

A Machine-Learning Framework for Clustering and Calibration of Roadway Performance Models with Application in the Large-Scale Traffic Assignment

by

Mohammad Amin Abedini

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science
Graduate Department of Civil and Mineral Engineering
University of Toronto

© Copyright by Mohammad Amin Abedini 2022

A Machine-Learning Framework for Clustering and Calibration of Roadway Performance Models with the Application in Large-Scale Traffic Assignment

Mohammad Amin Abedini

Master of Applied Science

Graduate Department of Civil and Mineral Engineering
University of Toronto

2022

Abstract

The roadway performance model is one component of travel demand model systems, whose primary purpose is to replicate the congestion effect in traffic assignment. In this sub-model, accurate estimates of roadway travel time (delay), which is sensitive to road traffic volumes, have a principal role in properly assigning trips to various paths through the road network. Historically, one of the main challenges in developing these volume-delay models was the availability of data. Thanks to emerging data sources, the speed and volume for a wide range of roadway segments in the network are available in excellent temporal and spatial resolution. This study proposes a multi-stage machine-learning-based framework to clean, classify and calibrate roadway performance models of various roadway functional classes in the Greater Toronto and Hamilton Area. Recurrent spatial and temporal trends of roadway performance are further investigated, and distinctive patterns are observed for road segments with specific physical attributes.

Acknowledgments

I had the good fortune to collaborate with highly skilled researchers and take part in advanced projects while doing my master's degree at the Department of Civil and Mineral Engineering at the University of Toronto. The constant guidance, encouragement, and compassion of numerous persons who helped our research reach this accomplishment were important.

First and foremost, I would like to express my sincere gratitude to my supervisor Prof. Eric J. Miller for his continuous support and exceptional guidance over the past two years. Prof. Miller was more than just my academic supervisor; his knowledge, dedication, encouragement, and patience helped me tremendously and broadened my perspective on both academic and everyday life.

Secondly, I want to extend my gratitude to the StreetLight Data team for granting us the opportunity to use their on-demand mobility analytics platform. The milestone of this thesis could not have been attained without having access to their comprehensive data. I am grateful for the support Peter Stokes and Antonio Gittens gave us throughout this project, both by providing us with an academic license and by helping us out when we were using the platform.

Thirdly, I want to express my profound thanks to Prof. Marianne Hatzopoulou, my second reviewer, for taking the time to look over my thesis. Additionally, I appreciate the teaching assistantship positions that Professors Michael Seica, Baher Abdulhai, and David Hubbel offered to me during my master's studies.

Fourthly, I am thankful for all awards, scholarships and fellowships that funded my master's studies at the University of Toronto: CIV/MIN merit-based admission scholarship, Alan Tonks Scholarship in Transportation Engineering, Queen Elizabeth II Scholarship in Science and Technology, Research and Teaching Assistantships and University of Toronto Fellowships (Department of Civil and Mineral Engineering).

Due to the Covid-19 pandemic, my experience at the University of Toronto was mainly virtual, but I have some great friends and colleagues who helped me get through this difficult time. I would especially like to thank Mohammad Haghghi and Ladan Berahman for their outstanding support, which has made my academic journey joyful and unforgettable. The UT-ITE student chapter team and other colleagues in the University of Toronto Transportation Research Institute (UTTRI) and graduate student lounge also deserve a special mention.

Last but not least, I would like to share this accomplishment with my parents and siblings. I want to thank them sincerely and wholeheartedly for their constant support and everlasting love.

Table of Contents

Acknowledgments.....	iii
Table of Contents.....	v
List of Tables.....	vii
List of Figures.....	viii
Chapter 1.....	1
1 Introduction.....	1
1.1 Background and Motivation.....	1
1.2 Objectives (Research Questions).....	4
1.3 Thesis Structure.....	5
Chapter 2.....	6
2 Literature Review.....	6
2.1 Traffic Assignment.....	6
2.2 Volume Delay Functions (VDFs).....	7
2.3 Factors Influencing VDFs.....	21
2.4 History of VDFs Used for Toronto.....	23
2.5 VDFs and Emerging Technologies.....	24
2.6 Data Sources Required for Calibration.....	24
2.7 Data Cleaning.....	25
2.8 VDF Calibration Method.....	27
2.9 Inaccuracies in VDF Calibration.....	31
2.10 Stochasticity and Uncertainty.....	32
2.11 Existing Application of Machine-Learning Methods.....	33
Chapter 3.....	35

