,p}}{\left(\bar{T}_s^{\,p}\right)^4} \times \left(L_s\right)^2 \tag{18}$$

Where:

$V_s^{u,p}$: Variance of speed (u) for segment s during time interval p

$\bar{T}_s^{\,p}$: Average travel time of segment s during time interval p

L_s: Length of Segment s

Table 6.3 presents an example of travel time and speed performance measures along the Gardiner Expressway (both eastbound and westbound) during AM Peak Periods of the Fall Season. In addition, the minimum sample size requirement and the observed number of hits (e.g. sample size) are provided in this table. As shown in this table, the provided sample size is significantly higher than the minimum required sample size. Appendix A provides the methodology for sample size requirement. It summary, the data collected satisfied the sample size requirements for more than 99.5% of the segments in all peak periods.

Table 6.3 Example of performance measures for Gardiner Expressway during AM peak periods of the selected week in the Fall season

Direction	Interchange (From/To)	Length (Km)	Average travel time (h:mm:ss)	Standard deviation of TT (sec)	Variance of TT (sec^2)	Average speed (Km/hr)	Standard deviation of Speed	Variance of speed	Minimum sample size	Hits
Westbound	DVP									
	Lower Jarvis	1.7	0:01:52	43.4	1883	54.7	21.2	450.2	57.8	564.3
	Spadina Ave	2	0:02:54	45.5	2070	41.1	10.8	117.1	26.6	542.8
	Jameson Ave	3.5	0:04:34	48.1	2313	46.3	8.1	64.9	11.6	787.5
	South Kingsway	3.3	0:03:16	27.3	745	60.4	8.5	71.4	7.5	872.2
	Islington Ave	3.8	0:03:29	23.9	571	64.6	7.5	56.1	5.2	940.1
	Kipling Ave	1.1	0:00:47	9.4	88	80.7	16.8	281.6	16.6	758.5
	Hwy 427	1.9	0:01:33	26.1	681	74.4	20.7	429.7	29.8	598.2
	Total/Average	17.3	0:18:25			56.37				

Direction	Interchange (From/To)	Length (Km)	Average travel time (h:mm:ss)	Standard deviation of TT (sec)	Variance of TT (sec^2)	Average speed (Km/hr)	Standard deviation of Speed	Variance of speed	Minimum sample size	Hits
Eastbound	Hwy 427									
	Kipling Ave	2.0	0:01:38	19.5	380	69.7	14.6	212.0	16.8	941.1
	Islington Ave	1.1	0:01:27	29	841	44	15.2	230.2	45.7	879.4
	Lake Shore Blvd (split)	2.8	0:04:41	39.4	1552	35.9	5.0	25.4	7.6	1052.1
	Jameson Ave	4.3	0:08:28	58.7	3445	30.5	3.5	12.4	5.1	797.9
	Spadina Ave	3.5	0:04:20	35.2	1239	48.6	6.6	43.2	7.0	797.9
	York St	1.0	0:00:56	18.4	338	64.2	21.4	459.1	42.8	464.3
	Lower Jarvis	1.0	0:00:42	15.2	231	86.8	30.9	953.2	48.6	406.6
	DVP	1.7	0:01:10	7.99	63	87.9	10.1	101.8	5.1	390.5
	Totals/Averages	17.4	0:23:21			44.72				

6.7 Implementation Details: A Note on the Degree of Freedom

The accuracy of the calibration process highly depends on the quantity and quality of the observed available data for calibration. Traditionally, the OD estimation problem was highly underspecified as the number of unknown parameters (i.e. OD pairs) was significantly higher than the number of observations (e.g. traffic data from loop detectors). Within the incorporation of the enriched travel time/speed data into the calibration process, this research aimed to minimize the degree of freedom of the subject network (i.e. difference between the number of observed and unknown parameters), and ultimately, steered the *estimated* OD flows to the *true* (and unknown) OD pairs.

As indicated earlier in Section 6.2, the *theoretical* number of unknown OD flows is 2809 (53×53) for each time interval. However, 2140 of those OD pairs are not feasible (i.e. with zero counts). Therefore, the *actual* total number of unknown parameters is:

$$(669 \times 24) + (5 \times 24) = 16{,}176.$$

On the other hand, the number of observed parameters is as follows:

- *Observed traffic sensor data:* The loop detectors are able to provide 67 sets of link flow counts for 24 time intervals. Therefore the total sensor measurement is 67×24=1,608.

- *Observed turning movement counts:* Given the available aggregated peak hour (i.e. 7:00-8:00 and 4:30-5:30) turning movement for 356 links (from 60 intersections), the total observed count measurements from this data source would be 356×2=712.

- *Speed data:* Among the 1,483 segments, the 15-minutes aggregated speed data were available for 438 segments in the study area. Therefore, the total speed available data is 438×24=10,512.

In summary, the total number of available data for calibration is equal to $1{,}608 + 712 + 10{,}512 = 12{,}832$. The degree of the freedom for the calibration problem would be equal to

$16{,}176 - 12{,}832 = 3{,}344$, which elaborates the complexity of the calibration process comparing to the synthetic test network with the degree of freedom of 20.

It is important to note that in the initial runs, the assumption of fixed driver parameters and route choice parameters was relaxed by allowing these parameters to change each 15 min. That was based on the assumption that driver's route choice and vehicle following parameters are a function of congestion. In addition, the weight factors were determined as part of the optimization function, rather than through trial and error. Sections 6.9.1 and 6.9.2 discuss further the results of these tests.

6.8 GA Control Parameters

Similar to Section 5.2, the GA control parameters were selected after some initial trials. Table 6.4 summarizes the simple GA (SGA) and parallel GA (PGA) control parameters. It should be noted that the GENOTRANS library enhanced the trial and error process to select the appropriate control parameters for the large-scale network.

Table 6.4 Simple and parallel GA control parameters

Category	Control Parameter	Selected Value
Simple GA	Selection method	Ranked-based selection mechanism
	Cross-over	α-blend crossover (α=0.2)
	Crossover rate	P_c= 90%
	Mutation method	Self-adaptive Gaussian
	Mutation rate	P_m= 5%
	Population size	100
	Number of generations	50
	Number of simulation runs for each chromosome evaluation in Paramics	3
Parallel GA	Island topology	Fully connected topology
	Migration policy	Good migrants replace bad individuals
	Migration rate	8%
	Migration frequency (epoch interval)	5
	Number of demes	4
	Deme size	25

As is apparent from this table, the number of population and generation are not significantly greater than the simple synthetic network, as the computational constrains assume an important role in the convergence speed of the algorithms. The above GA control parameters were found to be effectively covering the search space for the optimal solution. A bigger search space in turn implies a larger population for evaluation, consequently meaning that more computation resources are needed to evaluate a single generation; thus a longer time to reach convergence. Given the available number of processors and Paramics licenses in the HPC, the size of the network, and the level of congestion during peak hours, the experiments in this chapter were designed using the selected GA control parameters.

6.9 Calibration Results

Similar to previous chapter, a number of experiments were carried out to demonstrate the transferability of the calibration methodology to a large-scale complex network:

- *Experiment 1:* compare the calibration results from the loop detector and turning movement counts to the multi-source traffic data based on distributed computing.
- *Experiment 2:* evaluate the effect of parallelization structure of GA population on the calibration accuracy using multi-source data (i.e. DGA vs. PDGA).
- *Experiment 3:* OD estimation without count and turn data but using the raw speed data (segmentation)

6.9.1 Experiment I: Calibration based on Multi Source Traffic Data Using Distributed Computing

In this experiment, the calibration results from the loop detector and turning movement counts were compared with results of the multi-source traffic data. The main objective of this experiment was to quantify the impact of incorporating the enriched speed data into the calibration process. The experiment was conducted in the HPC and other available Paramics resources. The number of available observed data from loop detector and turning movement count is equal to 1,608+712=2,320. On the other hand, the number of available traffic data for the multi-source case study is equal to 16,176. Given the wide coverage of the speed data across the study area and full temporal distribution of the data across 24 departure intervals,

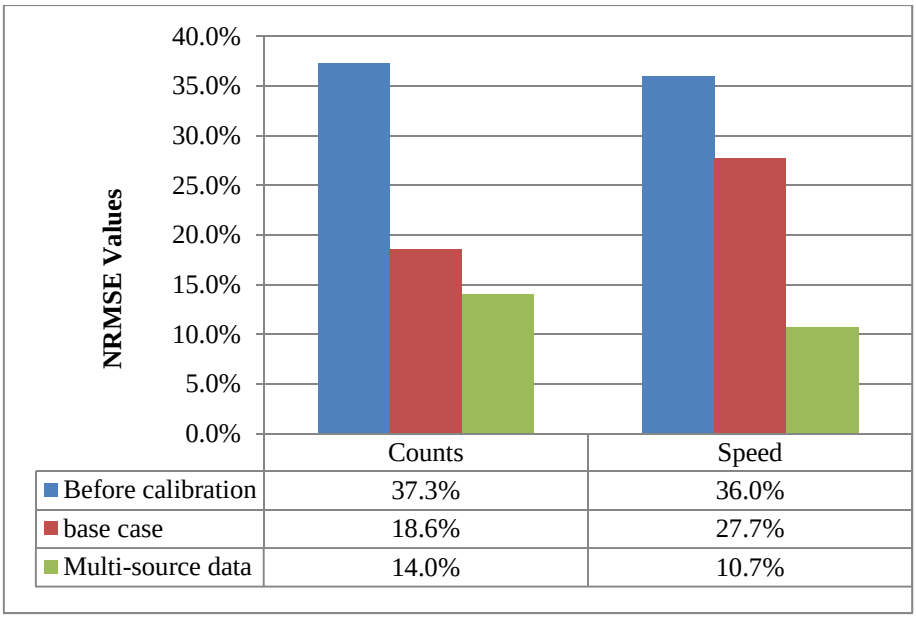
it was expected that the incorporation of the speed data significantly improve the calibration accuracy and find the solution efficiently.

In terms of computation time, the network was simulated a total of 15,000 times, which was the product of the number of generations (50), the number of chromosomes in each generation (100), and the number of DTA iterations in Paramics (3). The average CPU running time for the 6-hour simulation period (i.e. 24 time intervals of 15-minute each) was estimated to be an hour. Based on the available 80 distributed CPUs for running the simulation model, the calibration process took 198 hours (i.e. more than a week) to complete, compared to the estimated 15,000 hours running the process on a single CPU machine. In other words, the distributed GA was approximately 75 times faster than the simple GA. The final simulated dynamic OD flows, counts, speed data, and driver behavior parameters were obtained after running 46 generations. Table 6.5 present the calibration result statistics based on NRMSE measures of effectiveness for counts and speed values. For visual comparison, the MRMSE values for calibration are summarized in Figure 6.8.

Table 6.5 Calibration result statistics based on NRMSE

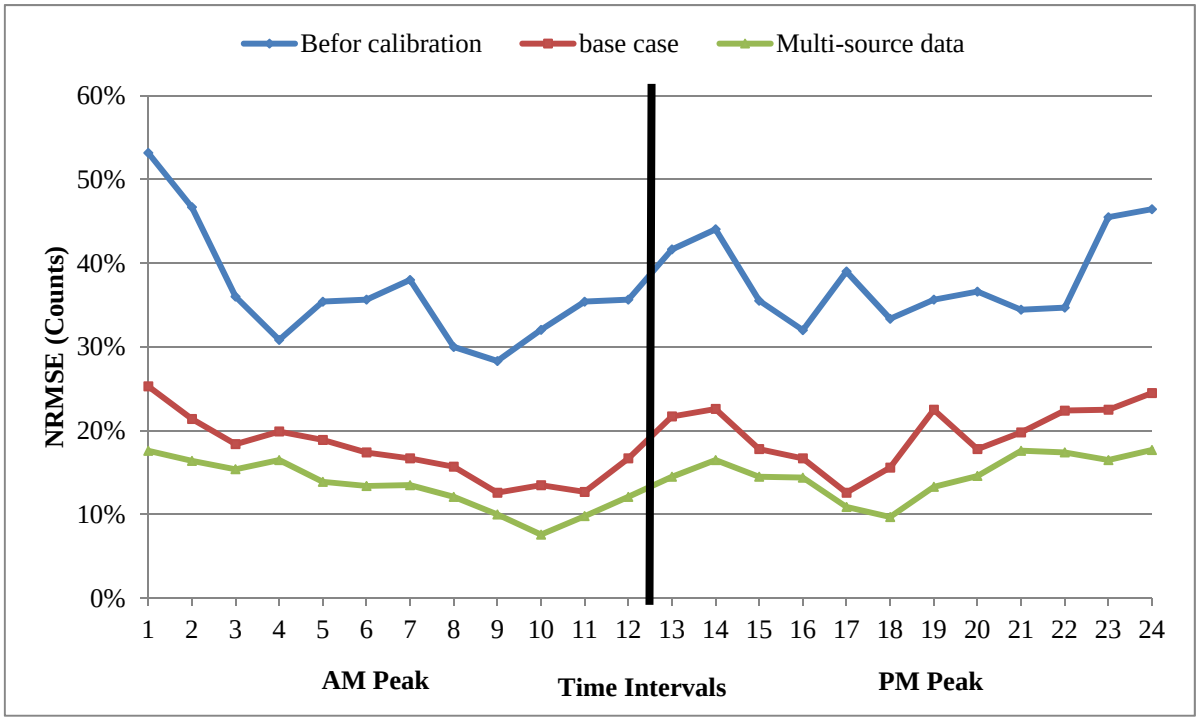
Category	Before Calibration	Base Calibration Case	Base Calibration Case + Speed Data	% change (Base Case vs. Before Calibration)	% change (Base Case vs. Base Case + Speed)
Count	37.3%	18.6%	14%	50.3%	24.6%
Speed	36%	27.7%	10.7%	23.2%	61.3%

In addition, Figure 6.9 presents the calibration results aggregated for each 15-minute interval of the analysis period (AM peak and PM peak) using the distributed GA (aggregated among all loop detectors and in-vehicle navigation system data). The number of intervals corresponds to the 24 simulation periods of 15-minutes, during AM and PM peak periods.

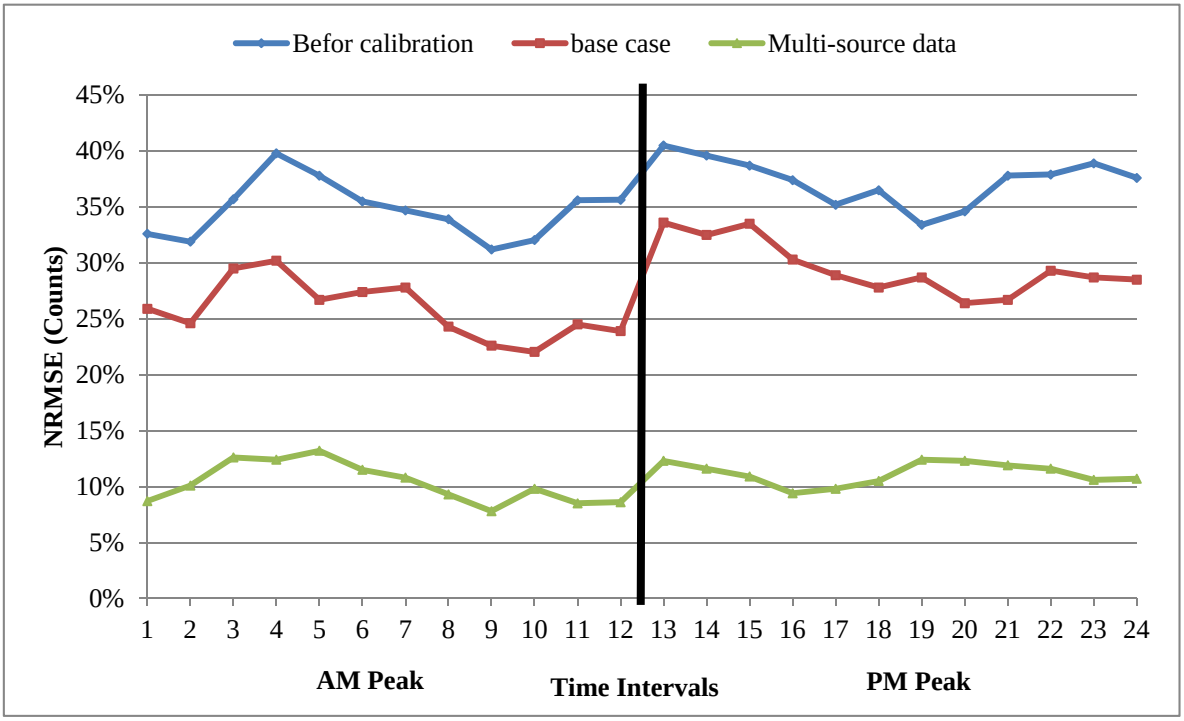


	Counts	Speed
Before calibration	37.3%	36.0%
base case	18.6%	27.7%
Multi-source data	14.0%	10.7%

Figure 6.8 Comparison of different scenarios based on NRMSE values



a. Count



b. Speed

Figure 6.9 Comparison between measures of effectiveness of different cases based on NRMSE for count and speed data

As is apparent from Table 6.5 and Figure 6.9, the simultaneous calibration process using the distributed GA performed as expected and had the capability of minimizing the discrepancy between the observed and simulated traffic conditions (i.e., OD flows, loop detector counts, and turning movements). According to Figure 6.9, the percentage change in terms of count NRMSE between the "base case" and the "before calibration" ranged between 35.4% and 67.7% for various time periods, with an average improvement of 50.3%. As noted earlier, the available traffic data from loop detector counts and turning movements were sparse relative to the number of unknown parameters and number of links in the network. However, such sensor coverage is not uncommon in the real-world situations with limited data. In this context, the proposed methodology was able improve the calibration accuracy by approximately 50%, comparing to the "before calibration" scenario.

Following the same trend as the count measures of effectiveness, the percentage change in terms of speed NRMSE between the "base case" and the "before calibration" ranged between 13.4% and 32.9% for various time periods, with an average improvement of 23.2%. It should be noted that for the "base case" scenario, the speed data was not incorporated into the calibration process. The speed NRMSE was calculated separately from the calibration process. In other words, it was found that the inclusion of count data (from loop detector and turning movements) into the calibration process improved accuracy not only in terms of lower count errors (i.e. 50.3%), but also lowered the travel time/speed errors (i.e. 23.2%).

In addition to conducting a before/after calibration study using traditional count data, this research also incorporated the speed traffic data obtained from in-vehicle navigation system data provider into the calibration process. According to Figure 6.9, the following conclusions can be made:

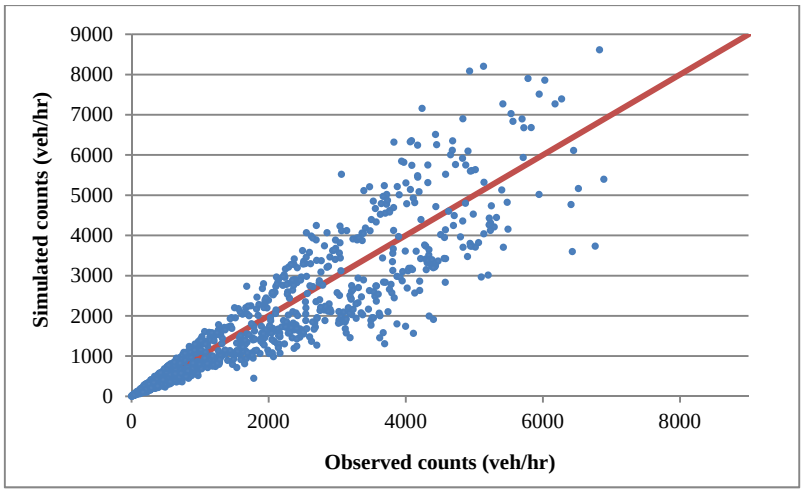
- Incorporation of the *speed* data increased the calibration accuracy and made a reduction in the *count* NRMSE measure of effectiveness by 24.6%, comparing to the "base case" scenario, and an overall improvement of 63.2% comparing to "before calibration" case. This observation is in line with the results of the previous case study indicating that the speed information provides a better indication of the quality

of traffic flow and moves in the right direction on the optimization surface more efficiently.

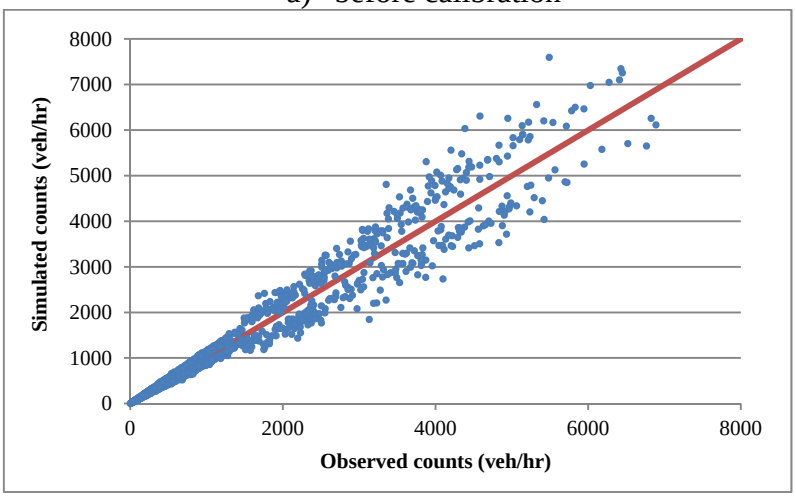
- Incorporation of the *speed* data increased the calibration accuracy and made a significant reduction in the *speed* NRMSE measure of effectiveness by 61.3%, comparing to the "base case" scenario, and an overall improvement of 70.0% comparing to "before calibration" case.

In summary, it was found that a) the proposed calibration methodology can efficiently estimate the unknown parameters and b) the inclusion of speed data into calibration process can significantly improve the quality of the solution and minimize the discrepancy between the observations and their simulated counterparts. It should be noted the previous studies in the subject of OD estimation and calibration of model parameters using the multivariate objective functions resulted in an improvement in the equivalent objective function ranging from 17% to 23% [30, 69, 73]. In this context, the proposed calibration methodology and the incorporation of the enriched speed data from GPS-enabled devices made a further progress towards the estimating the true dynamic OD flows and calibration the simulation model parameters.