3	Study Area and Dataset	35
3.1	Study Area	35
3.2	Data Sources	37
3.2.1	StreetLight Data	37
3.2.2	GTAModel V4.0 Outputs	39
3.2.3	Road Network and Street Attributes	39
3.2.4	DMTI Land-use Data	41
Chapter 4	42
4	Methodology	42
4.1	Calibration Framework Structure and Specifications	42
Chapter 5	51
5	Results and Discussion	51
5.1	Result	51
5.1.1	Calibrated Parameters for all Sampled Road Functional Classes.....	51
5.1.2	Road Physical Attributes Distribution among Each Cluster.....	60
Chapter 6	63
6	Conclusions, Limitations, and Future Opportunities	63
6.1	Conclusions.....	63
6.2	Limitations of the Study.....	63
6.3	Future Research Opportunities	64
References	67
Appendix 1	75

List of Tables

Table 3.1-Population of the Greater Toronto and Hamilton Area (GTHA)	36
Table 3.2-Sample dataset exported from StreetLight Data.....	38
Table 3.3-Link functional class and VDF definitions, Source: (UofT Travel Modelling Group (TMG), 2017).....	40
Table 4.1-Parameter optimization setting for highways	48
Table 4.2-Parameter optimization setting for all roadways excluding highways	48

List of Figures

Figure 2.1-Travel demand and infrastructure supply interaction, Source: (Bliemer et al., 2017)..	6
Figure 2.2-Different types of traffic assignment models, Source: (Bliemer et al., 2017)	7
Figure 2.3-Typical BPR function curves with different beta values	15
Figure 2.4-Typical BPR function curves with different alpha values	15
Figure 2.5-Typical Conical function curves	17
Figure 2.6-Typical Akçelik function curves	20
Figure 2.7-Three regimes in speed vs. v/c ratio coordinate plane, Source: (Wu et al., 2020).....	28
Figure 2.8-Illustration of peak hour, congestion periods and peak period, Source: (Wu et al., 2020)	30
Figure 3.1-GTHA Municipalities, Source: (Government of Canada, 2022)	35
Figure 3.2-GTHA road network, Source: (UofT Travel Modelling Group (TMG), 2017)	36
Figure 4.1-Three-stage clustering and calibration framework.....	42
Figure 4.2-GTAModel Emme-based Road segments with V/C >0.9.....	43
Figure 4.3-Scores-loadings biplot.....	44
Figure 4.4-PCA scores plot after applying Bayesian Gaussian mixture model clustering with n_cluster=2.....	45
Figure 4.5-modified two-regime Greenshield traffic flow model calibrated with empirical data	47
Figure 4.6-Daily calibrated speed-density diagram for one road segment before and after applying modified hierarchical clustering	49
Figure 4.7-Speed-density diagram for all road segments after applying modified hierarchical clustering with n_cluster = 7.....	50

Figure 5.1- Free flow speed based on NCS16 (left) and calibrated data (right), for all VDF codes	52
Figure 5.2-Capacity based on NCS16 (left) and calibrated data (right), for all VDF codes	53
Figure 5.3- Existing and calculated values difference of free flow speed and capacity, for all VDF codes	53
Figure 5.4-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , for all VDF codes.....	54
Figure 5.5-Box plot of calibrated free flow speed (left) and capacity (right) classified by the number of lane, for all VDF codes.....	54
Figure 5.6-Free flow speed based on NCS16 (left) and calibrated data (right), VDF code: 13 ...	55
Figure 5.7-Capacity based on NCS16 (left) and calibrated data (right), VDF code: 13.....	56
Figure 5.8-Number of lanes, VDF code: 13	56
Figure 5.9- Existing and calculated values difference of free flow speed and capacity, VDF code: 13.....	57
Figure 5.10-Free flow speed based on NCS16 (left) and calibrated data (right), VDF code: 50 .	58
Figure 5.11- Capacity based on NCS16 (left) and calibrated data (right), VDF code: 50.....	58
Figure 5.12-Number of lanes, VDF code: 50	59
Figure 5.13- Existing and calculated values difference of free flow speed and capacity, VDF code: 50.....	59
Figure 5.14-Traffic flow fundamental diagrams of all sampled road segments after applying modified hierarchical clustering with n_cluster = 4	60
Figure 5.15-Clustered road segments, n_cluster = 4	61
Figure 5.16-Distribution of 'number of lanes' within each cluster.....	61

Figure 5.17-Distribution of posted speed within each cluster	62
Figure 5.18-Distribution of land use category within each cluster	62
Figure 6.1-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 11	75
Figure 6.2-Calibrated capacity (right) and free flow speed (left), VDF code: 11	75
Figure 6.3-Number of lanes, VDF code: 11	76
Figure 6.4-Existing and calculated values difference of free flow speed and capacity, VDF code: 11.....	76
Figure 6.5-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 11.....	77
Figure 6.6-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 13	77
Figure 6.7-Calibrated capacity (right) and free flow speed (left), VDF code: 13	78
Figure 6.8-Number of lanes, VDF code: 13	78
Figure 6.9--Existing and calculated values difference of free flow speed and capacity, VDF code: 13.....	79
Figure 6.10-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 13	79
Figure 6.11-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 20	80
Figure 6.12-Calibrated capacity (right) and free flow speed (left), VDF code: 20	80
Figure 6.13-Number of lanes, VDF code: 20	81
Figure 6.14-Existing and calculated values difference of free flow speed and capacity, VDF code: 20.....	81
Figure 6.15-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 20	82

Figure 6.16-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 21	82
Figure 6.17-Calibrated capacity (right) and free flow speed (left), VDF code: 21	83
Figure 6.18-Number of lanes, VDF code: 21	83
Figure 6.19-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 21	84
Figure 6.20-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 22	84
Figure 6.21-Calibrated capacity (right) and free flow speed (left), VDF code: 22	85
Figure 6.22-Number of lanes, VDF code: 22	85
Figure 6.23-Existing and calculated values difference of free flow speed and capacity, VDF code: 22.....	86
Figure 6.24-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 22	86
Figure 6.25-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 30	87
Figure 6.26-Calibrated capacity (right) and free flow speed (left), VDF code: 30	87
Figure 6.27-Number of lanes, VDF code: 30	88
Figure 6.28-Existing and calculated values difference of free flow speed and capacity, VDF code: 30.....	88
Figure 6.29-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 30	89
Figure 6.30-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 31	89
Figure 6.31-Calibrated capacity (right) and free flow speed (left), VDF code: 31	90
Figure 6.32-Number of lanes, VDF code: 31	90

Figure 6.33-Existing and calculated values difference of free flow speed and capacity, VDF code: 31	91
Figure 6.34-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 31	91
Figure 6.35-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 40	92
Figure 6.36-Calibrated capacity (right) and free flow speed (left), VDF code: 40	92
Figure 6.37-Number of lanes, VDF code: 40	93
Figure 6.38-Existing and calculated values difference of free flow speed and capacity, VDF code: 40.....	93
Figure 6.39-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 40	94
Figure 6.40-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 42	94
Figure 6.41-Calibrated capacity (right) and free flow speed (left), VDF code: 42	95
Figure 6.42-Number of lanes, VDF code: 42	95
Figure 6.43-Existing and calculated values difference of free flow speed and capacity, VDF code: 42.....	96
Figure 6.44-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 42	96
Figure 6.45-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 50	97
Figure 6.46-Calibrated capacity (right) and free flow speed (left), VDF code: 50	97
Figure 6.47-Number of lanes, VDF code: 50	98
Figure 6.48-Existing and calculated values difference of free flow speed and capacity, VDF code: 50.....	98

Figure 6.49-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 50 99

Figure 6.50-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 51 99

Figure 6.51-Calibrated capacity (right) and free flow speed (left), VDF code: 51 100

Figure 6.52-Number of lanes, VDF code: 51 100

Figure 6.53-Existing and calculated values difference of free flow speed and capacity, VDF code: 51 101

Figure 6.54-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 51 101

Chapter 1

1 Introduction

1.1 Background and Motivation

In urban regions, traffic congestion is progressively becoming a severe problem. This phenomenon is causing a loss of productivity and travel time. Traffic congestion can be defined as travel time delay experienced by road users while travelling from origin to destination. The excessive traffic demand on capacity-restrained roadways is one of many factors that contribute to traffic congestion. If we consider the transportation system as a traditional economic system, traffic demand is one subsystem being served by another subsystem which is transportation infrastructural supply. In order to mitigate the congestion in cities, both sides of this system should be analyzed, and practical actions should be taken. So, planners and decision-makers need a tool to evaluate local transport needs and decide about policy changes and investments in future infrastructure development projects. Answering transportation-, land use- and environment-related research/policy questions is not possible unless they utilize an evidence-based tool. This is where the need for travel demand models emerges for making informed decisions. These models aim to predict future travel patterns by asking "what if" questions and seeing the interaction of hypothetical travel demand and network supply. A series of mathematical models form the travel demand model, and the last step is the trip assignment. This step is the most computationally-intensive phase of the travel demand model. This step captures the interaction of travel demand and transportation supply, and performance measures are exported for further transportation policy evaluation. For the Greater Toronto-Hamilton Area (GTHA), the operational travel demand model is GTAModel, which is an activity-based and agent-based microsimulation model. GTAModel is connected to Emme software for the traffic and transit assignment part.