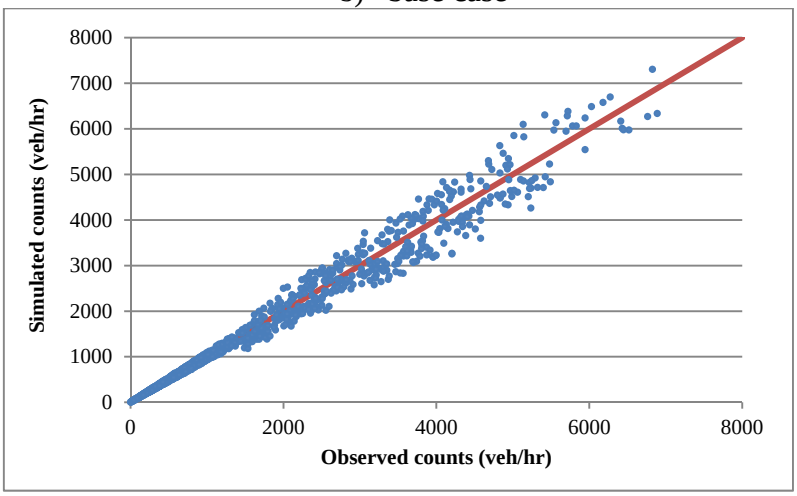
Figure 6.10 presents the plots of the observed *counts* (i.e. loop detector counts + turning movement count) and simulated entities for the before calibration, base calibration, and multi-source scenarios. In addition, the plots of the observed and simulated *speed* values in the multi-source scenario are presented in Figure 6.11. A visual comparison between these figures reveals that incorporation of the speed data from in-vehicle navigation system technologies significantly improved the calibration accuracy, in terms of NRMSE of both speed and count data. Appendix B presents these plots in more details.



a) before calibration

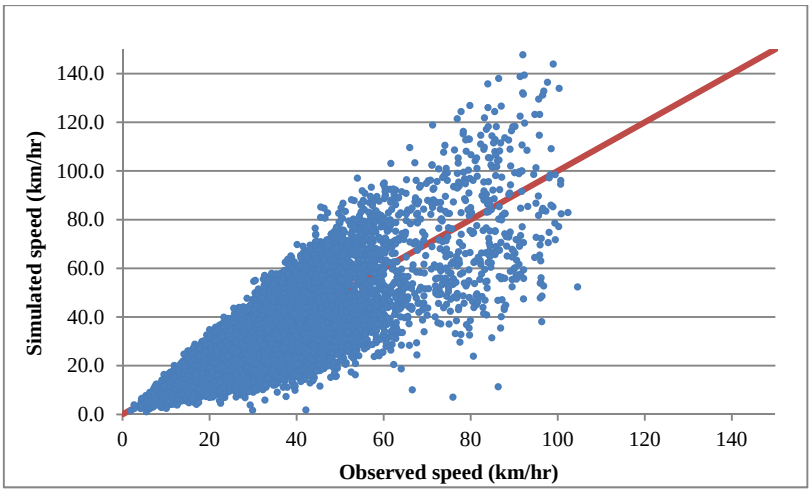


b) base case

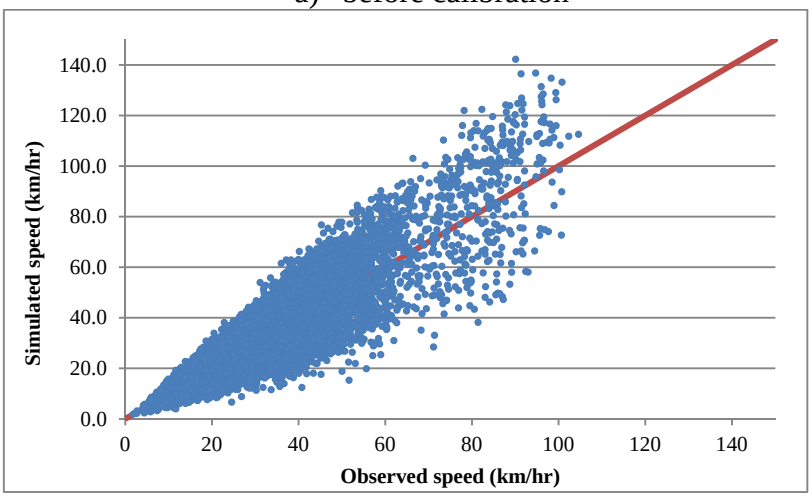


c) multi-source data

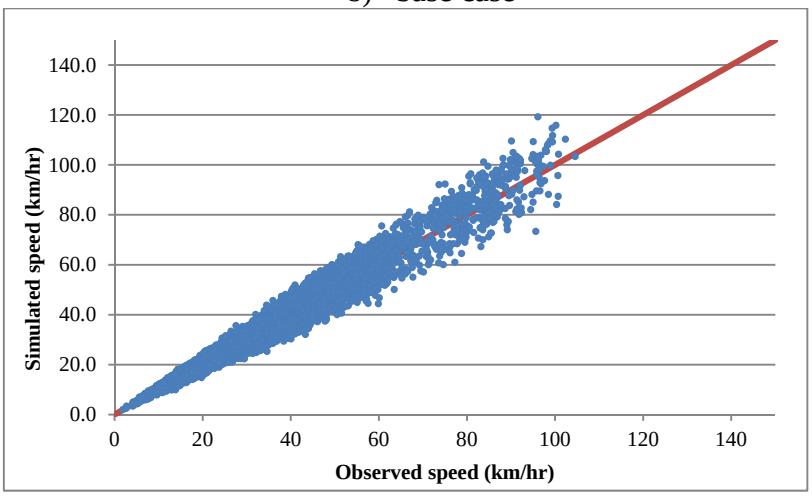
Figure 6.10 Comparison between simulated and observed counts (all three scenarios)



a) before calibration



b) base case



c) multi-source data

Figure 6.11 Comparison between simulated and observed speed values (all three scenarios)

As stated earlier ,the weighting factors, α_i, were used to determine the weight to place on the priori OD flows, link counts, turning counts, and speed values. In the previous case study with synthetic data, a sensitivity analysis was carried out to identify the weights given to different components of the objective function. However, this method was found to be repetitive and computationally intensive, which requires multiple evaluation of the objective function based on different weighting schemes. In addition, the analysis was based on limited combinations of the discrete weighting factors (i.e. 20%, 30%, and 50% in Table 5.8).

Considering the above limitations of the sensitivity analysis as well as the complexity of the Water Front network, the weighting factors were treated as unknown parameters (for both "base case" and "base case + speed data"). In this approach, the optimal values were obtained from the calibration process, based on NRMSE measures of effectiveness. Readers are referred to Section 3.2.2 for the detailed methodology. Table 6.6 presents the optimal weighting factors for the calibration of input parameters.

Table 6.6 Optimal weighting factors, α_i

Scenario	Weighting factors	Optimal values
Base case	α_1: a priori OD flows	0.327
	α_2: traffic counts from loop detectors	0.345
	α_3: turning movements at selected intersections	0.328
Multisource case (base case + speed data)	α_1: a priori OD flows	0.132
	α_2: traffic counts from loop detectors	0.218
	α_3: turning movements at selected intersections	0.198
	α_4: speed data from in-vehicle navigation systems	0.452

The results in Table 6.6 highlight the importance of the low weight given to the historical OD demand term in the objective function and a high weight to speed data derived from in-vehicle navigation devices. In other words, the calibration framework would be more reliant on the speed/travel time data as they are the direct measures of the link performances and highly affect the routing decision. These results show that the presence of reliable and wide spatial coverage of speed data from in-vehicle navigation systems improves the estimation of OD flows and calibration of driver behavior parameters and reduces the dependence on historical OD flows. Thus, it is possible to overcome the high dependency of the OD estimation problem on the historical OD flows and its accuracy.

Considering the above optimal weighting factors, Table 6.7 summarizes the fitness function values (i.e. Z in Equation 10) based on different scenarios. The calibration results in this table confirms the minimum dependency of the calibration process to the historical OD flows as the starting point and the improvement in the fitness function values by incorporating the speed data into the calibration process. In addition, Table 6.8 presents the calibrated driver behavior parameters for 24 time intervals, and separated for AM and PM peak periods (multi-source scenario).

Table 6.7 Optimal fitness function values (DGA)

Scenario	**Weighting factors (α_i)**				**Measure of effectiveness values (f_i)**				**Fitness Function Value (Z)**
	α_1	α_2	α_3	α_4	f_1	f_2	f_3	f_4	$Min\ Z = \sum_{i=1}^{4} \alpha_i f_i$
Base case	0.327	0.345	0.328	0	21.2%	17.2%	19.4%	0	19.2%
Multisource case (base case + speed data)	0.132	0.218	0.198	0.452	29.6%	12.5%	16.8%	10.7%	14.8%

Table 6.8 Calibrated driver behavior parameters (multi-source scenario)

Time interval	Vehicle-driving behavior				Route choice model parameters					
	Mean headway (sec)		Mean reaction time (sec)		Feedback (min)		Familiarity		Perturbation	
	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM
1	0.641	0.732	0.582	0.695	2.8	3.1	73.6%	73.5%	20.5%	19.8%
2	0.645	0.745	0.602	0.715	2.9	3.0	74.5%	76.4%	19.9%	19.5%
3	0.652	0.748	0.584	0.721	2.9	3.3	76.3%	74.5%	19.8%	19.2%
4	0.649	0.739	0.569	0.702	2.9	3.1	77.2%	76.6%	19.9%	18.5%
5	0.638	0.756	0.582	0.709	3.0	3.0	76.0%	77.5%	19.5%	18.9%
6	0.641	0.755	0.599	0.711	3.0	2.9	78.4%	76.5%	18.9%	19.8%
7	0.645	0.745	0.597	0.701	3.4	2.9	79.0%	76.4%	18.8%	18.7%
8	0.656	0.759	0.592	0.685	3.1	3.0	76.9%	75.6%	20.8%	19.4%
9	0.654	0.765	0.545	0.698	3.0	2.9	78.9%	73.8%	20.4%	19.5%
10	0.639	0.768	0.562	0.678	2.9	2.9	80.6%	78.5%	19.2%	19.4%
11	0.664	0.77	0.571	0.674	2.9	2.6	74.5%	72.6%	20.2%	20.3%
12	0.645	0.748	0.592	0.689	3.0	2.8	76.8%	74.8%	21.1%	19.7%
Average	0.647	0.753	0.581	0.698	2.98	2.96	76.9%	75.6%	19.9%	19.4%
Standard deviation	0.008	0.012	0.017	0.015	0.15	0.17	2.1%	1.7%	0.7%	0.5%
Coefficient of variation	1.2%	1.6%	2.9%	2.1%	5.0%	5.8%	2.7%	2.3%	3.7%	2.6%

Based on the final calibrated values, the following conclusions can be made:

- *Temporal variations of the driver behavior parameters in peak periods:* The driver behavior parameters (i.e. vehicle-driving behavior and route choice model parameters) remained relatively unchanged for the 15-minutes intervals of the AM and PM peak periods. For example, the coefficient of variation of the mean headways during the AM peak period is less than 2% (i.e. 1.2%). These findings supported the initial hypothesis that the general driving behavior is not subject to temporal variation in each 15-minutes. Figure 6.12 and Figure 6.13 present the temporal variations of the vehicle-driving behavior and route choice model parameters in the peak periods, respectively.

- *Comparing the driver behavior parameters in AM Peak vs. PM Peak:* As is apparent from the calibrated driver behavior parameters presented in Table 6.8, the first two parameters corresponding to driver behavior parameters are less for the AM peak period than the PM peak period (average mean headway of 0.649 for AM peak vs. 0.753 for PM peak). These findings support a recent research that showed that drivers' headway is subject to temporal variation [106]. One possible explanation is that during AM peak, drivers consist mainly of commuters who need to arrive to their destination on time. Thus, they might be willing to accept or rather force shorter headways when following other cars, merging, turning maneuvers or changing lanes. In addition, during the morning peak, drivers might be more alert than in the afternoon when they might be more tired after a long day of work. This might explain their shorter mean reaction time in the morning. That was also supported by a study that showed that driving performance is affected by time of day [107]. Unfortunately there is a lack of research in human factors to further support such findings. On the other hand, there is no significant difference between the route choice model parameters during AM and PM peak periods.

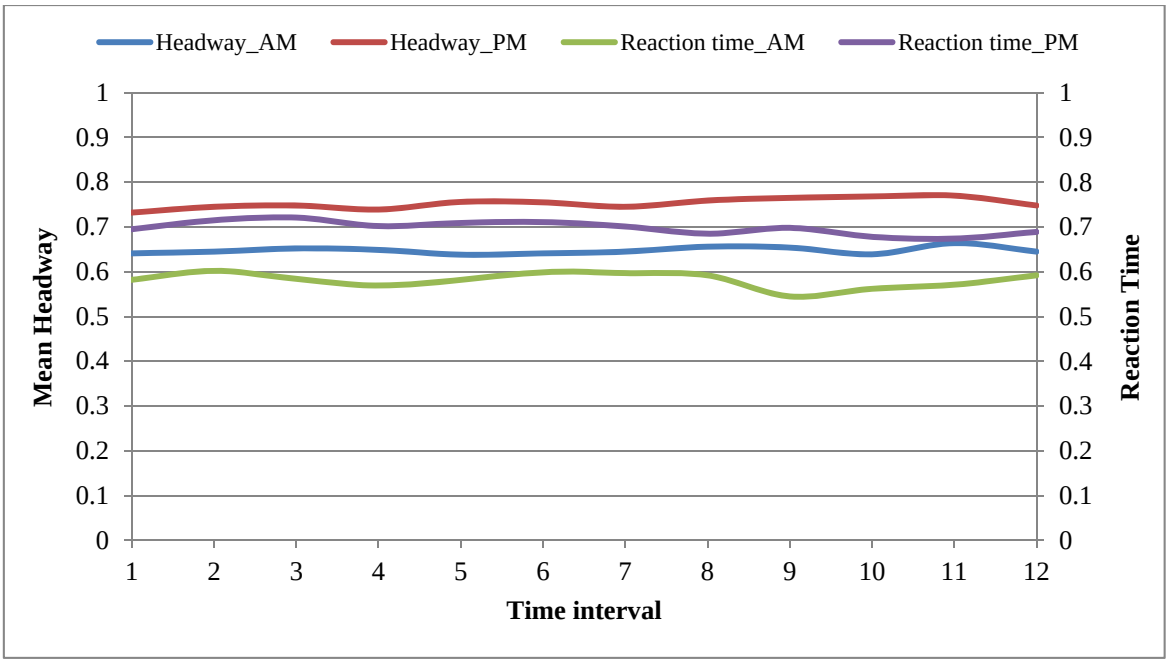


Figure 6.12 Temporal variations of the vehicle-driving behavior during peak periods

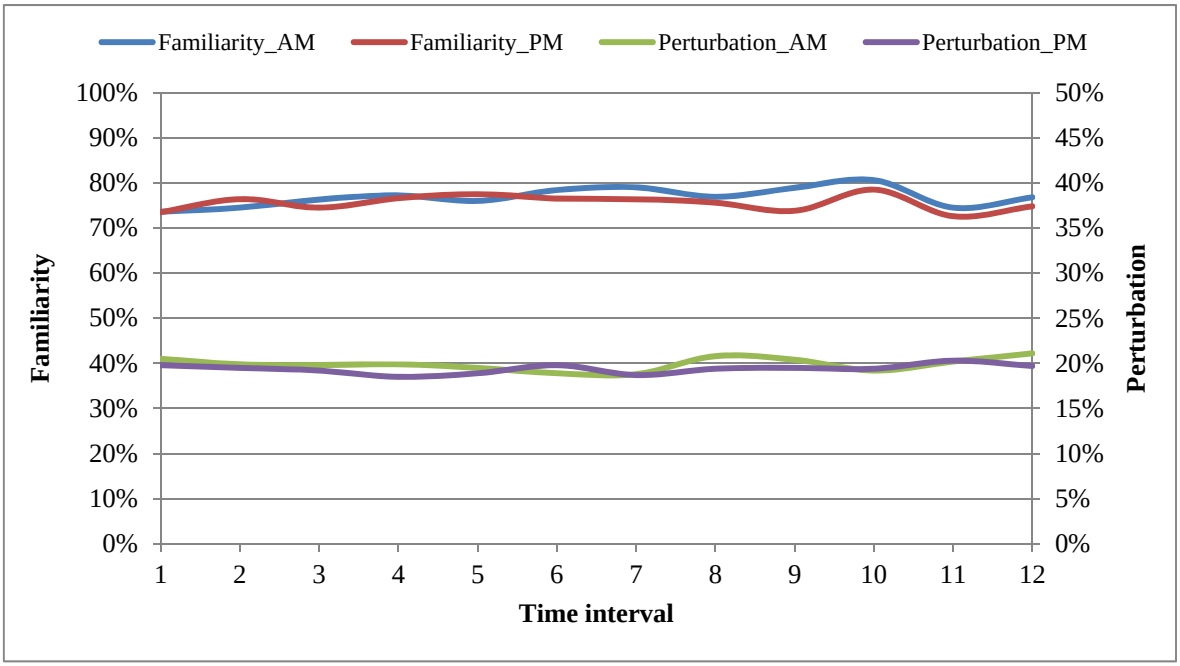


Figure 6.13 Temporal variations of the route choice model parameters during peak periods

As stated above, the vehicle-driving behavior parameters are less for the AM peak period than the PM peak period. In order to evaluate whether these parameters in the AM Peak period are statistically different from the PM peak period, F-test and t-test were conducted. The F-test is used to conduct a hypothesis test for equality of variances of the two samples

(AM vs. PM parameters). Also the t-test is the test for equality of two means regardless the variances are equal or different. Based on F-test, t-test, and 95% confidence interval, the driver behavior parameters were compared for AM and PM peak periods (Table 6.9).

Table 6.9 F-test and t-test results for comparing vehicle driving behavior parameters (multi-source case)

Parameter types	Parameter	F-test			t-test		
		F-value	F critical (95%)	Variance Comparison	t-value	t critical (95%)	Mean Comparison
Vehicle-driving behavior	Mean headway	2.25	2.82	Equal	25.77	2.20	Significantly different
	Mean reaction time	1.34	2.82	Equal	18.02	2.20	Significantly different
Route choice model parameters	Feedback	1.18	2.82	Equal	0.09	2.20	Equal
	Familiarity	1.10	2.82	Equal	1.69	2.20	Equal
	Perturbation	2.05	2.82	Equal	2.05	2.20	Equal

Based on the comparison results for different scenarios and measures of effectiveness, the following conclusions can be made:

- Route choice model parameters in the AM and PM peak periods were not statistically different at a 95% confidence interval (both variance and mean), suggesting that the drivers in the downtown area of Toronto were familiar with the roadway network and the variations in the level of congestion could not significantly affect their decision in choosing a different route for reaching a destination.

- In contrast to the route choice model parameters, it was found that, the average of the driving behavior parameters in the AM and PM peak periods were statistically different at a 95% confidence interval. These findings support the earlier observations

that the vehicle-driving behavior parameters are less for the AM peak period than the PM peak period.

6.9.2 Experiment II: Calibration based on Parallel Distributed GA

The purpose of this experiment is to evaluate the effect of the parallelization scheme of GA population on the quality of the solution by comparing the performance of the DGA with the PDGA. The calibration process was based on the available traffic data from multiple sources (i.e. speed, loop detector counts, and turning movements).

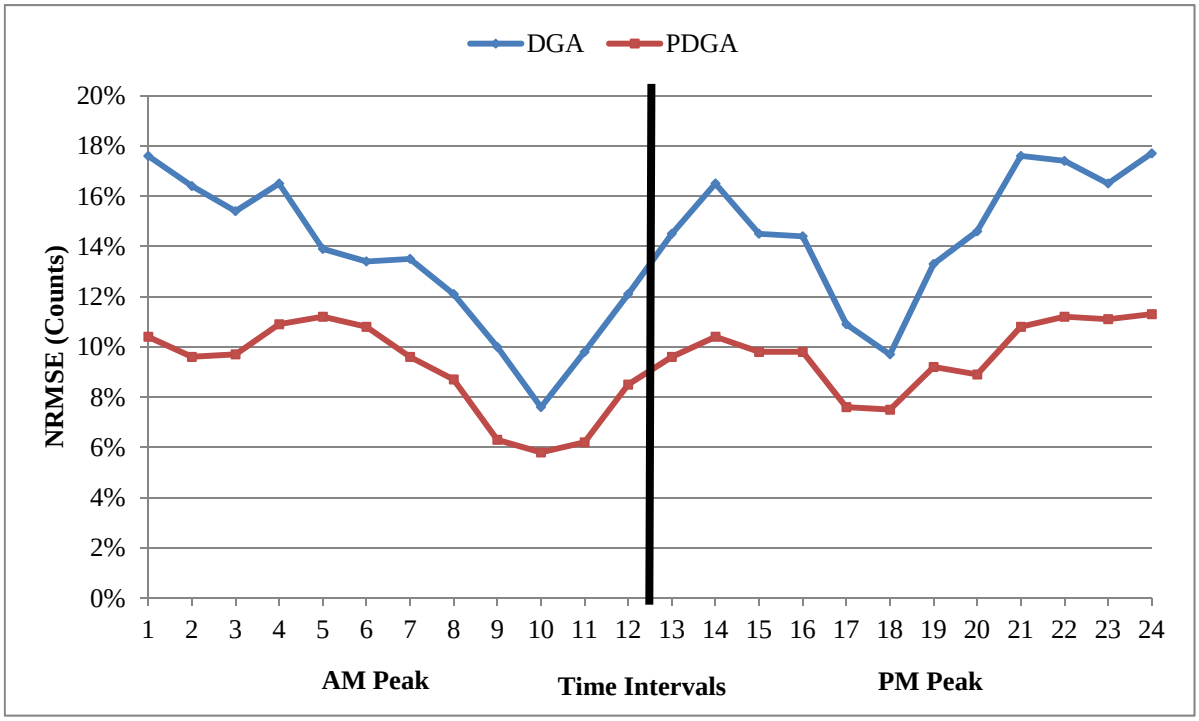
Based on the calibrated model parameters in the previous experiment (Table 22), the driver behavior and route choice model parameter were set to be fixed during the 15 minutes intervals of each peak period. Therefore, the total number of unknown parameters decreased from 16,176 to 16,066, and the degree of freedom would be equal to 3234.