Travel time is the most important among the various performance measures generated by the travel demand model. This measure and other correlated performance measures, such as vehicle kilometres travelled (VKT), can help to evaluate and assess various questions about mobility, accessibility, and congestion, as well as the economic benefits and environmental impacts of improvement/expansion projects. Travel time is calculated by the volume delay function (VDF) while assigning vehicular trips to paths between origins and destinations. Travel time, or link cost

in general, is commonly driven by variables such as traffic flow (the result of assignment) and link distance, free flow speed, capacity, and a speed-flow relationship (constant parameters of the road segment). The main objective of VDFs is to reproduce realistic travel time and traffic flow in the traffic assignment and to guarantee the assignment's convergence towards user equilibrium.

The volume–delay function, one of the most fundamental and extensively used functions within traffic modelling, has experienced minimal modification since it was invented in the 1960s, despite significant technological advancements in transportation systems and travel modelling. When comparing the various types of link performance models, it is found that researchers tend to prefer roadway performance models that better represent vehicle queuing, such as the one that exists in dynamic traffic assignment. On the other hand, transportation planners are still using traditional and simplistic methods, such as the BPR function (Huntsinger & Rouphail, 2011), which is used in static traffic assignments. This gap between the widespread development by researchers and limited application by planners for practical purposes motivates us to prepare a data-driven pipeline for the clustering and calibration of roadway performance models that are utilized in both static and dynamic traffic assignments. The expectations of both parties will be met as a result of this study.

The outputs of the travel demand models will not be helpful until the supply and demand for transportation are accurately calibrated and implemented within the overall modelling framework. For the GTHA, the calibration of the demand side of the travel demand model has been a significant undertaking. However, the supply side is seldom calibrated due to the scarcity of data. For Toronto, the last calibration occurred in 1992, when the tangent function was implemented and tested in the Emme traffic assignment software (Cheah et al., 1992). Before the current study, speed-flow relationships and parameters were presumed constant over each road functional class without being calibrated with actual empirical measurements. Real-world scenarios involve many variations in volume-delay characteristics, making it appear shortsighted to adopt a general VDF for the whole city without critically investigating the performance of the different road types. In this study, VDF calibration and clustering were done to apply them in the GTAModel-based Emme assignment.

The Greater Toronto and Hamilton Area (GTHA) is well-known as one of the largest metropolitan areas in North America and is also the test bed of this research. According to statistics, the GTHA

population constituted 51.19 percent of the population in Ontario and 19.68 percent in Canada in 2021 (Statistics Canada, 2021). In addition, the GTHA's population has grown by 4.7% between 2016 and 2021 (Statistics Canada, 2021). The GTHA has faced several issues due to its rapid population development, including air pollution, housing shortages, and health care, but most notably, massive travel demand and traffic congestion on the region's transportation networks. Considerable investments have been made to upgrade transportation networks in this region. However, transportation plans and policies must be thoroughly scrutinized to guarantee that the advantages can be realized without incurring adverse consequences. One of the consequences that should be addressed before it gets worse is traffic congestion. According to the 2016 Transportation Tomorrow Survey (TTS), it has been reported that 86 percent of morning trips surveyed were made by motorized mode in the region (Malatest, 2018). Hence, since most trips are made by vehicles, creating a mobile environment and supporting more sustainable travel patterns can ease the community's movement through an urban area and reduce traffic congestion and air pollution.

GPS-enabled devices have been used extensively in recent years, generating an extensive source of digital footprint. These emerging data sources can be used to calibrate travel demand models, which can address the limitation of existing data sources. These passive location traces have great geographical and temporal coverage, making them an excellent tool for evaluating the performance of the transportation infrastructure in the study area under various operating conditions. In this study, the University of Toronto has the academic license for Streetlight Data, which is an online transportation big data analytics platform. This served as motivation for exploring the database and then using it for this study's objectives.

Given the disparity between academic and operational use of transportation supply models, as well as the need to assess ongoing and future transportation expansion projects, this study was motivated to develop a simplified yet realistic transportation performance model to help the planners to have better insight into the transportation system performance and to develop robust and adaptive plans. This research focuses on a data-driven approach and presents a multi-stage clustering and calibration framework for constructing transportation performance models. Regarding the applicability and spatial transferability, this data collecting and processing pipeline may be applied elsewhere.

1.2 Objectives (Research Questions)

This study aims to provide a theoretically solid and computationally efficient roadway performance measure for the GTHA to accurately capture cumulative vehicle arrival in the link and its propagation in the transportation system. A review and comparison of the existing highway performance metrics are needed to identify the gap and highlight potential limitations. Based on the literature review, the traditional simplistic BPR function appears to have significantly been used in practice in different parts of the world. Further review of the literature reveals that many of the existing roadway performance metrics were not even calibrated with local speed-flow data. This leads to the objective of this study, which is to develop a data-driven, well-behaved, simplified and effective volume delay function to be used in the static traffic assignment for the region. To achieve this objective, this study builds a practice-ready framework for the clustering, calibration and validation of a new form of locally derived volume delay function for different roadway functional classes in the GTHA.