Comparing DGA and PDGA

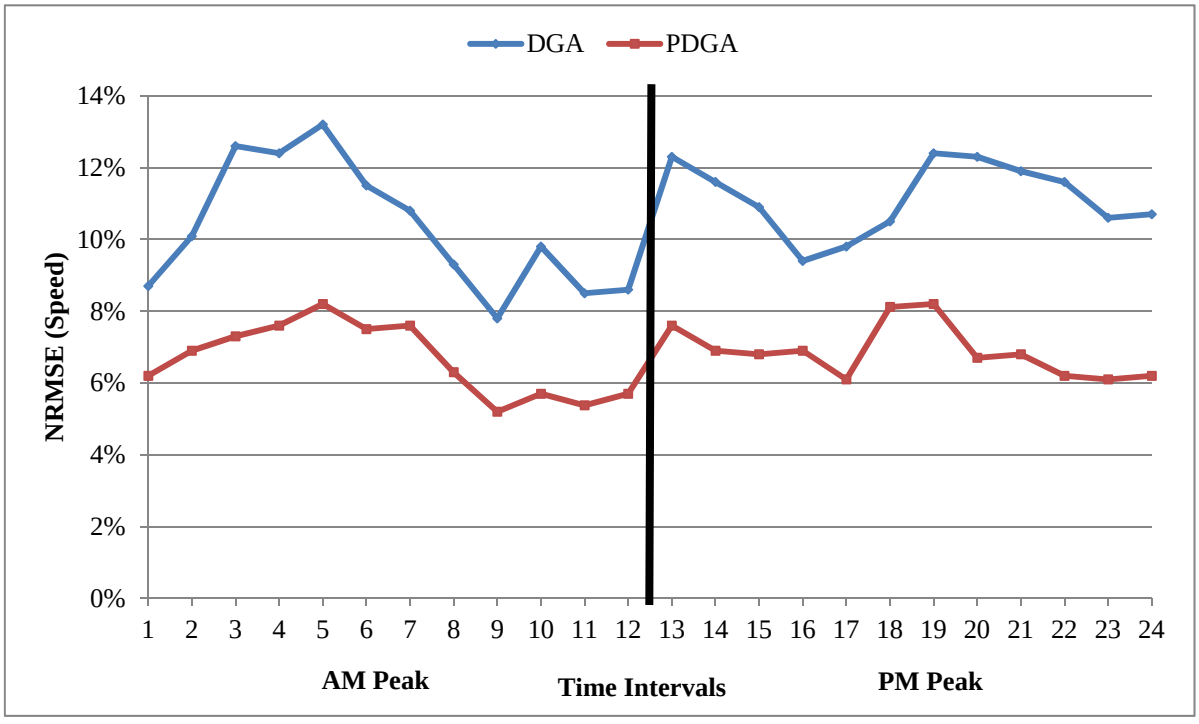
Following the methodology set forth in Section 3.2.2 and the parallel GA control parameters in Table 6.4, the final simulated dynamic OD flows, counts, speed data, and driver behavior parameters were obtained. Table 6.10 presents the calibration result statistics for counts and speed values. In addition, the calibration result statistics for DGA were adopted from Table 6.5 and Table 6.6 and incorporated in Table 6.10 for the direct comparison between DGA and PDGA. For the visual comparison, the count and speed NRMSE values for each 15-minute interval of the analysis period (AM peak and PM peak) are presented in Figure 6.14.

Table 6.10 Comparison between DGA and PDGA calibration result statistics based on counts and speed NRMSE values

Category	DGA	PDGA	Improvement percentage
Count	14%	9.4%	33%
Speed	10.7%	6.8%	37%



c) Count



d) Speed

Figure 6.14 Comparison between measures of effectiveness of DGA and PDGA based on NRMSE for count and speed data

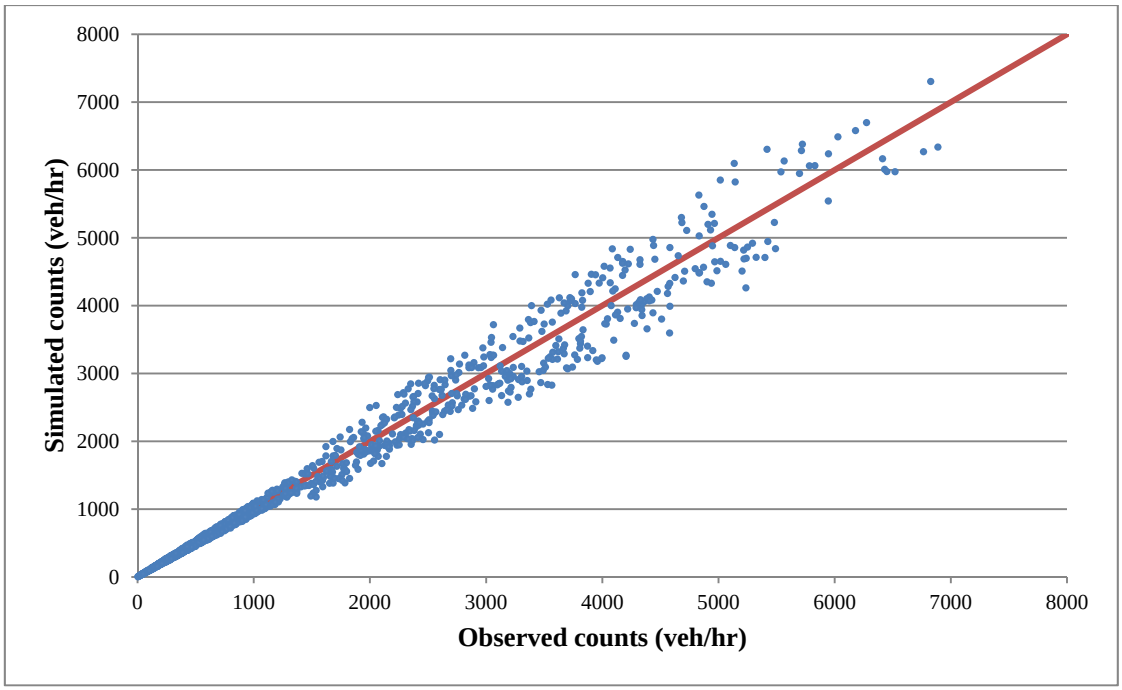
The reported results in Table 6.10 for the Water Front network is in line with the findings of the previous simple case study, suggesting that the parallel GA leads to further reduction in fitness function results, as compared to the single GA running in the parallel processors.

According to Figure 6.14, the improvement in terms of count NRMSE values varies depending on the simulation periods ranging from 19.1% and 41.4%, as compared to the single population GA with an average improvement of 33.2%.

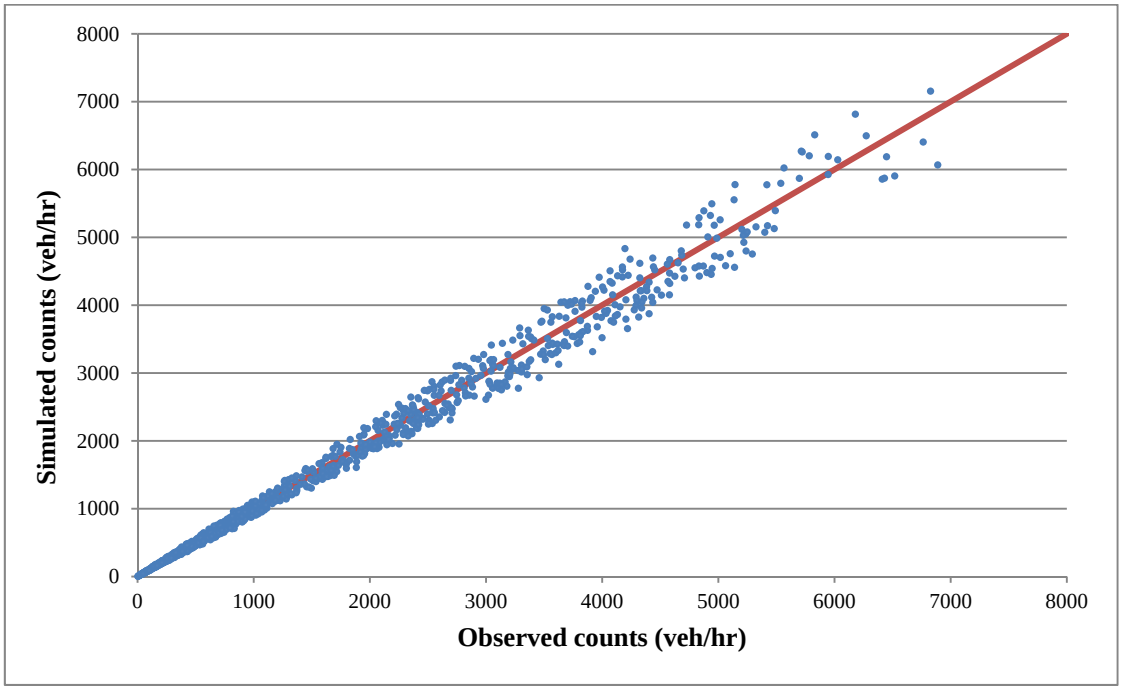
Following the same trend as the count performance measures, Table 6.10 and Figure 6.14 show that parallelization of GA result in an overall reduction of speed performance measures as compared to the DGA. The percentage change of speed NRMSE ranged between 23.4% and 47.1% for various time periods, with an average improvement of 37.7%.

In summary, it was found that the parallelization structure of GA can significantly improve the quality of solution, even for the large-scale network (i.e. 33% and 37% for count and speed NRMSE, respectively), which resulted in the NRMSE of less than 10% for both count and speed measurements.

Figure 6.15 presents the plots of the observed counts and simulated entities for the PDGA and DGA calibration scenarios, using multi-source traffic data. In addition, Figure 6.16 compares the observed and simulated speed values for DGA and PDGA scenarios. A visual comparison between these figures reveals that PDGA significantly outperforms the DGA in terms of NRMSE of both count and speed data Appendix C presents these plots in more details. Figure 6.17 presents the evolution of the fitness function with each GA generations.

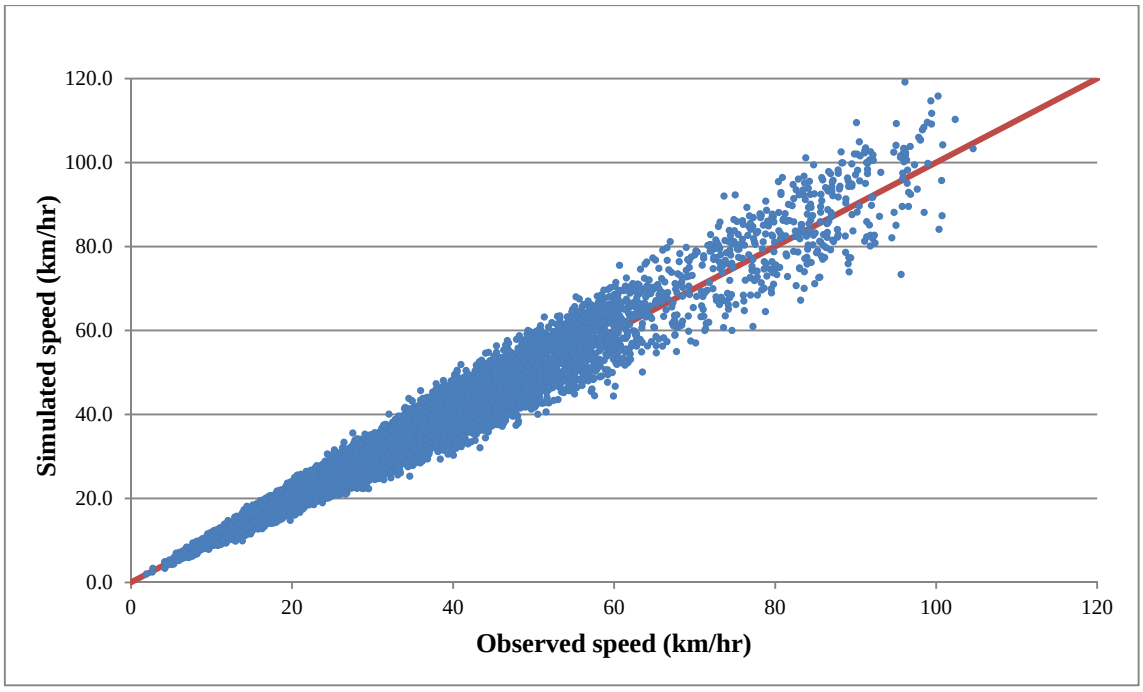


a) DGA

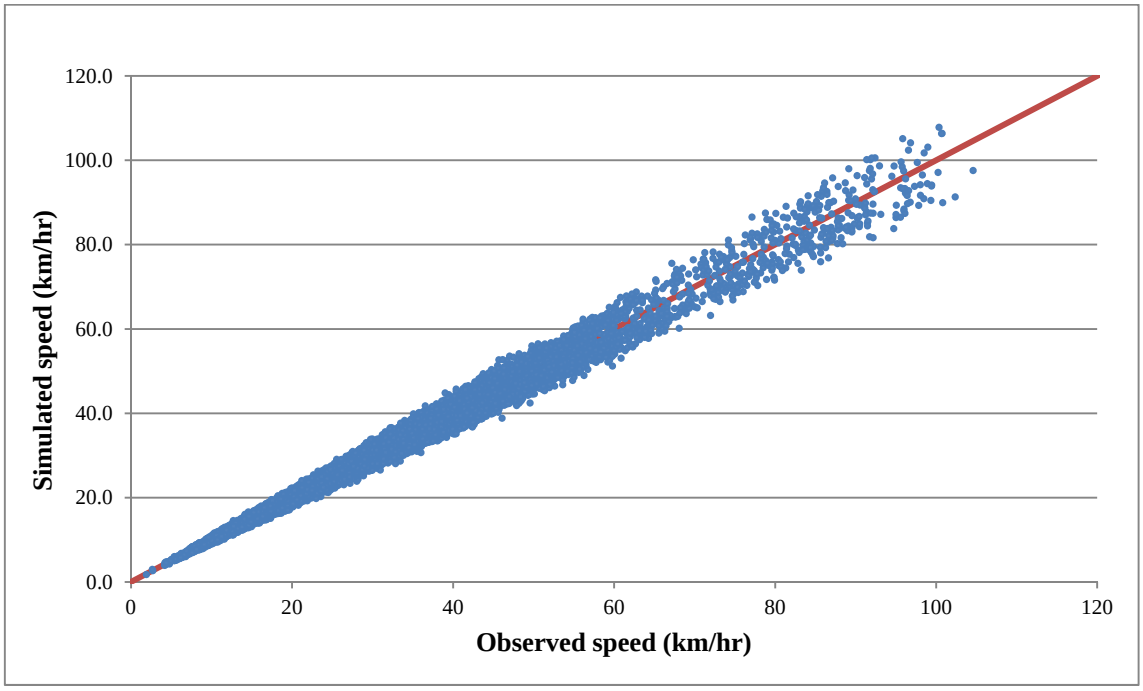


b) PDGA

Figure 6.15 Comparison between simulated and observed counts using multi-source data (DGA vs. PDGA)



a) DGA



b) PDGA

Figure 6.16 Comparison between simulated and observed speed using multi-source data (DGA vs. PDGA)

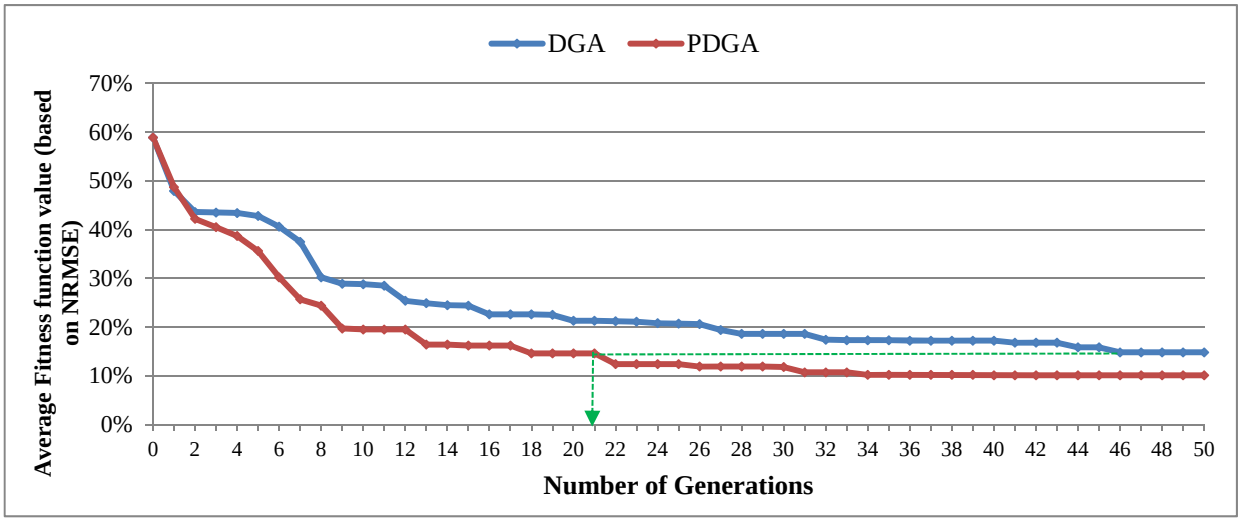


Figure 6.17 Effect of parallelization and distributed computing on GA convergence and quality of solutions (PDGA vs. DGA/SGA)

Based on the evolution of the objective function in different generations, the following observations can be made:

- *Termination point:* As is apparent from Figure 6.17, the PDGA required a fewer number of generations to find the optimal fitness function value comparing to the DGA (i.e. 40 vs. 46). In addition, the PDGA could produce the same quality as the DGA (fitness of 14.8%) in 21 of the number of generations comparing to 46 for the DGA (see dotted line), i.e. the PDGA required approximately half the number of generations to find the optimal fitness function values, comparing to DGA.

- *Quality of solution:* PDGA outperformed the DGA in terms of fitness function value by producing a 31.5% reduction in the fitness function values (fitness of 14.8% vs. 10.1%).

- *Convergence speed:* As noted earlier, the convergence speed of the algorithms were evaluated by the total CPU time (in hours) to reach the global optimal value. While the DGA significantly reduced the computation time comparing to SGA (198 hours vs. estimated 15,000 hours), the multi-deme structure of GA populations did not have any major impact on the convergence speed of the algorithm in terms of total CPU time (i.e. 198 and 214 hours for DGA and PDGA, respectively). While PDGA

requires more communication time between the core and slave processors for creating the necessary files/directories, fewer number of generation was required for convergence of this algorithm. This result was in line with the previous synthetic case study.

Optimal Weighting Factors

As stated earlier in Section 6.9.1, for the Water Front network, the weighting factors were treated as unknown parameters and their optimal values were obtained from the calibration process. Table 6.11 presents the optimal weighting factors for the calibration of input parameters.

Table 6.11 Optimal weighting factors, α_i

Algorithm	Weighting factors	Optimal Values
PDGA	α_1: a priori OD flows	0.103
	α_2: traffic counts from loop detectors	0.187
	α_3: turning movements at selected intersections	0.184
	α_4: speed data from in-vehicle navigation systems	0.526

The results presented in Table 6.11 is consistent with the previous experiment (Table 6.7), which highlight the importance of the low weight to the historical OD demand term in the objective function and a high weight to speed data derived from in-vehicle navigation devices. In other words, the calibration framework would be more reliant on the speed/travel time data as they are the direct measures of the link performances and highly affect the routing decision. These results show that the presence of reliable and wide spatial coverage of speed data from in-vehicle navigation systems improves the estimation of OD flows and calibration of driver behavior parameters and reduces the dependence on historical OD flows. Thus, it is possible to overcome the high dependency of the OD estimation problem on the historical OD flows and its accuracy.

As noted earlier, the PDGA outperformed the DGA in terms of fitness function value by further minimizing the discrepancy between the observed data and their simulated counterparts. The optimal weighting value of the speed data was increase from 0.452 in the previous experiment (i.e. DGA) to 0.526, while producing a lower NRMSE value (i.e. an average improvement of 37.7% from 10.7% to 6.8%).

Considering the above optimal weighting factors, Table 6.12 summarizes the fitness function values (i.e. Z in Equation 10) based on the parallelization structure of GA (i.e. PDGA) comparing to the results of the simple GA running in multiple processors (i.e. DGA). The calibration results in this table confirm that with parallelization of GA structure, the quality of the solution in terms of fitness function values (i.e. Z) can be improved. This was achieved by reducing the NRMSE of each component of the objective function and the changes in the corresponding weighting factors of each component. In summary, the PDGA reduced the total fitness function value by 32%, comparing to the DGA.

Table 6.12 Optimal fitness function values for multi-source scenario (DGA vs. PDGA)

Algorithms	**Weighting factors (α_i)**				**Measure of effectiveness values (f_i)**				**Fitness Function Value (Z)**	**Improvement Percentage (PDGA vs. DGA)**
	α_1	α_2	α_3	α_4	f_1	f_2	f_3	f_4	$Min\ Z = \sum_{i=1}^{4} \alpha_i f_i$	
DGA	0.132	0.218	0.198	0.452	29.6%	12.5%	16.8%	10.7%	14.8%	31.5%
PDGA	0.103	0.187	0.184	0.526	28.5%	9.2%	10.5%	6.8%	10.1%	

Table 6.13 presents the calibrated driver behavior parameters, separated for AM and PM peak periods. As noted earlier in the previous experiment, it was found that the driver behavior and route choice model parameters remained relatively unchanged during the 15 minutes intervals of each peak periods; however, the driver behavior parameters were found to be significantly different between AM Peak and PM Peak. As shown in Table 6.13, the calibrated vehicle-driving behavior parameters are less for the AM peak period than the PM

peak period (i.e. 14.5% less for mean headway and 16.0% for mean reaction time). The result of this experiment was in line with the previous findings, suggesting that drivers were accepting shorter headways during AM peak period comparing to PM peak period. In addition, during the morning Peak, drivers might be more alert than in the afternoon when they might be more tired after a long day of work. This might explain their shorter mean reaction time. Therefore for the purpose of this experiment, driver behavior parameters were set to be fixed during the 15 minutes intervals of each peak period.

On the other hand, the results show no significant difference between the route choice model parameters during AM and PM peak periods (less that 3% for all cases).

Table 6.13 Calibrated driver behavior parameters

Parameter Types	Description	Optimal Value		Difference between AM and PM
		AM	PM	
Vehicle-driving behavior	Mean headway (sec)	0.662	0.774	14.5%
	Mean reaction time (sec)	0.590	0.702	16%
Route choice model parameters	Feedback (min)	2.89	2.82	2.5%
	Familiarity	77.4%	75.7%	2.2%
	Perturbation	18.4%	17.9%	2.8%

Dynamic OD Flows

In conjunction with the driver behavior and route choice model parameters, the off-line dynamic OD flows were estimated for the 24 time intervals on 15-minute each. It should be noted that the simulated OD flows represent the historical trip pattern for the typical weekdays in the Fall season of 2012. Table 6.17 and Table 6.18 present the summary of the estimated hourly OD flows during AM and PM peak hour for the Water Front network, respectively. In addition, the seed OD matrix of 2006 for the subject network is presented in Table 6.16.