With the support of the newly calibrated road-type-specific performance models, this study explores the spatial and temporal variation of GTHA transportation infrastructural supply to answer the following questions:

- What mathematical formulation best fits speed flow data points in uncongested and congested regimes?
- What is the relationship between the roadway performance model and roadway physical attributes?
- What is the accuracy of the existing speed-flow relationships?
- How roadway performance varies spatially within the GTHA?
- How long-term planned and unexpected short-term incidents are affecting roadway performance?
- Which unsupervised machine learning methods (connectivity-based, centroid-based, hierarchical-based and density-based approaches) perform better in clustering link fundamental diagrams?
- Do the existing roadway functional classes in the GTHA's Network Coding Standard 2016 (NCS16) have the same traffic performance (capacity, etc.)? If not, what makes them different? Which clusters should be recommended?

- Is it possible to use one model for several types of roadways?

1.3 Thesis Structure

This study explores emerging data sources for developing roadway performance models. The structure of the thesis is as follows: Chapter 2 provides a literature review of roadway performance measures, from the history of this concept to the existing and presently-used functions in the traffic assignment module, with the emphasis on data acquisition, calibration and evaluation of this supply measures. The study area and data preparation are described in Chapter 3, including the introduction of the multiple data sources used in this study. The descriptive statistics of the sample dataset are provided as well. To explain the structure of the calibration framework, Chapter 4 describes the structure and specification of model calibration and verification. Chapter 5 presents the final results and effectiveness of the proposed method for the sampled road functional classes as well as generalized link performance models for the region. Conclusions and highlights of findings for this study are wrapped up in chapter 6, with insight on the future avenue for further research.

Chapter 2

2 Literature Review

2.1 Traffic Assignment

Traffic assignment models are used to predict future traffic flows and travel times for long-term strategic transportation planning. Traffic assignment models constitute two sub-models: route choice and network loading (Figure 2.1). The former sub-model uses the travel demand matrix and travel time as an input and calculates the path traffic volumes. The latter sub-model use path flows and supply constraint, road geometry and traffic operating conditions to generate travel time (Bliemer et al., 2017).

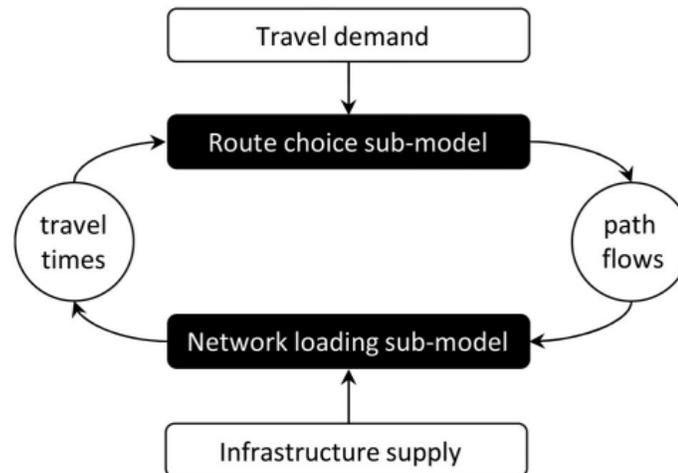


Figure 2.1-Travel demand and infrastructure supply interaction, Source: (Bliemer et al., 2017)

Traffic assignment models can be categorized into different groups in terms of their spatial, temporal and behavioural dimensions (Figure 2.2). Based on section flow capacity limit and queue storage limit, one can consider four levels of spatial dimension: unrestrained, capacity restrained, capacity constrained, and capacity and storage constrained. In terms of the temporal dimension, traffic assignment models can be categorized into three levels: static, semi-dynamic and dynamic. Regarding behavioural assumptions, one can classify the traffic assignment model into three groups: equilibrium, one shot and all-or-nothing (Bliemer et al., 2017).