Table 6.14 Summary of dynamic OD travel demand during AM peak periods (Fall 2012)

Peak period trips			
6:30 AM – 9:30 AM	To external gateways	To study area zones	**Total**
From external gateways	37,616	19,625	57,242
From study area zones	8,224	678	8,902
Total	45,840	20,303	66,143
Hourly trips			
6:30 AM – 7:30 AM	To external gateways	To study area zones	**Total**
From external gateways	10,260	5,345	15,605
From study area zones	2,667	180	2,848
Total	12,928	5,525	18,453
7:30 AM – 8:30 AM	To external gateways	To study area zones	**Total**
From external gateways	14,763	7,528	22,291
From study area zones	3,038	269	3,307
Total	17,801	7,797	25,598
8:30 AM – 9:30 AM	To external gateways	To study area zones	**Total**
From external gateways	12,593	6,753	19,345
From study area zones	2,519	229	2,747
Total	15,111	6,981	22,093

Table 6.15 Summary of dynamic OD travel demand during PM peak periods

Peak period trips			
3:30 PM – 6:30 PM	To external gateways	To study area zones	**Total**
From external gateways	43,960	9,854	53,814
From study area zones	18,036	805	18,841
Total	61,997	10,658	72,655
Hourly trips			
3:30 PM – 4:30 PM	To external gateways	To study area zones	**Total**
From external gateways	13,319	2,957	16,276
From study area zones	5,555	246	5,801
Total	18,874	3,203	22,077
4:30 PM –5:30 PM	To external gateways	To study area zones	**Total**
From external gateways	16,316	3,715	20,031
From study area zones	6,505	295	6,800
Total	22,821	4,010	26,831
5:30 PM – 6:30 PM	To external gateways	To study area zones	**Total**
From external gateways	14,326	3,181	17,507
From study area zones	5,976	264	6,240
Total	20,302	3,445	23,747

Table 6.16 2006 seed OD matrix for the Water Front network

AM peak hour trips			
7:30 AM – 8:30 AM	To external gateways	To study area zones	**Total**
From external gateways	13,615	7,178	20,793
From study area zones	2,641	251	2,892
Total	16,256	7,429	23,685
PM peak hour trips			
4:30 PM – 5:30 PM	To external gateways	To study area zones	**Total**
From external gateways	15,037	3,148	18,185
From study area zones	5,942	273	6,215
Total	20,979	3,421	24,399

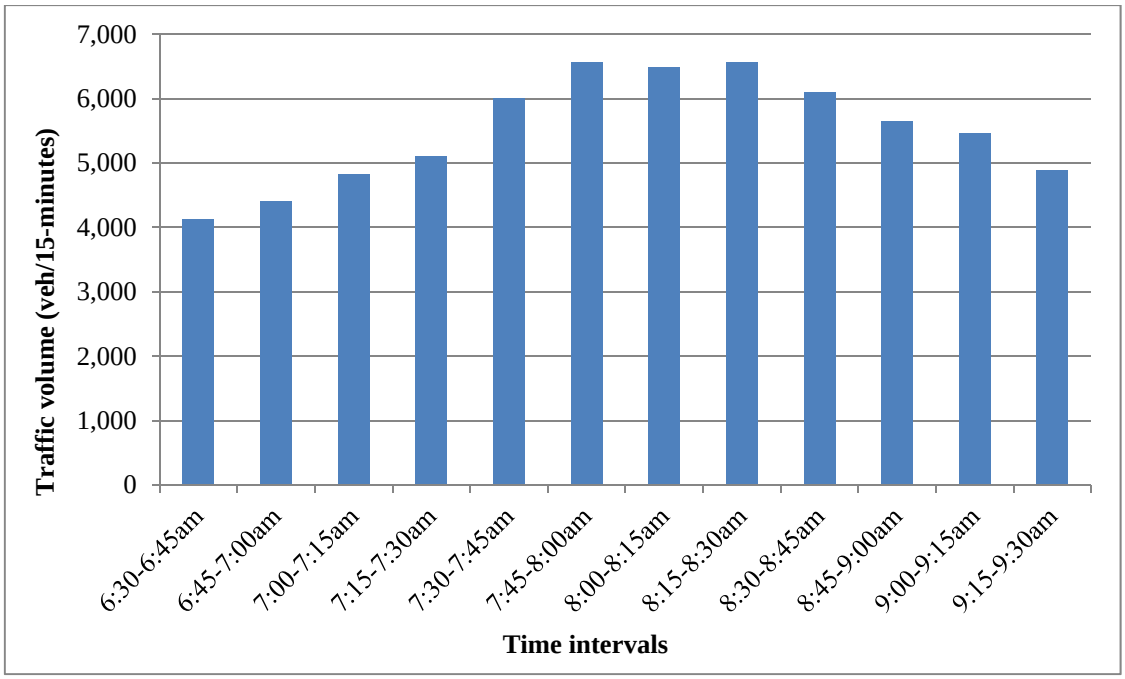
Based on the results presented in the above tables, the following observations can be made:

- *Traffic growth in peak hours*: during the AM and PM peak hours, the total trips from 2006 to 2012 in the subject network were estimated to have increased by 8.1% and 10%, respectively (i.e. from 23,685 to 25,598 for AM peak and from 24,399 to 26,831 for PM peak period). In other words, the average annual growth rate for the AM and PM peak hours were found to be 1.25 and 1.51, respectively.

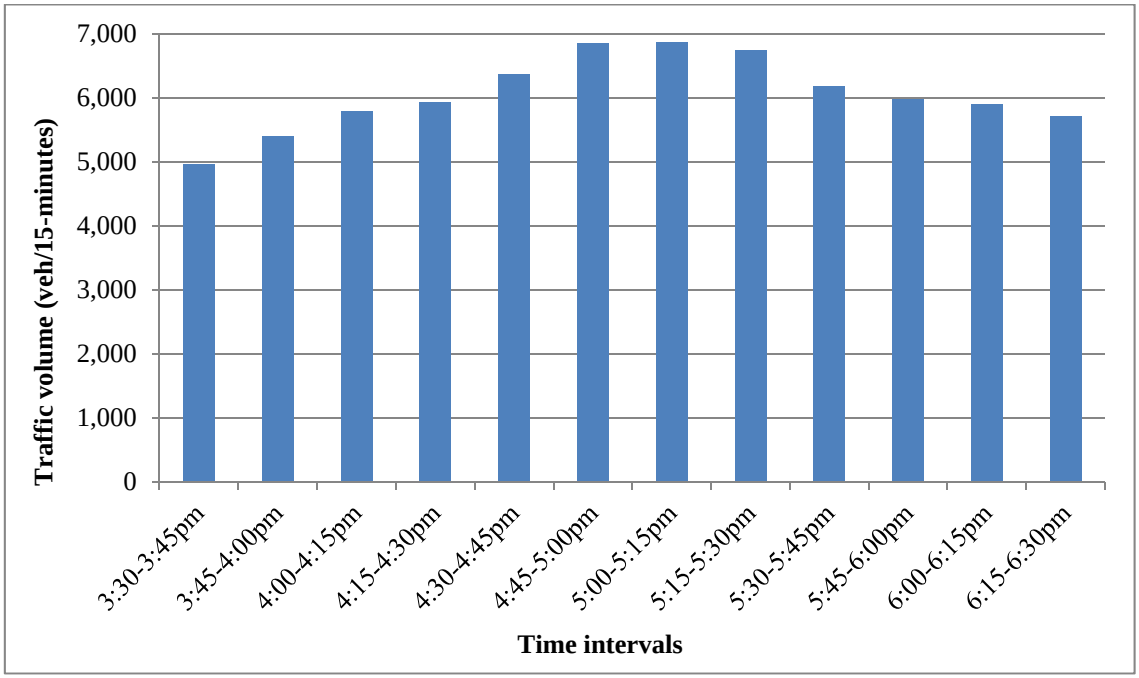
- *Dynamic temporal distribution*: the total trips in the AM and PM peak hours were found to be 38.7% and 36.9% of the total trips in the 3-hour peak period, respectively. In 2006, 40% of the peak period trips occurred in the peak hour, while this ratio for PM peak was estimated to be 38%. The distributions of the total trips in the 3-hour AM peak period were found to be 27.9% (6:30am-7:30am), 38.7% (7:30am-8:30am), and 33.4% (8:30am-9:30am). On the other hand, the distributions of the total trips in

the 3-hour PM peak period were found to be 30.4% (3:30pm-4:30pm), 36.9% (4:30pm-5:30pm), and 32.7% (5:30pm-6:30pm). Overall, the smoother trip distribution of the total traffic during the peak periods might be associated with flexible working hours in recent years.

In addition to the above hourly flows, Figure 6.18 presents the temporal variations of the total trips within each 15-minutes interval. Finally, the summary of the dynamic OD flows for each 15-minute interval is presented in Appendix D.



AM peak period



PM peak period

Figure 6.18 Temporal variations of the 15-minutes dynamic OD flows

6.9.3 Experiment III: Testing the Impact of Link Segmentation

As noted earlier in Section 6.6.4, the private vendor provided the raw speed traffic data was for each link in the GIS map, ranging from 5 meters up to 100 meters. For the past two experiments, the speed data from the GPS in-vehicle navigation systems were aggregated to constitute the interchange-to-interchange or intersection-to-intersection routes in the study area[5]. This step was found to be essential for the calibration of the large-scale Water Front network in terms of computation time and evaluation of the fitness function. However, the aggregation of consecutive link travel time/speed data compromises the accuracy of the observed data and increases the variance of the traffic data. In addition, one of the challenges with the previous experiments was that the covariance of travel time associated with two consecutive links was not known. As a result, variance of a route consisting of a number of smaller links was assumed to be the sum of variances of travel times.

Considering the above limitations and challenges with aggregation of speed data, the purpose of experiment was to investigate the impact of the inclusion of the disaggregated speed data into the calibration process, comparing to the aggregated speed data. The logic behind this experiment is that basing the calibration process on disaggregate speed data reflects important information that would be masked if this data is aggregated. The disaggregate speed data includes valuable information capable to reflect the congestion level, queuing formation, occurrence and location of shockwave and delays on segment by segment basis. Since the disaggregate speed data was extracted from trajectory data of individual drivers, it is also capable to reflect more closely the driver behavior parameter as affected by congestion. In addition, it will add to the observability of the problem by reducing the number of degrees of freedom, as the segments can be broken down into smaller links and more observations will be available. Furthermore, while traffic flow information is not a reliable indicator of congestion, speed data are more reliable indicator of congestion. Speed data can be used easily to derive the flow information from the fundamental flow-speed

[5] Readers are referred to Figure 36 for the definition of links, segments, and routes.

relationship. That the presence of traffic data is believed to be redundant and thus the use of disaggregate speed data would remove the need for traffic counts.

Since each segment in the study area is comprised of multiple smaller links, the calibration of the entire network with the disaggregated speed data would be computationally burdensome. Therefore, a small fraction of the Water Front network was chosen for this experiment.

The comparison of the two scenarios (i.e. with and without aggregation of speed data) is based on the NRMSE of each component of the objective function, their associated weighting factors, and the total objective function value (i.e. Z in Equation 10). For this experiment, the historical OD flows used is identical to the fraction of the initial perturbed OD flows in the previous experiment. The calibration is based on the PDGA with the same SGA and PGA components presented in Table 17. The simulation was set for the AM peak period. It should be noted that the optimization formula only incorporates the traffic data from speed measurements and historical OD flows. In other words, the loop detector counts and turning movements were excluded for this experiment to test whether the unique OD flows can be estimated from the given enriched speed measurements. Figure 6.19 presents the selected subset of the Water Front network for this experiment with available disaggregate, speed measurements for each segment. The network coded in Paramics contains 17 OD pairs, 325 segments comprised to 1462 links, 36 junctions, 255 nodes, and 22.1 km of roadway. The average segment length and link length in this network are approximately 68m and 15m, respectively.

The simulation period took place between 6:30 to 9:30 for the AM peak period with a half an hour warm-up. The observed speed data was acquired for the same time period as the previous experiments (i.e. 15^{th}-19^{th} of October, 2012, aggregated for each 15 minutes interval during the AM peak period). In this experiment and since the results of the previous experiments showed no significant change in the value of both driver behavior and route choice model parameters within a given peak period, these parameters behavior were set to be fixed during the AM peak period.

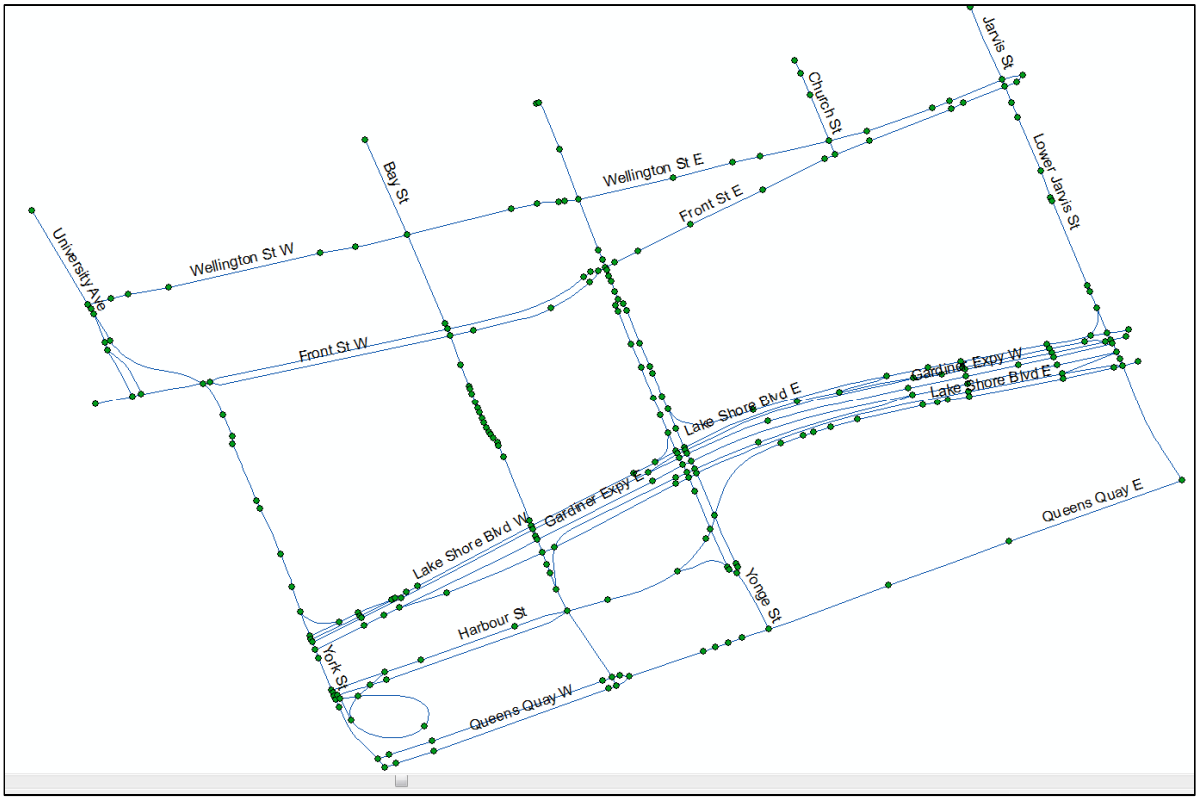


Figure 6.19 Subset of the Water Front network in ArcGIS

As noted earlier, the degree of freedom for the calibration problem would be the discrepancy between the number of unknown parameters and the number of available observed data. Among 325 segments of the small network, the 15-minutes aggregated speed data were available for 92 segments in the study area. Therefore for the 1st scenario (i.e. aggregated speed data), the total available speed data was 92×12=1,104. Considering that the number of unknown OD flows is 272, the total number of unknown parameters would be $(272 \times 12) + 5 = 3{,}269$. Therefore, the degree of freedom for the calibration process is equal to $3{,}269 - 1{,}104 = 2{,}165$.

As stated earlier, the traffic data was available for 92 segments in the subject network. These 92 segments consist of 349 smaller links. Among the 349 links, there were many links with zero number of observations for the 12 time intervals of the peak periods. This was identified as one of the limitations of the data provider, which is due to an inaccurate and/or missing signal of the GPS devices in the urban canyon (i.e. high-rise buildings in Downtown Toronto). In these cases, the average speed data were obtained from adjacent *links* to acquire

the speed data for each *segment*. In addition, the vendor requires a six-month waiting time to provide an enrich traffic data with an acceptable number of observations, while the traffic data was acquired after a three-month waiting period.

Based on the above noted reasons, the total available speed data was 2,424 (i.e. 58% coverage of the 349 links during 12 time intervals). The degree of freedom for the 2nd scenario would be equal to $3{,}269 - 2{,}424 = 845$. Therefore, the incorporation of the disaggregated traffic data into the calibration process significantly decreased the degree of freedom.

Following the methodology set forth in Section 3.2.2, the final simulated dynamic OD flows, speed data and driver behavior parameters were obtained for both scenarios. Table 6.17 compares the fitness function values (i.e. Z in Equation 10) of the above noted scenarios. For the visual comparison, the speed NRMSE values for each 15-minute interval of the AM peak period are presented in Figure 6.20. In addition, Table 6.18 summarizes the calibrated driver behavior parameters of the two scenarios along with the calibrated results from Experiment II.

Table 6.17 Optimal fitness function values for two scenarios

Scenario	Description	Weighting factors (α_i)		NRMSE values (f_i)		Fitness Function Value (Z)	Improvement Percentage (scenario 2 over scenario 1)
		α_1 (a priori OD flows)	α_4 (speed data)	f_1	f_4	$Min\ Z = \alpha_1 f_1 + \alpha_4 f_4$	
1	**Aggregated speed data**	0.124	0.876	25.4%	7.6%	9.80%	24.08%
2	**Raw speed data**	0.098	0.902	22.6%	5.8%	7.44%	

Table 6.18 Comparing driver behavior parameters of the two scenarios with previous experiment

Parameter Types	Description	Experiment III: Subset of the Water Front network		Experiment II (AM peak): Water Front network
		Scenario 1: Aggregated speed data	Scenario 2: Raw speed data	
Vehicle-driving behavior	Mean headway (sec)	0.672	0.676	0.662
	Mean reaction time (sec)	0.602	0.611	0.590
Route choice model parameters	Feedback (min)	2.77	2.78	2.89
	Familiarity	78.3%	78.8%	77.4%
	Perturbation	18.6%	18.9%	18.4%

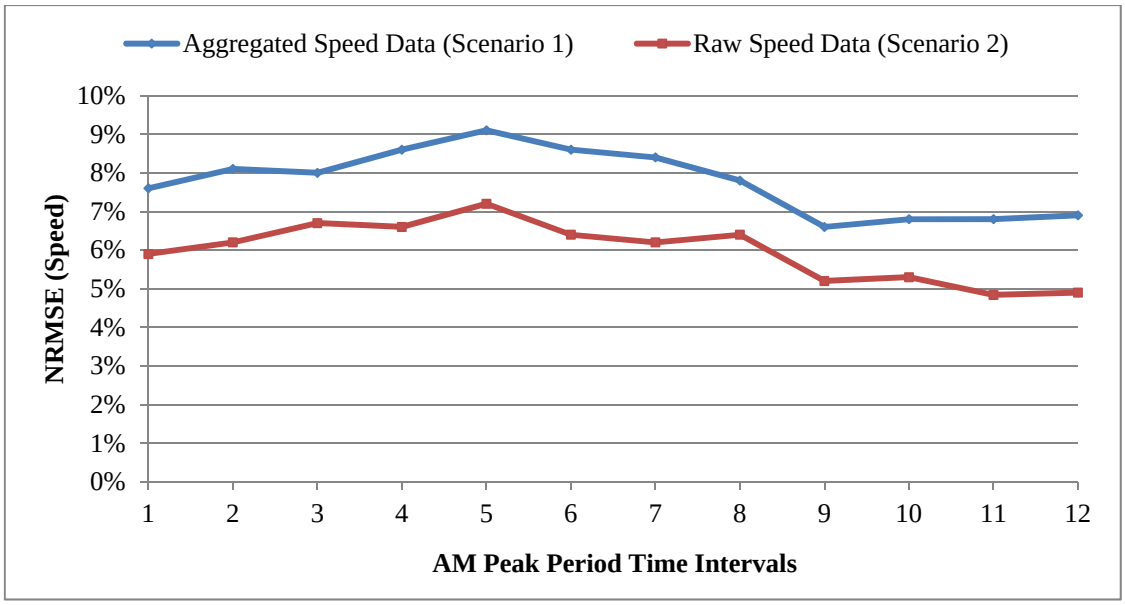


Figure 6.20 Comparison between NRMSE of speed data for the two scenarios

Based on the calibration results presented in Table 6.17, Table 6.18, and Figure 6.20, the following observations were made:

- For both scenarios, the low weight of the historical OD demand term in the objective function and a high weight to speed data were found to be consistent with the previous experiments, which highlights the importance of the quality and spatial distribution of the observed data for the calibration process.

- The incorporation of the disaggregated speed data into the calibration process reduced the NRMSE of speed data from the average of 7.6% for the 1st scenario to 5.8% for the 2nd scenario (i.e. 23.7% improvement).

- The weighting factor of the speed data for the 2nd scenario is higher than the weighting factor for the 1st scenario (0.902 vs. 0.876).

- The total objective function value for the 2nd scenario is less than the objective function value for the 1st scenario (9.80% vs. 7.44%).

- As is apparent from Table 6.18, the calibrated driver behaviour and route choice model parameters of the two scenarios are very similar. In other words, there is no

evidence to support that the incorporation of the disaggregated speed data into the calibration process has a significant impact on the final driver behavior parameters. In other words, the major contribution of this experiment was to evaluate the dependency of the calibration process to the historical OD flows, without having a significant impact on the final results (i.e. driver behavior parameters and OD flows).

- As expected, the calibrated driver behavior parameters of the two Experiments (Experiment II and III) are very close, as the network for segmentation was a subset of the Water Front network.

In summary, the incorporation of the enriched raw speed data without aggregation into the calibration process improved the calibration accuracy by reducing the NRMSE of each component of the objective function and the changes in the corresponding weighting factors of each component.