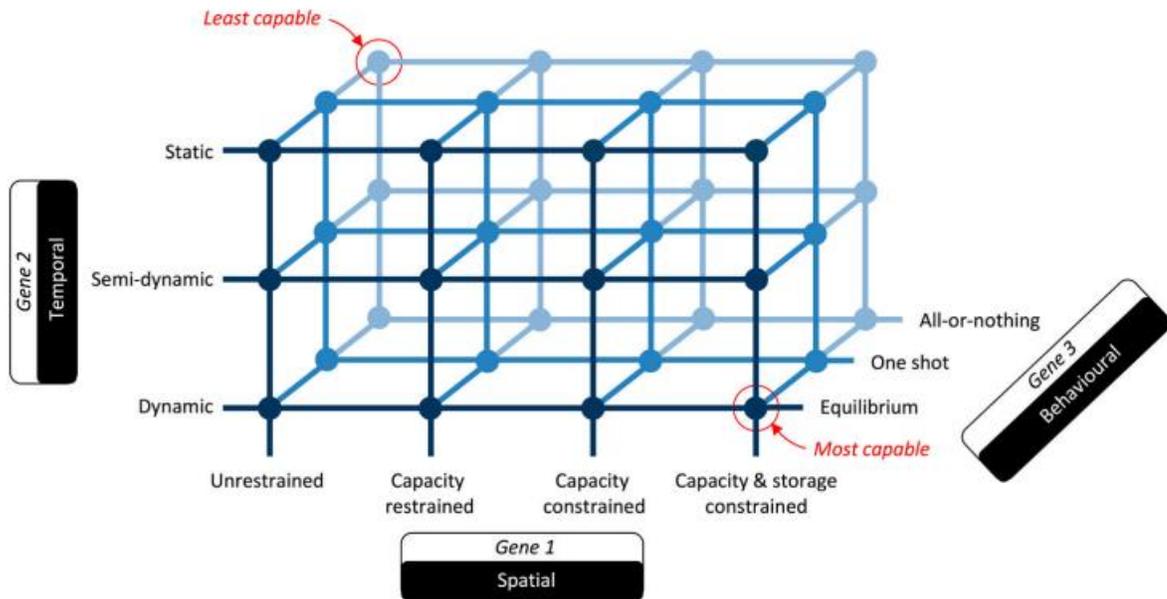


Figure 2.2-Different types of traffic assignment models, Source: (Bliemer et al., 2017)

2.2 Volume Delay Functions (VDFs)

Volume–delay functions, as the name implies, use nonlinear mathematical expressions to connect two key traffic factors. The independent variable, volume, represents the degree of traffic demand, while the dependent variable, delay, represents the slowing of traffic speeds as demand rises. The term "time delay" is a little ambiguous. A temporal baseline should be established from which the delay is calculated. In the context of vehicular journey time, this time baseline is the free flow condition in the roadway. Drivers set their speed in response to the speed limit and/or their interactions with other vehicles. When there is no barrier in the roadway, people will pick a comfortable speed, known as free flow speed, either consciously or inadvertently. Once the vehicular demand surpasses the free flow threshold, queues start to develop in the traffic stream. As congestion propagates in the link, drivers will experience a delay in their time of arrival to the destination, which is a disutility for their trip. In traffic studies and planning, we want to recreate this congestion effect in the simulated virtual environment. This has been approached in a variety of ways. Historically, as the traffic volume was closely correlated with travel time delay, a simplified model based on the macroscopic behaviour of vehicles was proposed. This macroscopic relationship is mostly referred to as a volume delay function (VDF) in the literature. This concept is also known by a variety of other names: link capacity function, volume delay relationship, travel time performance model, link cost function, delay function, link congestion function, road

resistance function, travel time-flow relationship, road impedance function, travel time estimation model, link latency function, macroscopic cost flow function.

Various formulations were proposed for this concept based on theoretical and/or empirical background (Branston, 1976). In the late 1950s, a very simplified form of linear function was proposed by (Irwin et al., 1961). Later, other researchers proposed some complex forms, e.g. exponential, logarithmic and hyperbolic, for better representation of traffic behaviour (Mosher Jr, 1963; Overgaard, 1967; Smock, 1962). Two functions suggested in the 1960s, BPR and Davidson, have gained much attention and discussion since they were proposed (Bureau of Public Roads, 1964; Davidson, 1966).

There are two main approaches to VDF formulation: mathematical and theoretical. The former approach uses the mathematical structure to replicate empirical data. This model only employs free flow speed and capacity. In some cases, traffic signal frequency and the effects of mixed traffic are considered in the formulation. In the second approach, the formulations are developed based on queuing principles and may explicitly consider the effect of signal timing and capacity on delay. These models employ free flow speed, capacity and green split parameters.

In travel demand models, VDFs serve a specific purpose. It is used as the objective function in the travel time minimization problem in the traffic assignment (Fukushima, 1984). The VDF is a mathematical model used to calculate the performance of available road infrastructure that serves city-wide trips. This allows macroscopic traffic models to capture congestion effects, which is needed for transportation supply evaluation. Throughout the modelling process, a generalized cost function is generated for each link, and after summing the discretized link costs, a total path cost between the origin and destination pairs is calculated. Then, by adjusting vehicle routing, the generalized path costs of competing paths are evenly balanced (Saw et al., 2015). In a simulated virtual environment, there are several approaches to replicate traffic behaviour. From static to dynamic and micro to macro, it all depends on the purpose of the study. The VDF function is an essential input for the standard mathematical programming formulation in static traffic assignment (Patriksson, 1994; Sheffi, 1985). To be more explicit, VDFs are employed in the macroscopic, static, deterministic, capacity-restrained, and aggregate user equilibrium traffic assignment as a fundamental highway performance model. VDFs are also utilized in agent-based modelling and semi-dynamic traffic assignment. The major goal of VDFs is to ensure that the Wardropian

algorithm converges to user-equilibrium, as well as to recreate realistic travel time and traffic flow in traffic assignment.