The results of this experiment was as expected since a) the segments were broken down into smaller links and the calibration process was based on the discrepancy between the observed raw speed data and their simulated counterparts, b) the degree of freedom for the 2nd scenario was significantly less than the 1st scenario (845 vs. 2,165). This experiment highlighted the significant impact of the aggregation level of the observed data on the calibration results and made a further progress towards the application of the "observability", which allowed the modeler to test if the dynamic OD flows can be estimated from the high quality speed data with an acceptable coverage of the network.

While the incorporation of the disaggregated speed data improved the quality of the solution, the convergence speed of the two scenarios was found to be similar. In terms of CPU time, the final OD flows and model parameters for the 1st scenario were obtained after 65 hours, while for the 2nd scenario the calibration process was completed after 71 hours. For the 2nd scenario, the elapsed time (T_E) and communication time (T_C) for creating the files/directories necessary for each processor in the available slave list would be greater than the 1st scenario because of the significant increase in the number of observations. However, the degree of freedom for the 2nd scenario is significantly less than the 1st scenario, which resulted in a fewer number of generation for convergence of the algorithm (30 for 2nd scenario vs. 38 for

the 1st scenario). Figure 6.21 presents the evolution of the fitness function with each GA generations.

In summary, the results of this experiment revealed that the incorporation of the disaggregated speed data for smaller links in a mediums-sized network can improve the calibration accuracy (24.08%) without having a major impact on the computational time for the subject network. As noted earlier, the major challenges with the incorporation of disaggregated speed data into the calibration process would be the extensive modifications on the encoded network in a simulation environment and in order to ensure the consistency with the acquired speed data, especially for a large-scale complex network.

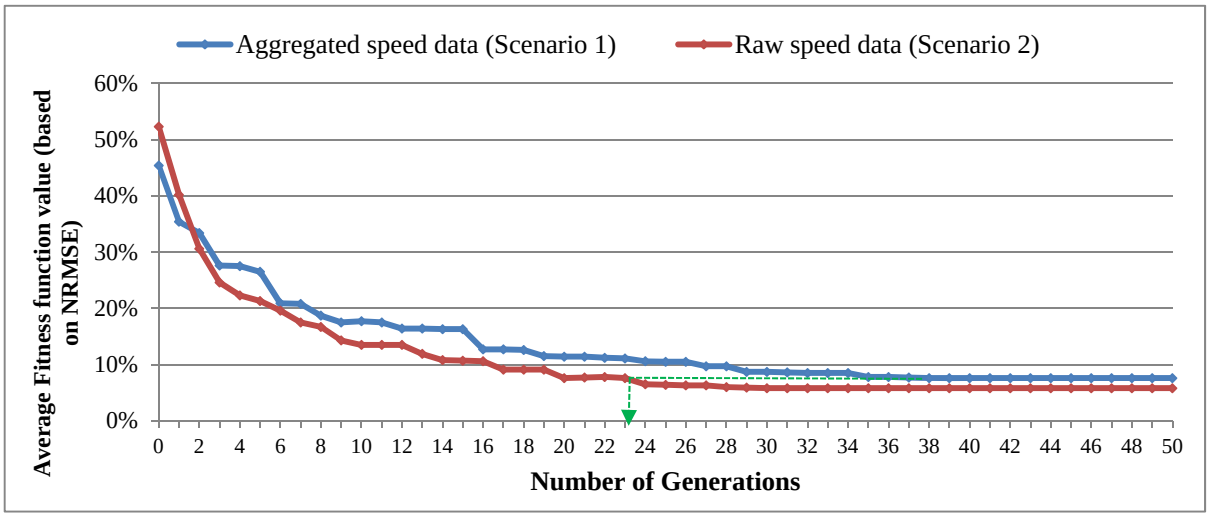


Figure 6.21 Comparing the evolution of the fitness function of the two scenarios

6.10 Summary

The purpose of this chapter was to demonstrate the feasibility of the proposed calibration approach in a large-scale network using the traffic data from various sources. Special considerations were given to the enriched speed data from in-vehicle navigation systems to improve the calibration accuracy and minimize the discrepancy between the observed traffic data and their simulated counterparts. Several simulation experiments were performed to achieve the objectives of this chapter. In summary, the following observations were made:

- The inclusions of speed data into calibration process significantly improved the quality of the solution in terms of count and speed NRMSE comparing to the “base case” scenario (i.e. loop detector and turning movement counts).

- The high network coverage of speed data from in-vehicle navigation systems was shown to be capable in decreasing the dependency of the estimation on historical OD flows while improving the calibration accuracy in terms of the pre-defined measures of effectiveness.

- The driver behavior parameters (i.e. vehicle-driving behavior and route choice model parameters) remained relatively unchanged for the 15-minutes intervals of the AM and PM peak periods. However, the vehicle-driving behavior parameters (i.e. mean headway and reaction time were found to be significantly different from AM peak period to PM peak period.

- Parallelization structure of GA (PDGA) significantly improved the quality of solution, comparing to DGA, even for the large-scale network. However, the total CPU time for convergence of the two algorithms was not significantly different.

- PDGA required a fewer number of generations to find the optimal fitness function value comparing to the DGA.

- The incorporation of the enriched raw speed data without aggregation into the calibration process improved the calibration accuracy and reduced the dependency of the calibration process to the historical OD flows by 24%, without having a major impact on the computational time for a medium-sized network.

Chapter 7: Conclusions and Recommendations

This final chapter concludes the thesis with a summary of the context and scope of this research, the findings of the case studies, the major contributions of this thesis, and some future research directions.

7.1 Context and Scope of Research

During the last decade, many studies have focused on the calibration of demand and supply parameters of dynamic traffic assignment (DTA) models, in order to replicate prevailing traffic conditions and drivers' behaviors. In the microscopic context, the demand variables are the time-dependent origin-destination (OD) flows for the period of interest, while the driver behavior and route choice model parameters are considered as supply parameters. Review of the previous studies revealed the following gaps in the context of calibration of DTA model parameters:

- The majority of the earlier calibration efforts rely on the iterative sequential calibration of demand and supply model parameters. However, this approach ignores the presence of interaction among the various demand and supply parameters and may lead to suboptimal solutions.
- Recent simultaneous calibration frameworks of DTA model parameters are highly dependent on the quality of historical information (e.g. historical OD flows, historical driver behavior parameters). Therefore, these techniques are mostly applicable when the starting point is close to the optimal one.
- Traditional calibration frameworks proposed in most previous studies ignore the reliability of the different components of the objective function (e.g. historical OD flows, count data from loop detectors). A limited number of studies have considered trial-and-error methodology with limited combinations of weights given to different components of the objective function (i.e. manual adjustments).
- Most previous efforts focused on the loop detector data available at the aggregate level, with traffic data collection mostly relying on surveys and vehicle counts that are costly and time-consuming. They have, therefore, been applied infrequently and often do not capture the full range of demand and supply patterns of the network. The

extensive deployment of emerging wireless technologies have resulted in the collection and archiving of time-varying traffic data at the network level and across multiple days, providing enrich datasets for the calibration of complex DTA models.

Based on the above-noted gaps in the context of DTA calibration, this study has focused on the simultaneous calibration of demand and supply parameters of the DTA model in a microscopic context that can capture the interactions between all demand and supply parameters. The calibration process has been formulated as a multi-objective optimization problem that incorporates the traffic data from multiple sources and allocates relative weights to different terms of the objective function. One of the benefits of the proposed calibration framework is that it can incorporate the traffic data from these new sources (upon availability) and significantly improve calibration accuracy. In addition, the proposed framework is flexible and can incorporate any other measure of reliability into the calibration process.

A genetic algorithm (GA) was selected as a suitable solution algorithm for the resulting nonlinear stochastic optimization problem. The quality of the solution and convergence speed of a GA is further enhanced by dividing the GA population into multiple demes and running the GA on a high-performance computer (HPC) cluster with multiple processors.

For this research, traffic data for a real-world network were available from three different sources: loop detector counts, turning movement counts at signalized intersections, and speed data from in-vehicle navigation system technology. The roles of the enriched speed data from in-vehicle navigation system technology are improved calibration accuracy and the application of the observability in order to minimize the dependency of the calibration process on historical OD flows.

The proposed methodology was applied in a synthetic case study as well as a complex real-world network in the business district core of downtown Toronto, Ontario, Canada. The results obtained from the case studies are briefly described in the next section.

7.2 Summary of Findings

Several conclusions are drawn from the experiments performed for the network of the synthetic case study, and are summarized below:

- In the various experiments conducted, the simultaneous calibration process using GA revealed a remarkable improvement over the starting values, in terms of calibration accuracy. In addition, the incorporation of the AVI data into the calibration process further enhanced the quality of solution by further minimizing the objective function. It can be argued that speed information, in contrast to flow measurements, can clearly distinguish between congested or uncongested conditions. Thus, the incorporation of AVI readings improves the solution quality by decreasing the number of local point solutions.
- In the tests conducted, a superior quality solution and savings in computation costs and resources were achieved for the distributed computing and hybrid runs. The distributed GA (DGA) model was 10 times as efficient as the simple GA (SGA). The parallel distributed GA (PDGA) had a better performance than the DGA in terms of fitness function value (26.9%) and termination point (i.e. required number of generations). However, because of its requirement for parallel processing, the convergence speed of PDGA in terms of CPU running time was an hour more than the DGA (9 vs. 11 hours).
- Experiment results revealed the importance of the weights of the objective function and their impact on the solution quality. A weighting scheme of (0.2, 0.3, 0.5) provided the best results for the apriori demand, traffic counts, and speed values respectively. Therefore, it was found that more weight can be given to the speed/travel time data from AVI sensors as the direct measures of the link performance, which can affect the routing decision, comparing to the traffic counts from inductive loop detectors.

Considering the outcomes of the synthetic network, different experiments were performed for the real-world complex network to test the performance of the proposed approach with a larger number of variables and where drivers had more sophisticated route choice alternatives

to reach their destination. Several conclusions are drawn from the experiments performed for the Water Front network in Toronto, and are summarized below:

- The proposed calibration approach was successfully transferable to large-scale network. The incorporation of the in-vehicle navigation system speed data significantly improved the calibration accuracy, in terms of normalized root mean square error (NRMSE) for both count and speed measurements. In addition, a significantly higher quality of solutions and lower computation times were achieved by augmentation of GA with efficiency-enhancement techniques of parallelization and distributed computing.
- In the experiment examining the simulation supply model parameters, it was found that general driving behavior parameters were not subject to temporal variation in each 15-minute intervals of the peak periods. However, the mean headways and reaction times were found to be lower for the AM peak period compared to those of the PM peak period.
- The incorporation of the speed data from in-vehicle navigation system technology significantly reduced the dependency of the calibration process to historical OD flows. In the test conducted with more refined disaggregated speed data for smaller links, it was found that the dependency of the calibration process to the apriori OD information can be reduced to less than 10% for a subset of the large-scale network.

7.3 Research Contributions

As noted earlier in the chapter, this research identified a number of gaps in the context of calibration of DTA model parameters. Previous works on DTA-based dynamic OD estimation and calibration of model parameters were highly dependent on the quality of historical information, especially dynamic OD flows. Therefore, in the absence of such information, the estimated OD flows and model parameters cannot truly replicate the current traffic conditions. In addition, the weighting factors for different components of the objective function were identified through manual adjustments and limited sensitivity analyses. However, such approaches were found to be computationally intensive and inaccurate. Finally, considering the above-noted limitation, the majority of the past studies only

considered loop detector counts for calibration of the DTA model parameters on small-sized or medium-sized networks for one hour simulation period.

The outcome of this research is a step towards filling in the gaps in these areas. In summary, the primary contributions of this research are as follows:

General simultaneous calibration framework

In contrast to previous DTA calibration framework which required manual adjustments, this thesis addresses the more general case of multi-objective optimization problem that simultaneously captures the non-linear interactions between demand and supply parameters, uses model outputs to directly capture complex relationships between the traffic data and model parameters, and incorporates different types of traffic data into the calibration process without any limitations. As is shown in various experiments, this resulted in a more complete representation of the state of the traffic network by taking advantage of all surveillance data and, thus, reducing the possibility of suboptimal solutions. Additionally, this research defines a mathematical formulation to quantify the "reliability" of the observed measurements and estimated the optimal weighting factors though the calibration process. Thus, this research proposes a novel way to automate the calibration framework which significantly reduces the computation costs and improves the accuracy and quality of the final solutions. It should be noted that proposed framework is flexible to incorporate any other measure of reliability into the calibration process. Furthermore, constraints can be added to delimit the size of a feasible search space.

Off-line calibration framework

One of the major contributions of this thesis is the creation of reliable off-line dynamic OD flows and Paramics model parameters that can be used as a priori estimates for the on-line calibration process. In other words, the outputs of the off-line calibration process can be used for real-time OD estimation along key corridors, incident management, reduction in unexpected congestion and, ultimately, provision of real-time traffic data to travelers.

Application of advanced GA to the calibration process

The thesis proposes an innovative application of GA methods and their advanced features to the problem of DTA model calibration in a large-scale network. The application of GA with parallel computing and multi deme architecture results in a superior quality solution and savings in computation costs and resources for all conducted experiments.

Minimum dependency to historical information by incorporation of enriched speed data

This research addresses the problem of a DTA-based approach with inaccurate apriori information. In fact, conventional DTA-based calibration approaches place high expectations on the quality of apriori information. Therefore, improvement over the starting point is usually limited as the adopted optimization techniques are likely to find the nearest local optimum. However, the proposed GA techniques are multipoint searching methods, having global and probabilistic search capabilities. Additionally, this research incorporates the enriched speed data from in-vehicle navigation systems into the calibration process and demonstrates the significant impact of such data on the quality of solution. In fact, this thesis addresses one of the biggest challenges of the OD estimation and calibration of model parameters by minimizing the impact of apriori OD information to less than 10% for a large-scale network.

Testing the impact of link segmentation

This research further improves the calibration accuracy with the incorporation of the enriched raw speed data without aggregation for smaller links into the calibration process. This further reduced dependency of the calibration process on historical OD flows in a medium-sized network.

Temporal variation of driver behavior parameters

While the OD estimation problem was extensively reviewed in the past, very few studies have investigated the effect of time of day on the driving performance. This research takes a further step towards analyzing the temporal variations of the driving behavior of travelers, especially during different time intervals of peak periods. The evaluation of driver behavior

parameters reveal that mean headway is less for the AM peak period than the PM peak period, suggesting that drivers are willing to accept or rather force shorter headways. In addition, the mean reaction time of drivers is found to be less for the AM peak period than the PM peak period, as the drivers are more alert in the morning than in the afternoon when they might be more tired after a long day of work. On the other hand, there is no significant difference between the route choice model parameters during AM and PM peak periods.

7.4 Research Challenges

This section of the thesis highlights some of the challenges and lessons learned during the course of this research, summarized as follows:

- *Segment definition:* Discrepancies in the defined segments between the acquired speed data and the coded network in the simulator created some issues and challenges when mapping the speed data to the Water Front Network. For example, the aggregation of consecutive link travel time/speed data compromised the accuracy of the observed data and increased the variance of the traffic data.

- *Level of aggregation of the speed data:* The quality of the solution and the dependency of the calibration process to the historical OD flows were limited by the aggregation of purchased data, provided spatially on an interchange-to-interchange basis and temporally on a 15-minutes basis. A finer aggregation of data (i.e. links) and more frequent temporal groupings (e.g. 5 minutes) will significantly improve the calibration accuracy and create a more refined dynamic OD matrix.

- *Covariance of segments:* One of the challenges with this research was that the covariance of travel time associated with two consecutive links was not known. As a result, variance of a segment consisting of a number of smaller links was assumed to be the sum of variances of travel times.

- *Number of available processors:* One of the limitations of this research was with the available number of processors in the HPC to run the simulation in parallel. Among the total 265 number of processors in HPC, only 80 processors were made available for this research, and among those, a number of CPUs in the HPC became unavailable

(non-responsive), therefore the simulation models were run in other available CPUs with Paramics licenses.

- Considering the limited number of available processors, the total efficiency of the cluster was approximately 49%. Upon the full availability of the cluster processors (i.e. 80), it is expected that the efficiency of the Cluster significantly increases to 90% [91][6]. Therefore, in an ideal situation where all the processing nodes can be assigned to the problem, the DGA can be more than 70 times faster than SGA

7.5 Future Research Directions

While this thesis contributes to the literature on the calibration of microscopic traffic simulation models in several ways, there are a number of topics for future research:

- In this research, two measures of effectiveness (i.e. the Geoffrey E. Havers statistic, GEH, and normalized root mean square error, NRMSE) were selected to quantify the relationship between the observed and simulated measurements for the synthetic network. Due to computational burdens, only one of them (i.e. NRMSE) was selected for the real-world complex network. More research regarding the use of different measures of effectiveness is necessary.
- In this research, a mathematical formulation was obtained from the literature to quantify the "reliability" of the observed measurements, based on the average and variance of both observed and simulated data. However, there is a lack of research and experimental studies in the definition of different measures of reliability and their combinations. Therefore, it is recommended the proposed weighting schema be validated with different measurements and the final calibration results compared based on different weighting formulas.
- The off-line calibration process was successfully demonstrated in a large-scale network in downtown Toronto, Canada, with a combination of both highways and

[6] Abd, H. M. A. E. H. (2010). *Optimization of Multimodal Evacuation of Large-scale Transportation Networks*, Doctoral dissertation, University of Toronto

arterials. The calibrated OD and Paramics model parameters can be used as the most recent sources of data for calibration of the entire network of downtown Toronto, based on a more refined and powerful HPC.

- One of the contributions of this research was that the enriched speed data from in-vehicle navigation system technology minimized the dependency of the calibration process on historical OD flows. It can be argued that the incorporation of traffic data from other reliable resources (e.g. traffic data from Bluetooth receivers) with the current disaggregated speed data may fully diminish the dependency of OD estimation on historical information. Considering emerging technologies and growing interest of different municipalities and jurisdictions in low-cost and accurate traffic data acquisition, it is possible to further expand this research and estimate the unique *true* dynamic OD flows without a priori information.
- As for the data obtained from in-vehicle navigation systems, the variance of a segment was considered a function of the variance of each link constituting the segment and the covariance between these links. However, the covariance terms were not known in this study. As a future work, it is recommended that the magnitude of travel time covariance using Bluetooth data obtained from the strategic routes be investigated, comparing the results with data from in-vehicle navigation system technology.
- As noted earlier, the outputs of the off-line calibration process can be used for real-time OD estimation along key corridors, incident management, reduction in unexpected congestion and, ultimately, provision of real-time traffic data to travelers. For these real-time operational purposes, it would be more desirable to update the dynamic OD flows within more refined time intervals (e.g. each 5 minutes). Therefore, for real-time operational analysis and incident managements, it is recommended that traffic data be acquired from Bluetooth technology (or other vendors in the market) and used to update the off-line OD flows every 5 minutes for key corridors in the study area.

Bibliography

1. Peeta, S., & Ziliaskopoulos, A. K. (2001). Foundations of dynamic traffic assignment: The past, the present and the future. *Networks and Spatial Economics, 1*(3), 233-265.

2. PTV Group, VISUM. http://www.english.ptv.de/cgi-bin/traffic/traf vision.pl. Accessed on May 4th, 2011.

3. EMME/2: INRO transportation forecasting software. http://www.inrosoftware.com/en/products/emme/index.php. Accessed on 4 April 2012.

4. Wang, Y., Messmer, A., & Papageorgiou, M. (2001). Freeway network simulation and dynamic traffic assignment with METANET tools. *Transportation Research Record: Journal of the Transportation Research Board, 1776*(1), 178-188.

5. Halati, A., Lieu, H., & Walker, S. (1997). CORSIM-corridor traffic simulation model. In *Traffic Congestion and Traffic Safety in the 21st Century: Challenges, Innovations, and Opportunities*.

6. Cameron, G. D., & Duncan, G. I. (1996). PARAMICS—Parallel microscopic simulation of road traffic. *The Journal of Supercomputing, 10*(1), 25-53.

7. Barceló, J., & Casas, J. (2005). Dynamic network simulation with AIMSUN. In *Simulation Approaches in Transportation Analysis*, 57-98, Springer US.

8. Yang, Q., & Koutsopoulos, H. N. (1996). A microscopic traffic simulator for evaluation of dynamic traffic management systems. *Transportation Research Part C: Emerging Technologies, 4*(3), 113-129.

9. Yang, Q., Koutsopoulos, H. N., & Ben-Akiva, M. E. (2000). Simulation laboratory for evaluating dynamic traffic management systems. *Transportation Research Record: Journal of the Transportation Research Board, 1710*(1), 122-130.

10. Fellendorf, M., & Vortisch, P. (2001). Validation of the microscopic traffic flow model VISSIM in different real-world situations. In *National Research Council (US). Transportation Research Board. Meeting (80th: 2001: Washington, DC). Preprint CD-ROM*.

11. Park, B. B., & Schneeberger, J. D. (2003). Microscopic simulation model calibration and validation: case study of VISSIM simulation model for a coordinated actuated signal system. *Transportation Research Record: Journal of the Transportation Research Board*, *1856*(1), 185-192.

12. Caliper, TransModeler traffic simulation software, http://www.caliper.com/transmodeler/. Accessed on April 27th, 2011.

13. Ahmed, K. I. (1999). Modeling drivers' acceleration and lane changing behavior, *Doctoral dissertation, Massachusetts Institute of Technology*.

14. Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H., & Mishalani, R. (1998, February). DynaMIT: a simulation-based system for traffic prediction. In *DACCORS Short Term Forecasting Workshop, The Netherlands*.

15. Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H. N., & Mishalani, R. (2002). Real time simulation of traffic demand-supply interactions within DynaMIT. *Applied optimization*, *63*, 19-34.

16. Mahmassani, H. S., Hu, T., & Jayakrishnan, R. (1995). Dynamic traffic assignment and simulation for advanced network informatics (DYNASMART).*Urban traffic networks: Dynamic flow modeling and control. Springer, Berlin/New York.*

17. Mahmassani, H. S. (2001). Dynamic network traffic assignment and simulation methodology for advanced system management applications. *Networks and Spatial Economics*, *1*(3-4), 267-292.

18. Mahut, M., Florian, M., Florian, D., Velan, S., & Tremblay, N. (2005). Equilibrium dynamic traffic assignment for large, congested networks. *INRO white paper*.

19. Florian, M. A., Mahut, M., & Tremblay, N. (2006). A simulation-based dynamic traffic assignment model: DYNAMEQ. *Centre for Research on Transportation*.

20. Dynamic Urban System for Transportation (DynusT), http://dynust.net/. Accessed July, 12th, 2010.

21. Chiu, Y. C., Nava, E., Zheng, H., & Bustillos, B. (2011). DynusT User's Mannual, University of Arizona.

22. Cascetta, E., Inaudi, D., & Marquis, G. (1993). Dynamic estimators of origin-destination matrices using traffic counts. *Transportation science*, *27*(4), 363-373.

23. Tavana, H. (2001). Internally-consistent estimation of dynamic network origin-destination flows from intelligent transportation systems data using bi-level optimization. *Ph.D. Dissertation, University of Texas at Austin.*

24. Zhou, X., Qin, X., & Mahmassani, H. S. (2003). Dynamic origin-destination demand estimation with multiday link traffic counts for planning applications. *Transportation Research Record: Journal of the Transportation Research Board*, *1831*(1), 30-38.

25. Nie, Y. (2006). A variational inequality approach for inferring dynamic origin-destination travel demand. *Ph.D. Dissertation, University of California Davis*.

26. Zhang, H. M., Nie, Y. M., & Qian, Z. (2008). Estimating time-dependent freeway origin-destination demands with different data coverage: sensitivity analysis. *Transportation Research Record: Journal of the Transportation Research Board*, *2047*(1), 91-99.

27. Kim, H., Baek, S., & Lim, Y. (2001). Origin-destination matrices estimated with a genetic algorithm from link traffic counts. *Transportation Research Record: Journal of the Transportation Research Board*, *1771*(1), 156-163.

28. Kattan, L., & Abdulhai, B. (2006). Non-iterative approach to dynamic traffic origin-destination estimation with parallel evolutionary algorithms. *Transportation Research Record: Journal of the Transportation Research Board*, *1964*(1), 201-210.

29. Stathopoulos, A., & Tsekeris, T. (2004). Hybrid meta-heuristic algorithm for the simultaneous optimization of the O–D trip matrix estimation. *Computer-Aided Civil and Infrastructure Engineering*, *19*(6), 421-435.

30. Vaze, V., Antoniou, C., Wen, Y., & Ben-Akiva, M. (2009). Calibration of dynamic traffic assignment models with point-to-point traffic surveillance. *Transportation Research Record: Journal of the Transportation Research Board*, *2090*(1), 1-9.

31. Cipriani, E., Florian, M., Mahut, M., & Nigro, M. (2011). A gradient approximation approach for adjusting temporal origin–destination matrices. *Transportation Research Part C: Emerging Technologies*, *19*(2), 270-282.

32. Zhou, X., & Mahmassani, H. S. (2005). Recursive approaches for online consistency checking and OD demand updating for real-time dynamic traffic assignment operation. *Transportation Research Record*, *1923*, 218-226.

33. Hu, S. R., Peeta, S., & Chu, C. H. (2009). Identification of vehicle sensor locations for link-based network traffic applications. *Transportation Research Part B: Methodological*, *43*(8), 873-894.

34. Zhou, X., Erdogan, S., & Mahmassani, H. S. (2006). Dynamic origin-destination trip demand estimation for subarea analysis. *Transportation Research Record: Journal of the Transportation Research Board*, *1964*(1), 176-184.

35. Verbas, İ. Ö., Mahmassani, H. S., & Zhang, K. (2011). Time-dependent origin-destination demand estimation. *Transportation Research Record: Journal of the Transportation Research Board*, *2263*(1), 45-56.

36. Chiu, Y. C., Zhou, L., & Song, H. (2010). Development and calibration of the anisotropic mesoscopic simulation model for uninterrupted flow facilities. *Transportation Research Part B: Methodological*, *44*(1), 152-174.

37. Tung, R., Wang, Z., & Chiu, Y. C. (2010). Integration of dynamic traffic assignment in a four step model framework: a deployment case study in Seattle model. In

Proceeding of the Third Conference on Innovations in Travel Modeling, Arizona State University, Arizona.

38. Tavana, H., & Mahmassani, H. S. (2000). Estimation and application of dynamic speed-density relations by using transfer function models. *Transportation Research Record: Journal of the Transportation Research Board*, *1710*(1), 47-57.

39. Huynh, N., Mahmassani, H. S., & Tavana, H. (2002). Adaptive speed estimation using transfer function models for real-time dynamic traffic assignment operation. *Transportation Research Record: Journal of the Transportation Research Board*, *1783*(1), 55-65.

40. Qin, X., & Mahmassani, H. S. (2004). Adaptive calibration of dynamic speed-density relations for online network traffic estimation and prediction applications. *Transportation Research Record: Journal of the Transportation Research Board*, *1876*(1), 82-89.

41. Doan, D. L., Ziliaskopoulos, A., & Mahmassani, H. (1999). On-line monitoring system for real-time traffic management applications. *Transportation Research Record: Journal of the Transportation Research Board*, *1678*(1), 142-149.

42. Peeta, S., & Bulusu, S. (1999). Generalized singular value decomposition approach for consistent on-line dynamic traffic assignment. *Transportation Research Record: Journal of the Transportation Research Board*, *1667*(1), 77-87.

43. He, R., Miaou, S., Ran, B., & Lan, C. (1999). Developing an on-line calibration process for an analytical dynamic traffic assignment model. In *78th Annual Meeting of the Transportation Research Board, Washington, DC*.

44. He, R. R., & Ran, B. (2000). Calibration and validation of a dynamic traffic assignment model. *Transportation Research Record: Journal of the Transportation Research Board*, *1733*(1), 56-62.

45. Hawas, Y. E. (2002). Calibrating simulation models for advanced traveler information systems/advanced traffic management systems applications. *Journal of transportation engineering*, *128*(1), 80-88.

46. Hawas, Y. E. (2000). Integrated traffic assignment and signal control for on-line operation. In *Proc., 7th World Congress on Intelligent Transportation Systems*.

47. Chen, Y. S., Van Zuylen, H. J., & Lee, R. (2006). Developing a large-scale urban decision support system. In *Control in Transportation Systems,* 11(1), 216-221.

48. Chu, L., Liu, H. X., Oh, J. S., & Recker, W. (2003). A calibration procedure for microscopic traffic simulation. In *Intelligent Transportation Systems, 2003. Proceedings. 2003 IEEE*, Vol. 2, pp. 1574-1579, IEEE.

49. Mahut, M., Florian, M., Tremblay, N., Campbell, M., Patman, D., & McDaniel, Z. K. (2004). Calibration and application of a simulation-based dynamic traffic assignment model. *Transportation Research Record: Journal of the Transportation Research Board*, *1876*(1), 101-111.

50. Mahmassani, H. S., Qin, X., & Zhou, X. (2004). DYNASMART–X evaluation for real–time TMC application: Irvine test bed. *Maryland Transportation Initiative, University of Maryland, College Park, Maryland*.

51. Balakrishna, R., Koutsopoulos, H. N., & Ben-Akiva, M. E. (2005). Calibration and validation of dynamic traffic assignment systems. In *Transportation and Traffic Theory. Flow, Dynamics and Human Interaction. 16th International Symposium on Transportation and Traffic Theory*.

52. Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H. N., & Mishalani, R. (2002). Real time simulation of traffic demand-supply interactions within DynaMIT. *Applied optimization*, *63*, 19-34.

53. Ben-Akiva, M., & Bierlaire, M. (2003). Discrete choice models with applications to departure time and route choice. In *Handbook of transportation science* (pp. 7-37). Springer US.

54. Kundie, K. K. (2002). Calibration of mesoscopic traffic simulation models for dynamic traffic assignment, *Doctoral dissertation, Massachusetts Institute of Technology*.

55. Box, M. J. (1965). A new method of constrained optimization and a comparison with other methods. *The Computer Journal*, *8*(1), 42-52.

56. Gupta, A. (2005). Observability of Origin-destination matrices for dynamic traffic assignment, *Doctoral dissertation, Massachusetts Institute of Technology*.

57. Mahanti, B. P. (2004). Aggregate calibration of microscopic traffic simulation models. *Master's thesis, Massachusetts Institute of Technology*.

58. Toledo, T., Ben-Akiva, M. E., Darda, D., Jha, M., & Koutsopoulos, H. N. (2004). Calibration of microscopic traffic simulation models with aggregate data. *Transportation Research Record: Journal of the Transportation Research Board*, *1876*(1), 10-19.

59. Toledo, T., Koutsopoulos, H. N., Davol, A., Ben-Akiva, M. E., Burghout, W., Andréasson, I & Lundin, C. (2003). Calibration and validation of microscopic traffic simulation tools: Stockholm case study. *Transportation Research Record: Journal of the Transportation Research Board*, *1831*(1), 65-75.

60. Darda, D. (2002). Joint calibration of a microscopic traffic simulator and estimation of origin-destination flows, *Doctoral dissertation, Massachusetts Institute of Technology*.

61. Jha, M., Gopalan, G., Garms, A., Mahanti, B. P., Toledo, T., & Ben-Akiva, M. E. (2004). Development and calibration of a large-scale microscopic traffic simulation model. *Transportation Research Record: Journal of the Transportation Research Board*, *1876*(1), 121-131.

62. Kim, S. J. (2006). Simultaneous calibration of a microscopic traffic simulation model and OD matrix, *Doctoral dissertation, Texas A&M University*.

63. Balakrishna, R. (2006). Off-line calibration of dynamic traffic assignment models, *Doctoral dissertation, Massachusetts Institute of Technology*.

64. Balakrishna, R., Ben-Akiva, M., & Koutsopoulos, H. N. (2007). Offline calibration of dynamic traffic assignment: simultaneous demand-and-supply estimation. *Transportation Research Record: Journal of the Transportation Research Board*, *2003*(1), 50-58.

65. Huyer, W., & Neumaier, A. (2008). SNOBFIT: stable noisy optimization by branch and fit. *ACM Transactions on Mathematical Software (TOMS)*, *35*(2), 9.

66. Spall, J. C. (1998). An overview of the simultaneous perturbation method for efficient optimization. *Johns Hopkins APL Technical Digest*, *19*(4), 482-492.

67. Pel, A. J., Bliemer, M. C., & Hoogendoorn, S. P. (2009). Hybrid route choice modeling in dynamic traffic assignment. *Transportation Research Record: Journal of the Transportation Research Board*, *2091*(1), 100-107.

68. Qian, Z. S., & Zhang, H. M. (2012). A hybrid route choice model for dynamic traffic assignment. *Networks and Spatial Economics*, 1-21.

69. Balakrishna, R., Antoniou, C., Ben-Akiva, M., Koutsopoulos, H. N., & Wen, Y. (2007). Calibration of microscopic traffic simulation models: Methods and application. *Transportation Research Record: Journal of the Transportation Research Board*, *1999*(1), 198-207.

70. Antoniou, C. (2004). On-line calibration for dynamic traffic assignment, Doctoral dissertation, *Massachusetts Institute of Technology*.

71. Antoniou, C., Ben-Akiva, M., & Koutsopoulos, H. N. (2005). Online calibration of traffic prediction models. *Transportation Research Record: Journal of the Transportation Research Board*, *1934*(1), 235-245.

72. Ashok, K., & Ben-Akiva, M. E. (2000). Alternative approaches for real-time estimation and prediction of time-dependent origin–destination flows. *Transportation Science*, *34*(1), 21-36.

73. Vaze, V. S. (2007). Calibration of dynamic traffic assignment models with point-to-point traffic surveillance. *MSc thesis, Massachusetts Institute of Technology*.

74. Huang, E. (2010). Algorithmic and implementation aspects of on-line calibration of dynamic traffic assignment, *Doctoral dissertation, Massachusetts Institute of Technology*.

75. Huang, E., Antoniou, C., Lopes, J., Wen, Y., & Ben-Akiva, M. (2010). Accelerated on-line calibration of dynamic traffic assignment using distributed stochastic gradient approximation. In *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on* (pp. 1166-1171). IEEE.

76. Hooke, R., & Jeeves, T. A. (1961). Direct search solution of numerical and statistical problems. *Journal of the ACM (JACM)*, *8*(2), 212-229.

77. Huang, E., Antoniou, C., Wen, Y., Ben-Akiva, M., Lopes, J., & Bento, J. (2009). Real-time multi-sensor multi-source network data fusion using dynamic traffic assignment models. In *Intelligent Transportation Systems, 2009. ITSC'09. 12th International IEEE Conference on* (pp. 1-6). IEEE.

78. Appiah, J. & Rilett L.R. (2010). Joint estimation of dynamic origin-destination matrices and calibration of micro-simulation models using aggregate intersection turn count data. *Transportation Research Board 89th Annual Meeting, Transportation Research Board Annual Meeting Paper*, 10-2764.

79. Omrani, R., & Kattan, L. (2012). Demand and supply calibration of dynamic traffic assignment models: past efforts and future challenges. *Transportation Research Record: Journal of the Transportation Research Board*, *2283*(1), 100-112.

80. Special Report 209: Highway Capacity Manual, 3rd Edition. (1994). Transportation Research Board, National Research Council, Washington, D.C.

81. Kattan, L. (2005). Dynamic traffic origin/destination estimation using evolutionary based algorithms, *Doctoral dissertation, University of Toronto.*

82. Cantu-Paz, E. (2000). Efficient and accurate parallel genetic algorithms, Vol. 1, Springer.

83. Holland J. (1975). Adaptation in natural and artificial system. *University of Michigan Press*, Ann Arbor.

84. Ma, T., & Abdulhai, B. (2002). Genetic algorithm-based optimization approach and generic tool for calibrating traffic microscopic simulation parameters. *Transportation Research Record: Journal of the Transportation Research Board*, *1800*(1), 6-15.

85. Maslov, I. V., & Gertner, I. (2006). Multi-sensor fusion: an evolutionary algorithm approach. *Information Fusion*, *7*(3), 304-330.

86. Goldberg, D. E. (2002). The design of innovation: lessons from and for competent genetic algorithms, *by David E. Goldberg* (Vol. 7). Springer.

87. Mohamed, M. S. M. (2007). Generic parallel genetic algorithms framework for optimizing intelligent transportation systems (GENOTRANS). In *Masters Abstracts International*, Vol. 46, No. 06.

88. Baker, J. E. (1985). Adaptive selection methods for genetic algorithms. In *Proceedings of the 1st International Conference on Genetic Algorithms*, pp. 101-111, L. Erlbaum Associates Inc..

89. Herrera, F., Lozano, M., & Verdegay, J. L. (1998). Tackling real-coded genetic algorithms: Operators and tools for behavioral analysis. *Artificial intelligence review*, *12*(4), 265-319.

90. Back, T., Hammel, U., & Schwefel, H. P. (1997). Evolutionary computation: Comments on the history and current state. *Evolutionary computation, IEEE Transactions on*, *1*(1), 3-17.

91. Abd, H. M. A. E. H. (2010). *Optimization of Multimodal Evacuation of Large-scale Transportation Networks*, Doctoral dissertation, University of Toronto.

92. Cantu-Paz, E. (2000). *Efficient and accurate parallel genetic algorithms* (Vol. 1). Kluwer Academic Publisher, Springer.

93. Tomassini, M. (1999). Parallel and distributed evolutionary algorithms: a review. *Evolutionary Algorithms in Engineering and Computer Science: Recent Advances in Genetic Algorithms, Evolution Strategies,* Evolutionary Programming, Genetic Programming, and Industrial Applications: 113-131.

94. Mohamed, M. (2007). *Generic parallel genetic algorithms framework for optimizing intelligent transportation systems (GENOTRANS),* Master of Science Thesis, University of Toronto.

95. Cantu-Paz, E. (1998). Using Markov chains to analyze a bounding case of parallel genetic algorithms. In *Genetic Programming: Proceedings of the Third Annual Conference*, pp. 456-462, San Francisco, CA: Morgan Kaufmann Publishers.

96. Brown, M., Fukui, K., & Trivedi, N. (2005). *Introduction to grid computing*. IBM, International Technical Support Organization.

97. Ivanov, N., *Real Time Big Data Processing with GridGain,* Grid Computing, http://www.gridgain.com/book/book.html, Accessed 12/07/2012.

98. Pop, F., Lovin, M. A., Cristea, V., Bessis, N., & Sotiriadis, S. (2012, July). Applications monitoring for self-optimization in GridGain. In *Complex, Intelligent and Software Intensive Systems (CISIS), 2012 Sixth International Conference on* (pp. 755-760). IEEE.

99. Dowling, R., Skabardonis, A., & Alexiadis, V. (2004). *Traffic analysis toolbox volume III: Guidelines for applying traffic micro-simulation modeling software* (No. FHWA-HRT-04-040).

100. Abdulhai B, Georgi A, Roorda MJ, Tarabain A, and Tang E. (2004). *Micro simulation of the Toronto waterfront revitalization plan: impact assessment of roadway configuration alternatives*. Final Report for the Toronto Waterfront Revitalization Corporation.

101. 2012 Travel Time Studies, Ministry of Transportation Ontario, Ontario, Canada.

102. Roorda, M. J., Hain, M., Amirjamshidi, G., Cavalcante, R., Abdulhai, B., & Woudsma, C. (2010). Exclusive truck facilities in Toronto, Ontario, Canada. *Transportation Research Record: Journal of the Transportation Research Board*, *2168*(1), 114-128.

103. Internet, ONE-ITS: The next level of traffic innovation, http://one-itswebapp1.transport.utoronto.ca/web/one-its, accessed 27/11/2012.

104. Omrani, R., Izadpanah, P., Hellinga, B., Hadayeghi, A., & Abdelgawad, H. (2013). Evaluation of wide-area traffic monitoring technologies for travel time studies. In *Transportation Research Board 92nd Annual Meeting* (No. 13-5273).

105. Hayya, J., Armstrong, D., & Gressis, N. (1975). A note on the ratio of two normally distributed variables. *Management Science*, *21*(11), 1338-1341.

106. Brackstone, M., Waterson, B., & McDonald, M. (2009). Determinants of following headway in congested traffic. *Transportation research part F: traffic psychology and behavior*, *12*(2), 131-142.

107. Lenné, M. G., Triggs, T. J., & Redman, J. R. (1997). Time of day variations in driving performance. *Accident Analysis & Prevention*, *29*(4), 431-437.

108. Theil, H. (1961). Economic forecasts and policy. North-Holland, Amsterdam, Netherlands.

Appendix A: Statistical Validity of Sample Size

The traffic data obtained from in-vehicle navigation system technology include the minimum, maximum, and average number of observations for each road segment within the study area. As stated earlier, the vendor was able to provide a significantly large number of observations per road section and per peak period.

In the probability theory, the central limit theorem (CLT) states that the mean of a sufficiently large number of independent random variables, each with a well-defined mean and well-defined variance, will be approximately normally distributed[7]. Most statistical text books have stated that if sample size is larger than 30, the distribution of sample mean X_n can assume to be normally distributed[8,9,10,11,12]. Given a large sample size from the data provider enables us to assume that the distribution of the sample can be approximated as a normal distribution.

Suppose a random variable X_p is normally distributed with unknown mean μ and unknown standard deviation S. Also, a sample data (X_s) of size n is collected from the population with the sample mean of $\bar{x}$ and a standard deviation of σ, with the following notations:

$$X_p \sim N(\mu, S^2) \tag{19}$$

$$X_s \sim N(\bar{x}, \sigma^2) \tag{20}$$

The difference between the sample mean and the true value of the population mean can be expressed as an error term (δ) as follows:

---

[7] Rice, J, Mathematical Statistics and Data Analysis (Second ed.), Duxbury Press, ISBN 0-534-20934-3, 1995.
[8] Becker, W. E., Statistics: for Business and Economics, South-Western, Cincinnati, pp.271-273, 1995
[9] Freund, J. E. and Perles, B. M., Statistics: A First Course, 7th ed., Prentice-Hall, New Jersey, pp.275-279, 1999.
[10] Hogg, R. V. and Tanis, E. A., Probability and Statistical Inference, 6th ed., Prentice-Hall, New Jersey, pp.307-313, 2001
[11] Levine, D. M., Ramsey, P. P. and Berenson, M. L., Business statistics for quality and productivity, Prentice-Hall, New Jersey, pp.259-264, 1995.
[12] Watson, C. J., Billingsley, P., Croft, D. J. and Huntsberger, D. V., Statistics: for Management and Economics, 5th ed., Prentice-Hall, New Jersey, pp.297-305, 1996.

$$-\delta \leq \mu - \bar{x} \leq \delta \quad (21)$$

$$\delta = {Z_{\alpha/2} \cdot \sigma}/{\sqrt{n}} \quad (22)$$

Where $Z_{\alpha/2}$ is the Z-score of the standard normal distribution table with the confidence interval of $100(1 - {\alpha}/{2})\%$, and α is the predetermined significance level (e.g. 0.05). For $\alpha = 0.05$, the $Z_{\alpha/2}$ value for the 95% confidence interval is 1.96.