In traffic flow theory, the relationship between speed, flow and density has been the subject of extensive research. These fundamental traffic flow metrics could be used in traffic assignment software with some modifications. There are two significant differences between the speed-flow relationship in traffic flow and the travel time-flow relationship used in traffic assignment. First, the speed-flow relationship is not strictly increasing and lacks one-to-one mapping (it has a backward-bending relationship), which makes it unsuitable for the minimization problem in the traffic equilibrium assignment (Jastrzebski, 2000). Hence, the volume-delay function was developed to fit the requirement of those algorithms. Volume-delay function and speed flow relationship have the same function in the uncongested part, but the congested region is where they diverge. Volume-delay function does not fit the curve with the speed-flow measurement points in the congested part. Secondly, the flow utilized in planning applications is demand flow, whereas the flow used in traffic flow theory is the detected/observed flow at a road segment. The detected and demand flows are equal prior to queue formation in the roadway. These two flows are no longer the same as the queue starts to develop. Traditional traffic flow measuring systems (such as single loop detectors) can only count served traffic, not demand traffic. The highest flow values recorded in the actual world have a limit, which is the highway's capacity. The flow that is computed in the planning applications, on the other hand, has no limit. Although there is field data and following that empirical approach for the uncongested regime, a theoretical approach should be chosen for the congested regime.

Demand flow is a term used in queuing theory, which is also reflected in the VDF concept. The demand is the volume that would pass through the corridor if there were no capacity constraints. Traditional traffic measurement devices cannot determine the demand flow. There could be two methods to derive demand from processed volumes. First, the actual demand can be calculated from queueing theory. Secondly, the density ratio can be used to estimate the demand. The first method is preferred if there is appropriate data for it. If the first method is not applicable, the density ratio can be used for demand estimation (Kucharski & Drabicki, 2017).

VDFs have some drawbacks and limitations:

1. The function is highly simplified and behaviorally unrealistic and should be used only as a part of the static traffic assignment (Gentile & Noekel, 2009).
2. (Gentile et al., 2005) claimed that static traffic assignment could not replicate the actual traffic volume and travel time. Additionally, they pointed out that the VDF's goal is to ensure that the Wardropian algorithm will converge to the user equilibrium rather than to produce actual travel times.
3. VDFs assume steady flow and queues, which is rarely observed in reality.
4. Delays at intersections due to competing traffic volumes from other links are not considered by the link VDFs.
5. VDFs have a vague definition of flow over capacity. In reality, the flow cannot exceed the capacity.
6. Queue formation, propagation, and dissipation are not captured well.
7. It does not consider the conceptual and measurement uncertainty in road capacity.
8. VDFs are not spatially transferable. They do not account for the change in vertical and horizontal geometry.
9. VDFs do not account for the speeds of different types of vehicles in the traffic stream.
10. There is a scaling inconsistency between VDF parameters and variables. Parameters are estimated at the area type and facility type level, whereas the speed and capacity are estimated at the link level.
11. VDF parameters are not sensitive to time of day traffic conditions.

Besides, (Horowitz, 1991) pointed out some limitations of VDFs in the travel forecasting model:

1. Any road segment's delay results from traffic flow exclusively on that particular road segment.
2. The Frank-Wolfe method employed in the equilibrium traffic assignment cannot consider the volume of numerous road segments for calculating the delay of one road segment.
3. The delay function must follow specific criteria to be employed in the equilibrium traffic assignment, such as not having any discontinuities, strictly rising, and being analytically integratable.
4. Many travel forecasting models permit only one functional form of VDF.

5. Many travel forecasting models cannot consider turn penalties as a function of turning volume.
6. In traffic assignment, volumes may be generated that exceeds the capacity. As a result, VDFs need to be able to calculate the travel time in the situation that volume exceeds capacity.