Based on the equation for the error term, Equation 21 can be rephrased as follows:

$$|\mu - \bar{x}| \leq \frac{Z_{\alpha/2} \cdot \sigma}{\sqrt{n}} \quad (23)$$

In Equation 23, the absolute difference the population and sample means (i.e. $|\mu - \bar{x}|$) is unknown. A 10% margin of error between the population mean and the sample mean were assumed in this research.

$$\frac{|\mu - \bar{x}|}{\bar{x}} \leq 0.1 \quad (24)$$

Equation 24 was then integrated into Equation 5 to determine the minimum required sample size (n_{min}), as follows:

$$\frac{Z_{\alpha/2} \cdot \sigma}{\sqrt{n} \times \bar{x}} \leq \frac{|\mu - \bar{x}|}{\bar{x}} \leq 0.1 \quad (25)$$

$$n \geq \left(\frac{Z_{\alpha/2} \cdot \sigma}{0.1\bar{x}}\right)^2 \quad (26)$$

$$n_{min} = \left(\frac{Z_{\alpha/2} \cdot \sigma}{0.1\bar{x}}\right)^2 \quad (27)$$

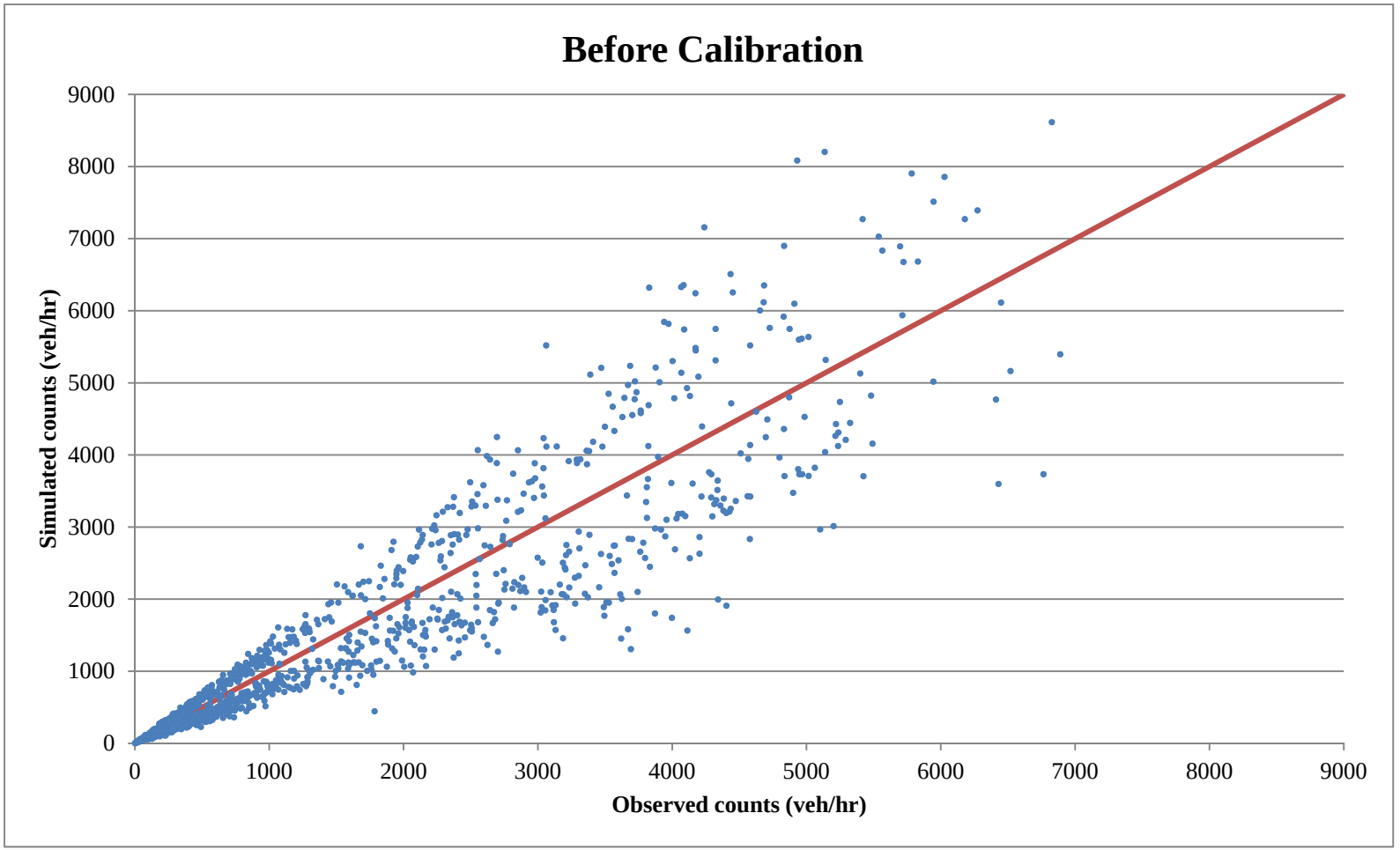
The following table presents an example of travel time and speed performance measures along the Gardiner Expressway during AM Peak Periods of the Fall Season. In addition, the minimum sample size requirement and the observed number of hits (e.g. sample size) are provided in this table. As shown in this table, the provided sample size is significantly higher

than the minimum required sample size. As noted earlier, the data collected satisfied the sample size requirements for more than 99.5% of the segments in all peak periods

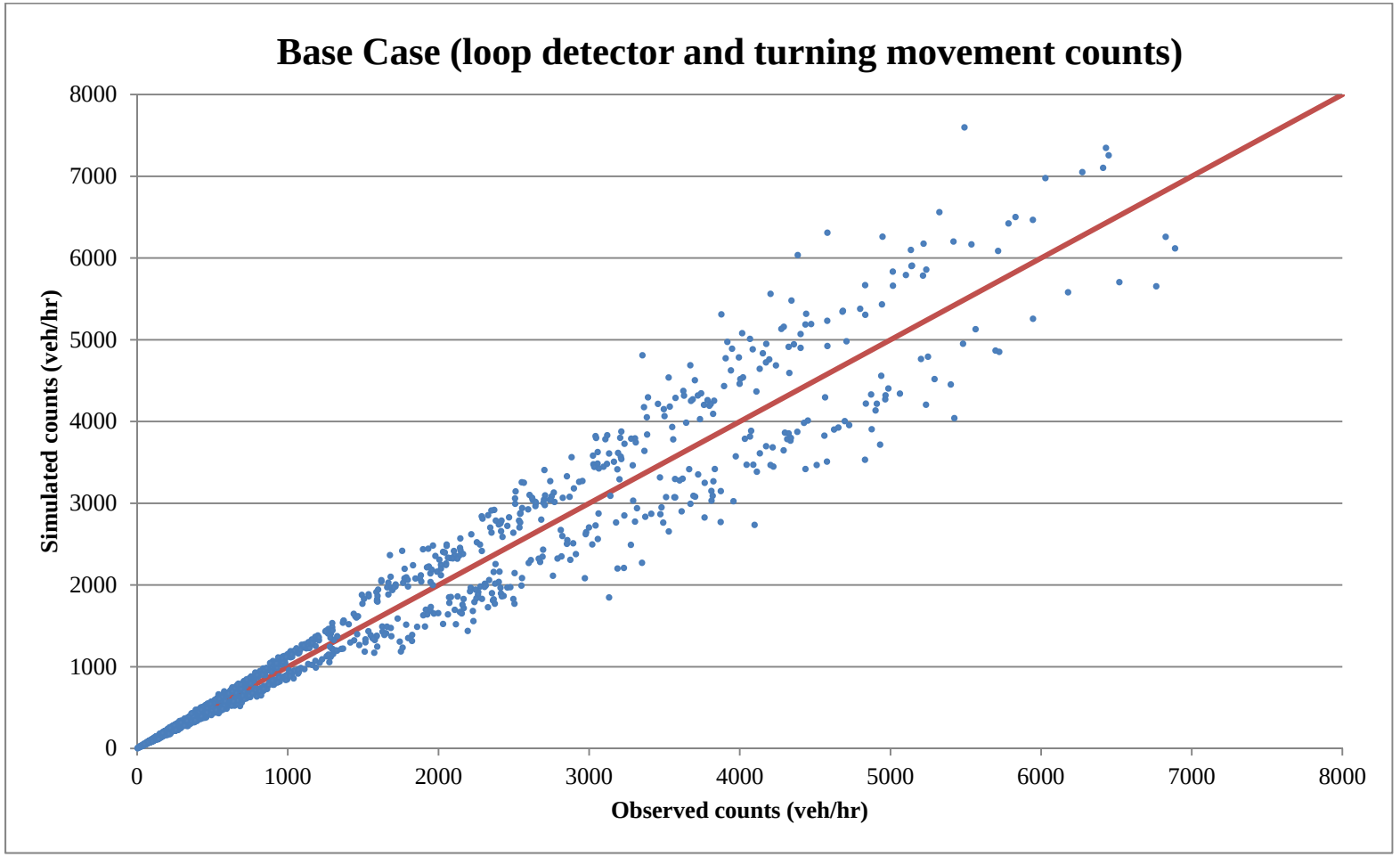
Example of minimum sample size and the number of observations

Direction	Interchange (From/To)	Length (Km)	Average travel time (h:mm:ss)	Standard deviation of TT (sec)	Average speed (Km/hr)	Standard deviation of Speed	Minimum sample size	Hits
Westbound	DVP							
	Lower Jarvis	1.7	0:01:52	43.4	54.7	21.2	57.8	564.3
	Spadina Ave	2	0:02:54	45.5	41.1	10.8	26.6	542.8
	Jameson Ave	3.5	0:04:34	48.1	46.3	8.1	11.6	787.5
	South Kingsway	3.3	0:03:16	27.3	60.4	8.5	7.5	872.2
	Islington Ave	3.8	0:03:29	23.9	64.6	7.5	5.2	940.1
	Kipling Ave	1.1	0:00:47	9.4	80.7	16.8	16.6	758.5
	Hwy 427	1.9	0:01:33	26.1	74.4	20.7	29.8	598.2
	Total/Average	17.3	0:18:25		56.37			

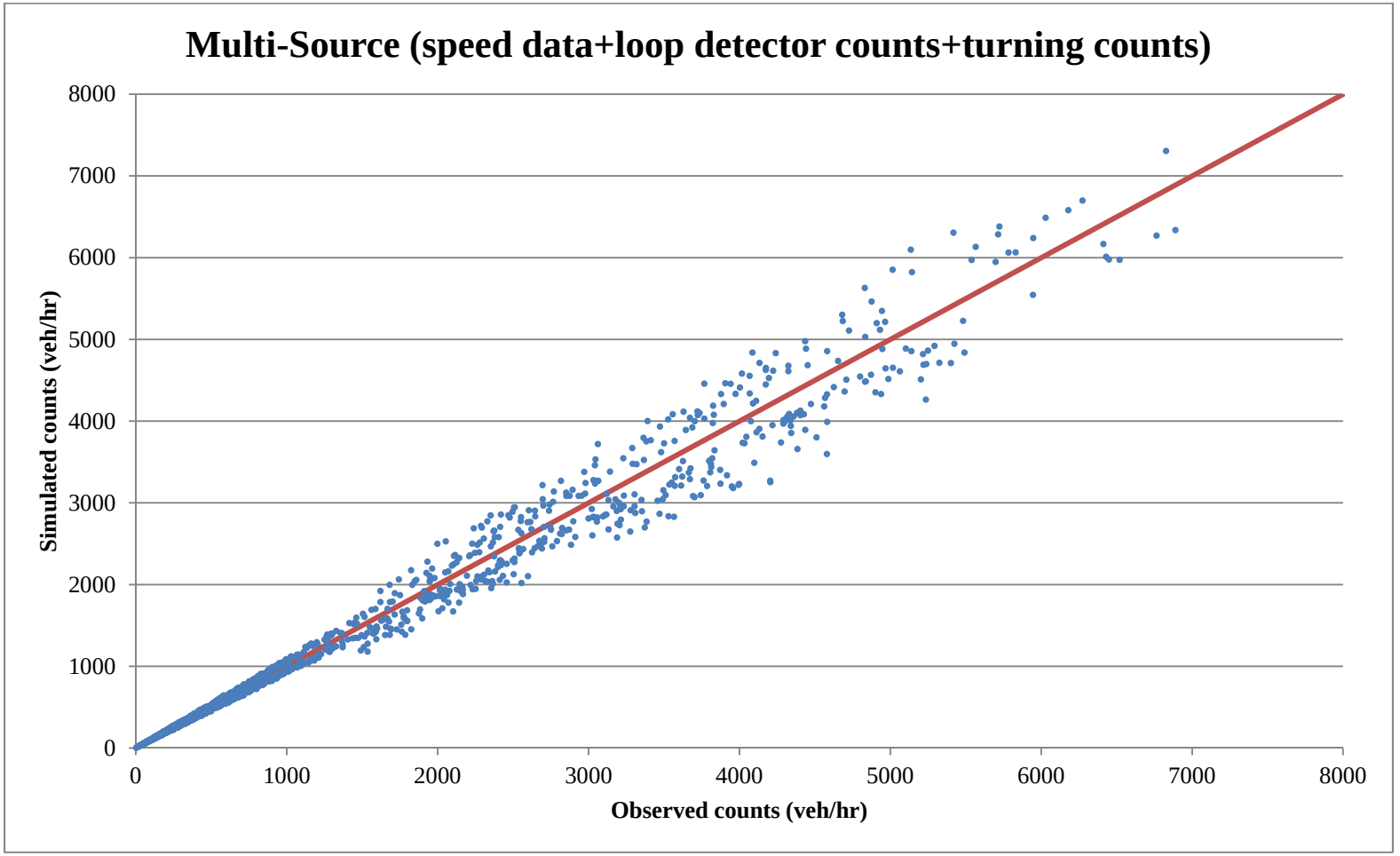
Appendix B: Comparison between Observed Counts/Speed Data and Their Simulated Counterparts for Different Scenarios (Experiment I)



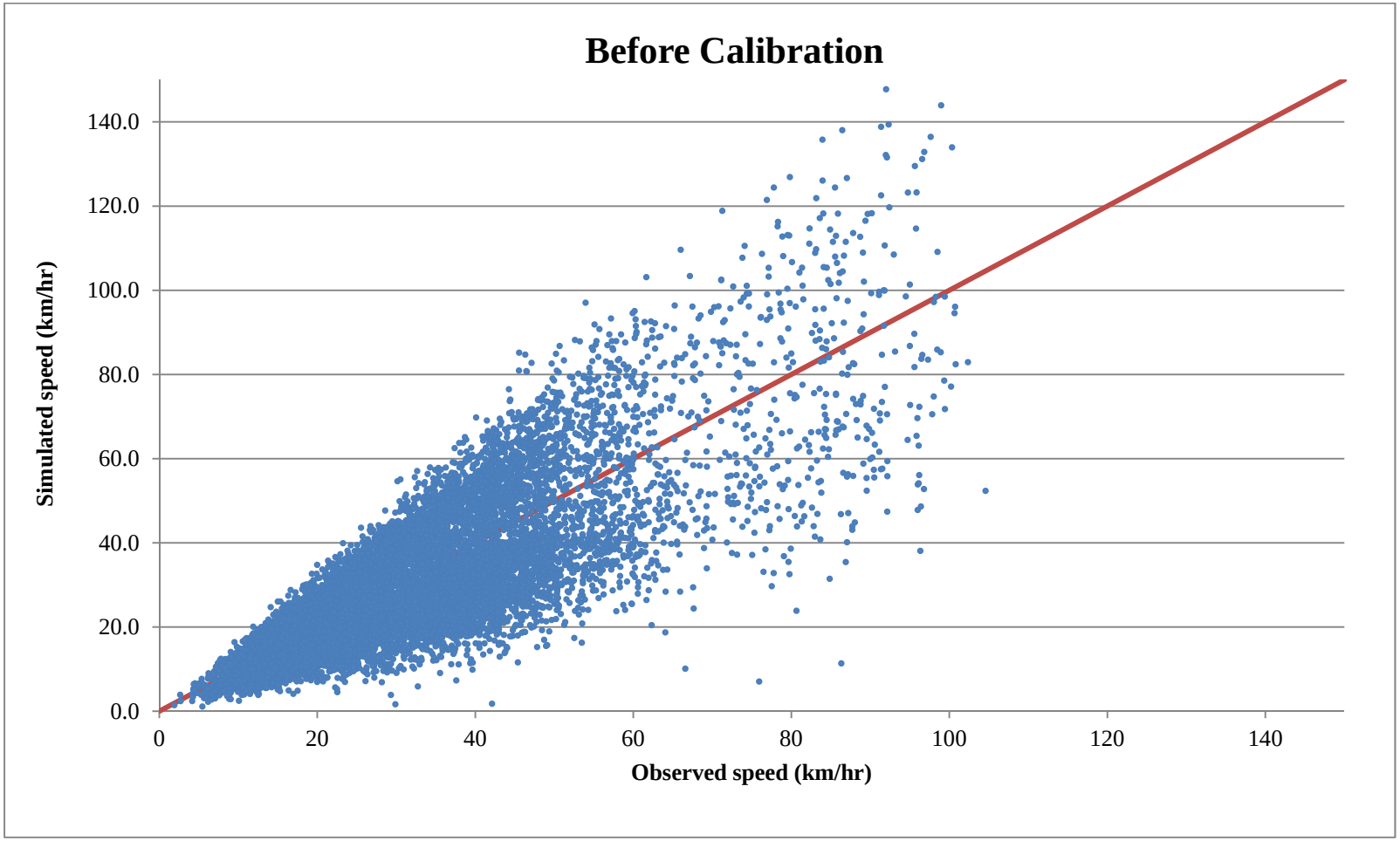
Comparison between simulated and observed loop detector and turning movements counts (Before calibration)



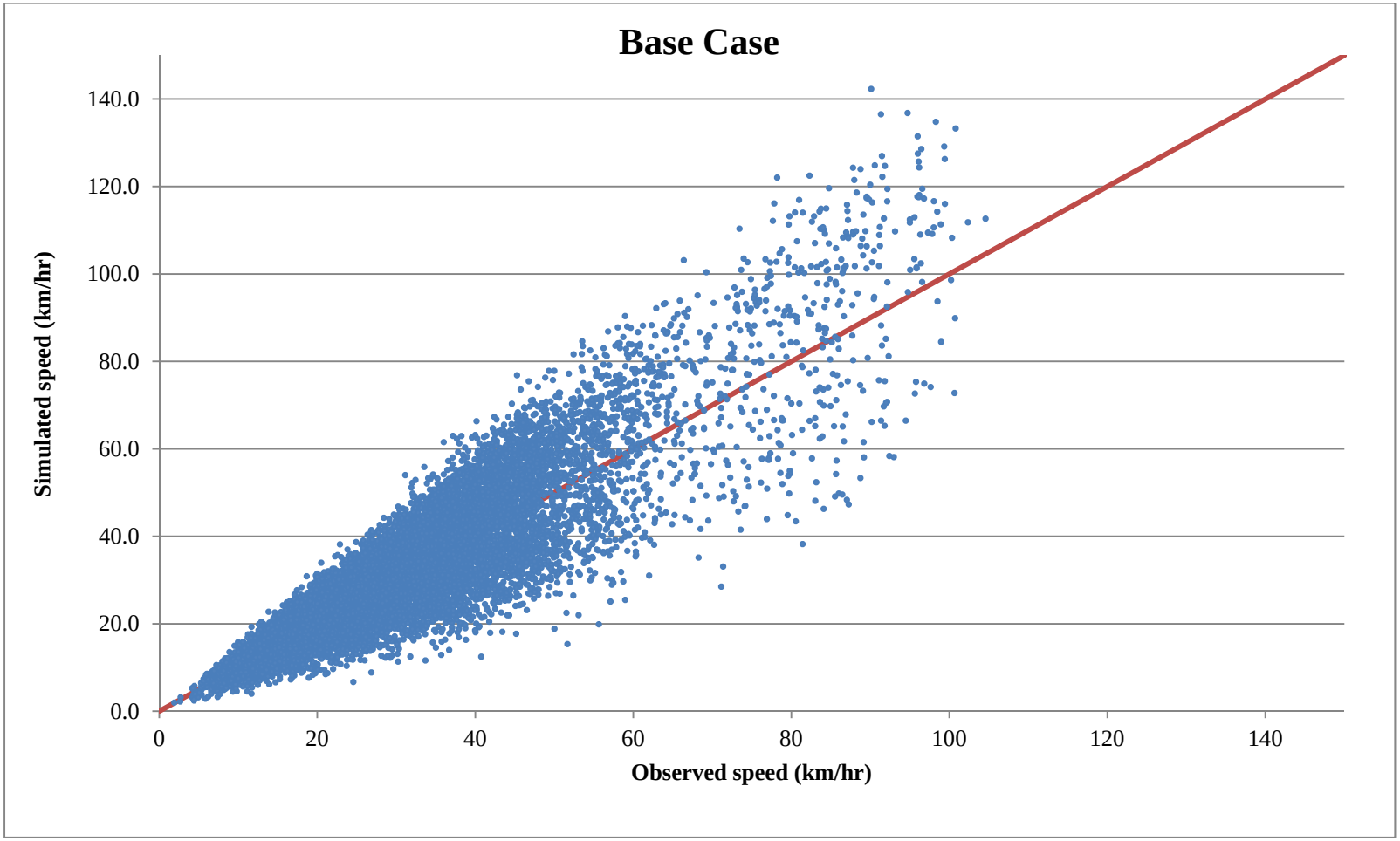
Comparison between simulated and observed loop detector and turning movements counts (Base case)



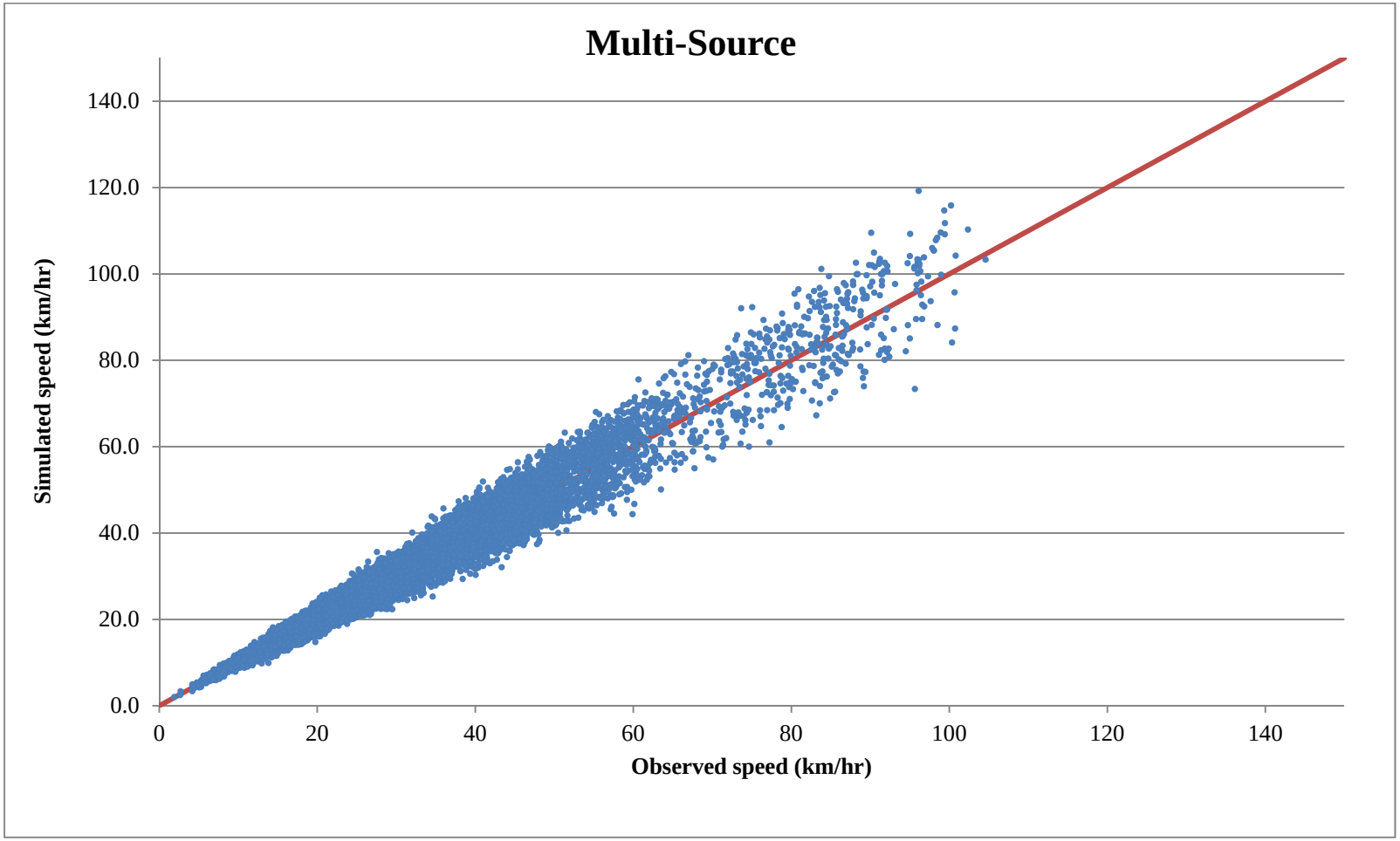
Comparison between simulated and observed loop detector and turning movements counts (Multi-source case)



Comparison between simulated and observed speed data from in-vehicle navigation systems (Before calibration)

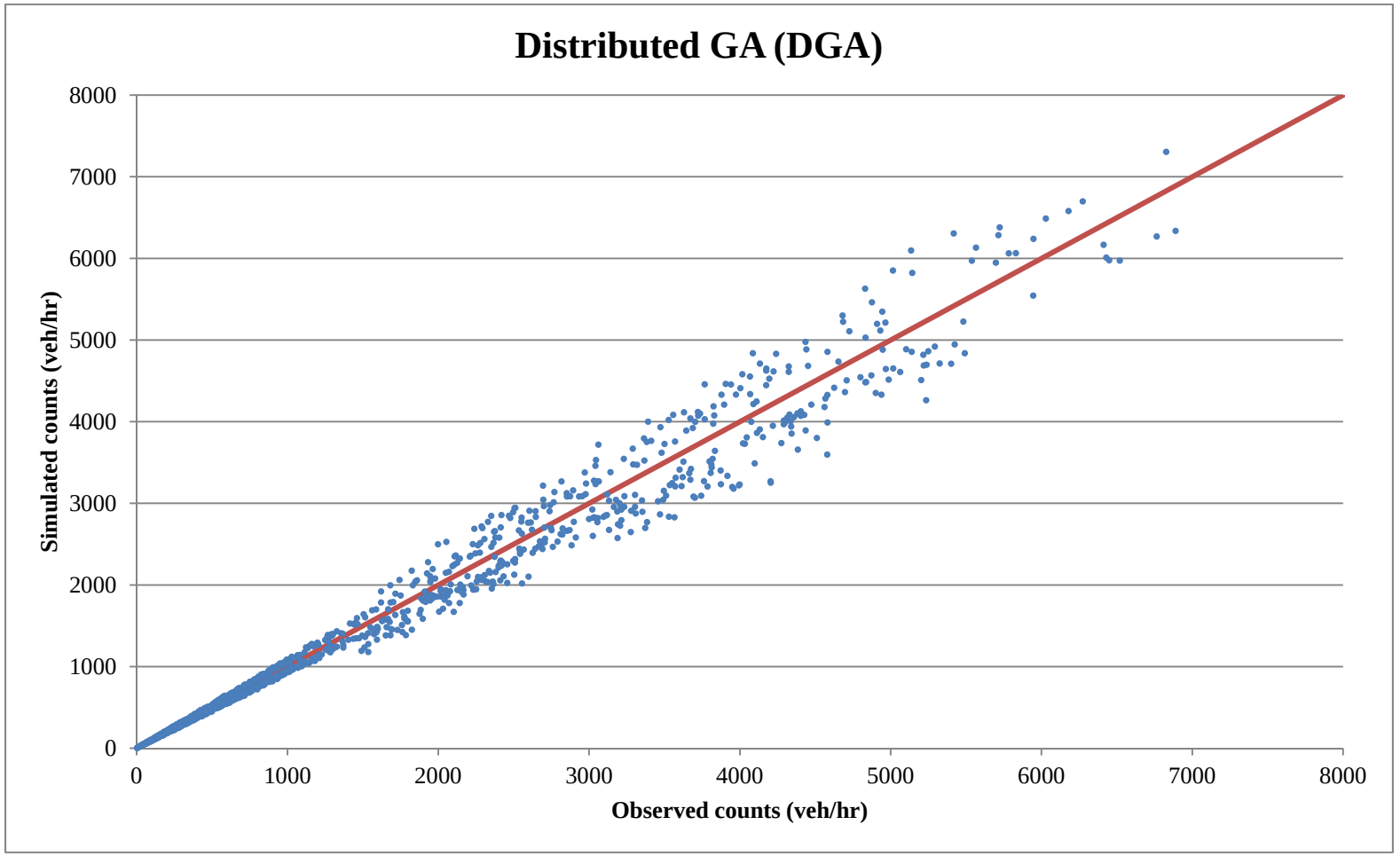


Comparison between simulated and observed speed data from in-vehicle navigation systems (Base case)

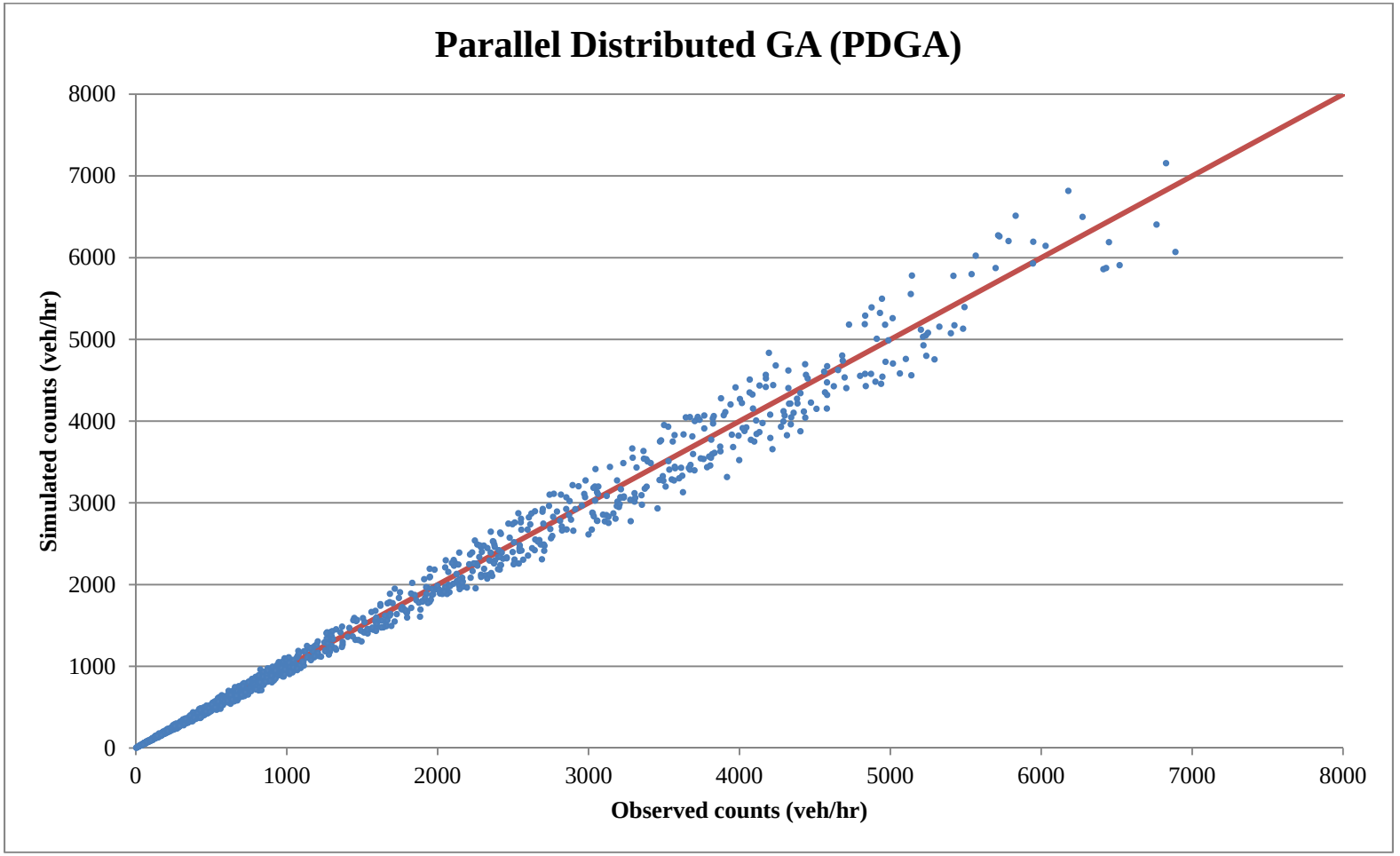


Comparison between simulated and observed speed data from in-vehicle navigation systems (Multi-source)

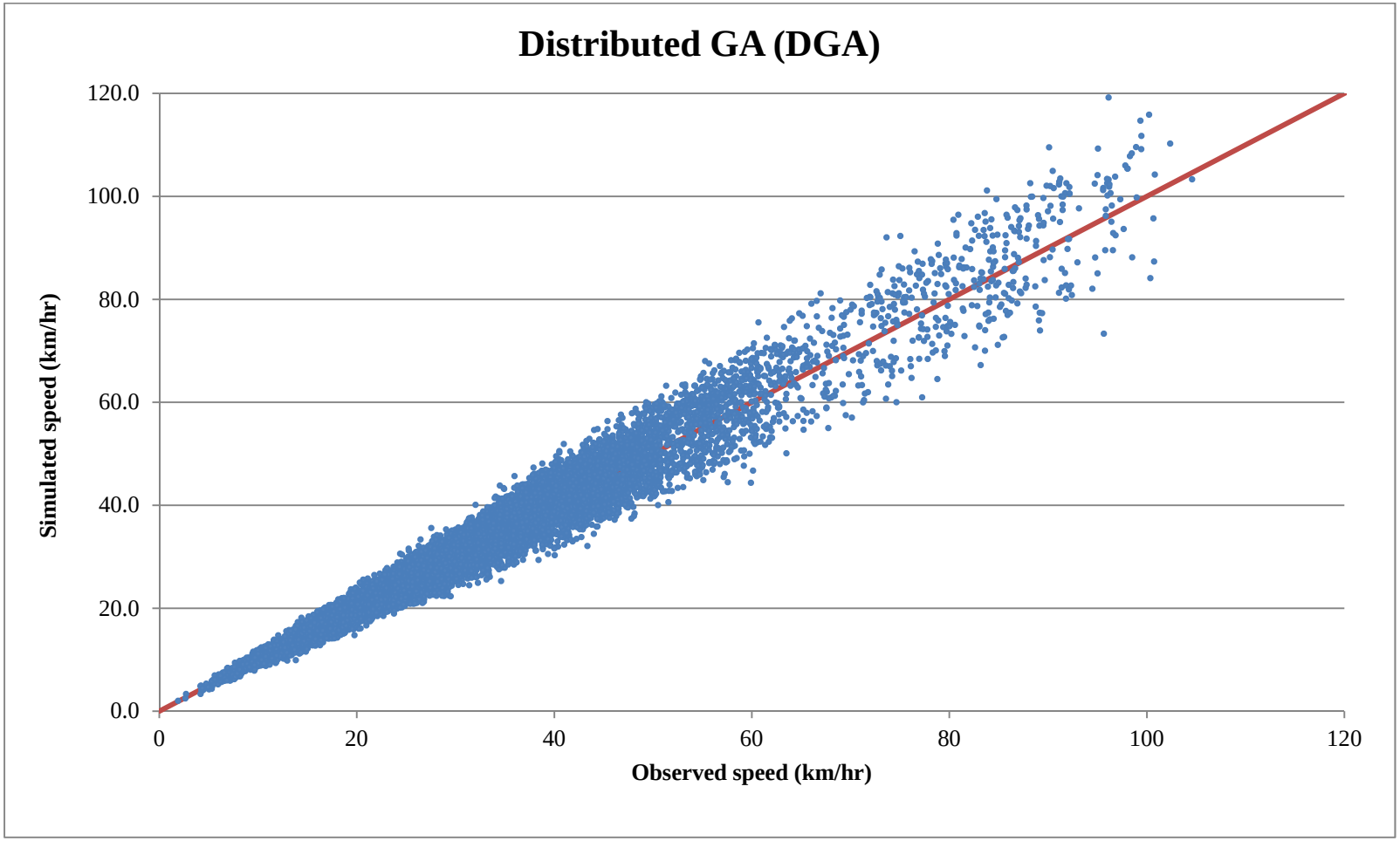
Appendix C: Comparison between Observed Counts/Speed Data and Their Simulated Counterparts for Multi-Source Scenario (Experiment II)



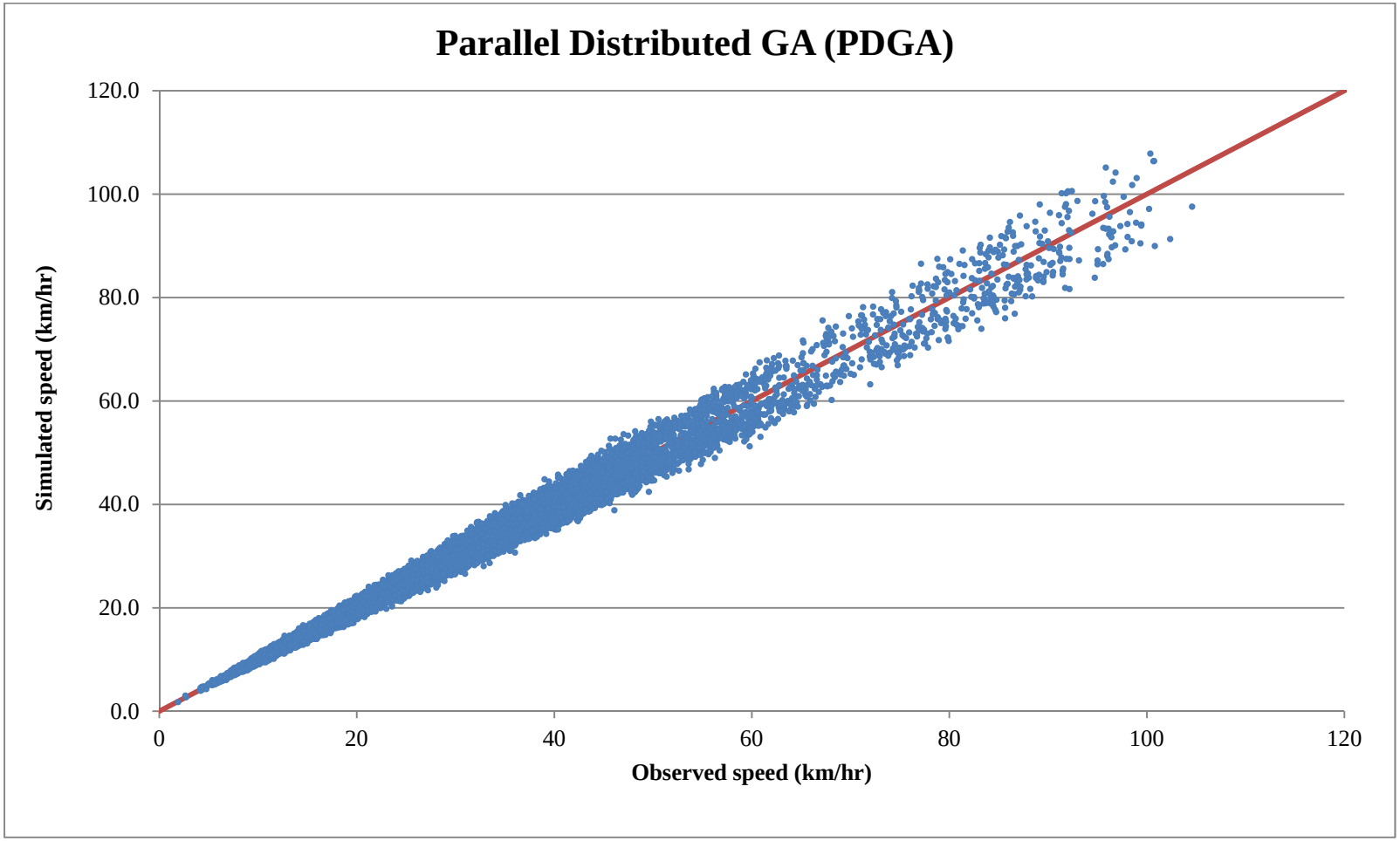
Comparison between simulated and observed counts using multi-source data (DGA)



Comparison between simulated and observed counts using multi-source data (PDGA)



Comparison between simulated and observed speed using multi-source data (DGA)



Comparison between simulated and observed speed using multi-source data (PDGA)

Appendix D: Simulated OD Flows for Different Time Intervals Using PDGA (Experiment II)

Dynamic OD trips for 15-minutes intervals (6:30AM-7:30AM)

6:30 AM – 6:45 AM	To external gateways	To study area zones	**Total**
From external gateways	2,225	1,234	3,459
From study area zones	634	29	663
Total	2,859	1,263	4,122
6:45 AM – 7:00 AM	To external gateways	To study area zones	**Total**
From external gateways	2,425	1,255	3,680
From study area zones	679	43	722
Total	3,104	1,298	4,402
7:00 AM – 7:15 AM	To external gateways	To study area zones	**Total**
From external gateways	2,698	1,389	4,087
From study area zones	689	52	741
Total	3,387	1,441	4,828
7:15 AM – 7:30 AM	To external gateways	To study area zones	**Total**
From external gateways	2,912	1,467	4,379
From study area zones	665	56	721
Total	3,577	1,523	5,100

Dynamic OD trips for 15-minutes intervals (7:30AM-8:30AM)

7:30 AM – 7:45 AM	To external gateways	To study area zones	**Total**
From external gateways	3,495	1,695	**5,190**
From study area zones	745	63	**808**
Total	**4,240**	**1,758**	**5,998**
7:45 AM – 8:00 AM	To external gateways	To study area zones	**Total**
From external gateways	3,845	1,868	**5,713**
From study area zones	778	75	**853**
Total	**4,623**	**1,943**	**6,566**
8:00 AM – 8:15 AM	To external gateways	To study area zones	**Total**
From external gateways	3,672	1,975	**5,647**
From study area zones	769	62	**831**
Total	**4,441**	**2,037**	**6,478**
8:15 AM – 8:30 AM	To external gateways	To study area zones	**Total**
From external gateways	3,751	1,990	**5,741**
From study area zones	746	69	**815**
Total	**4,497**	**2,059**	**6,556**

Dynamic OD trips for 15-minutes intervals (8:30AM-9:30AM)

8:30 AM – 8:45 AM	To external gateways	To study area zones	**Total**
From external gateways	3,521	1,825	**5,346**
From study area zones	697	61	**758**
Total	**4,218**	**1,886**	**6,104**
8:45 AM –9:00 AM	To external gateways	To study area zones	**Total**
From external gateways	3,225	1,698	**4,923**
From study area zones	674	54	**728**
Total	**3,899**	**1,752**	**5,651**
9:00 AM – 9:15 AM	To external gateways	To study area zones	**Total**
From external gateways	3,091	1,705	**4,796**
From study area zones	592	65	**657**
Total	**3,683**	**1,770**	**5,453**
9:15 AM – 9:30 AM	To external gateways	To study area zones	**Total**
From external gateways	2,756	1,525	**4,281**
From study area zones	556	49	**605**
Total	**3,312**	**1,574**	**4,886**

Dynamic OD trips for 15-minutes intervals (3:30PM-4:30PM)

3:30 PM – 3:45 PM	To external gateways	To study area zones	**Total**
From external gateways	2,958	678	**3,636**
From study area zones	1,270	58	**1,328**
Total	**4,228**	**736**	**4,964**
3:45 PM –4:00 PM	To external gateways	To study area zones	**Total**
From external gateways	3,258	685	**3,943**
From study area zones	1,395	62	**1,457**
Total	**4,653**	**747**	**5,400**
4:00 PM – 4:15 PM	To external gateways	To study area zones	**Total**
From external gateways	3,498	795	**4,293**
From study area zones	1,425	67	**1,492**
Total	**4,923**	**862**	**5,785**
4:15 PM – 4:30 PM	To external gateways	To study area zones	**Total**
From external gateways	3,605	799	**4,404**
From study area zones	1,465	59	**1,524**
Total	**5,070**	**858**	**5,928**

Dynamic OD trips for 15-minutes intervals (4:30 PM-5:30PM)

4:30 PM – 4:45 PM	To external gateways	To study area zones	**Total**
From external gateways	3,925	842	**4,767**
From study area zones	1,531	71	**1,602**
Total	**5,456**	**913**	**6,369**
4:45 PM –5:00 PM	To external gateways	To study area zones	**Total**
From external gateways	4,218	880	**5,098**
From study area zones	1,674	79	**1,753**
Total	**5,892**	**959**	**6,851**
5:00 PM – 5:15 PM	To external gateways	To study area zones	**Total**
From external gateways	4,115	1,009	**5,124**
From study area zones	1,665	76	**1,741**
Total	**5,780**	**1,085**	**6,865**
5:15 PM – 5:30 PM	To external gateways	To study area zones	**Total**
From external gateways	4,058	984	**5,042**
From study area zones	1,635	69	**1,704**
Total	**5,693**	**1,053**	**6,746**

Dynamic OD trips for 15-minutes intervals (5:30 PM-6:30PM)

5:30 PM – 5:45 PM	To external gateways	To study area zones	**Total**
From external gateways	3,714	877	**4,591**
From study area zones	1,526	58	**1,584**
Total	**5,240**	**935**	**6,175**
5:45 PM –6:00 PM	To external gateways	To study area zones	**Total**
From external gateways	3,605	829	**4,434**
From study area zones	1,476	65	**1,541**
Total	**5,081**	**894**	**5,975**
6:00 PM – 6:15 PM	To external gateways	To study area zones	**Total**
From external gateways	3,577	733	**4,310**
From study area zones	1,512	68	**1,580**
Total	**5,089**	**801**	**5,890**
6:15 PM – 6:30 PM	To external gateways	To study area zones	**Total**
From external gateways	3,430	742	**4,172**
From study area zones	1,465	73	**1,538**
Total	**4,895**	**815**	**5,710**