The travel time is calculated as follows:

$$t(v) = \frac{d}{s(v)}$$

Where d is the distance, s is the speed, and v is the traffic flow. Speed is a function of traffic volume based on the traffic flow fundamental diagram. So, travel time will be a function of traffic flow. Distance is unique for each road section. Hence, this formulation can be divided into static and dynamic parts. By multiplying free flow speed (s_0) to numerator and denominator, we can write the travel time of a road section as the multiplication of free flow travel time and a function of traffic volume

$$t(v) = \frac{d}{s(v)} = d * \frac{1}{s(v)} = \frac{d}{s_0} * \frac{s_0}{s(v)} = t_0 * f(v)$$

Generally, the travel time function is usually expressed as the multiplication of free flow travel time t_0 and normalized congestion function $f(\frac{v}{c})$:

$$t(v) = t_0 * f\left(\frac{v}{c}\right)$$

$\frac{v}{c}$ is the saturation rate. The relationship between saturation rate and travel time is nonlinear. The rate of change of travel time versus saturation rate is dependable on the magnitude of the traffic flow. Before the congestion regime, travel time slowly rises, but travel time rises steeply after passing a congestion threshold. The steepness of the VDF curve in the congested regime is closely related to road design (Akçelik, 1991). In literature, researchers have set a maximum cap on the steepness rate of a VDF to avoid the risk of overestimation of travel time (Spiess, 1990).

As travel time is directly correlated with the average speed of traffic flow, the VDF formulation is recommended to be written as the relationship between speed and flow rather than travel time and

flow (Nielsen & Jørgensen, 2008). A straightforward transformation can do this, and then this transformed formulation can be used in the calibration process:

$$s(v) = s_0 * f\left(\frac{v}{c}\right)$$

Where s_0 is free flow speed and $f\left(\frac{v}{c}\right)$ is normalized congestion function.

VDFs contains link/region-specific input parameters. Capacity, free flow speed, and some coefficients can be seen in most VDFs. Free-flow speed and capacity are fixed-variable values that are defined by road functional types.

Free flow speed (FFS):

When no other vehicles are in the roadway, such as during early morning commutes, a driver's free flow speed is the speed he or she selects. There are two approaches to determining free flow speed:

1. Velocity of vehicles in free-flow traffic (low density) or a top 5% (or 15%) percentile of the travel time among all observation points
2. Velocity based on the law limit

Capacity:

Based on the HCM definition, capacity is the maximum potential/ability of a roadway under normal road and traffic conditions to accommodate vehicular traffic flow. It is difficult to determine the capacity of the road segment because of its nondeterministic nature. The considerable variation in the capacity may come from the probabilistic character of the vehicle-to-vehicle and vehicle-to-built-environment interactions. Besides driving behaviour, environmental conditions (weather and vehicle composition) may influence the stochastic nature of this variable (Das & Chilukuri, 2020). Downstream control/restriction characteristics (traffic signal or geometry) influence the capacity of a road segment (Das & Chilukuri, 2020). Capacity is assumed to be constant in the traffic assignment problem, ignoring the fact that capacity depends on the composition of vehicles on the roadway (Das & Chilukuri, 2020). Capacity in HCM is determined by road features such as the number of lanes, speed, location, and gradient. Spiess defined capacity as the volume where the congested speed is half that of the free-flow speed. Other researchers

considered capacity as the average of the top 1% flow rate measured at the bottleneck (Huntsinger & Roupail, 2011).

There are some other methods for determining the capacity:

1. Selected maxima method: This method states that the maximum observed flow determines the capacity over a period of time (Gomes & Horowitz, 2009). This method assumes that actual road capacity is rarely seen due to outside variables (such as driver behaviour and weather conditions) that lead to less-than-ideal traffic circumstances.
2. Fundamental diagram method: This approach builds the diagram using the three fundamental components of traffic flow: speed, flow, and density (Rakha & Crowther, 2002). Mathematical models will be fitted to the data to determine the maximum point at which the curve turns, which this point is considered to be capacity.
3. The product limit method: This technique identified a traffic breakdown as a steep decline in speed, which can occur when demand exceeds supply. Based on the observed flows over a period of time, this approach can estimate the capacity (Kaplan & Meier, 1958).

(Spiess, 1990) coined some rules for VDFs. VDF formulation should have these mathematical and behavioural conditions:

1. $f(x)$ should be strictly and monotonically increasing. This is required for the convergence to a unique solution.
2. $f(0) = 1$ and $f(1) = 2$; To ensure that the proposed VDF is compatible with BPR-type functions, these requirements must be met.
3. $f'(x)$ should exist and be strictly increasing. This condition ensures the convexity of the VDF.
4. $f'(1) = \alpha$; which defines the magnitude of congestion effect when the capacity is reached.
5. $f'(x) < M\alpha$; this condition is defined to limit the steepness of the congestion curve, and as a result, it limits the VDF to have high values in the congested regime.
6. $f'(0) < 0$; this condition guarantees unique values for link volumes.
7. The computing time of evaluation of $f(x)$ should not be greater than the evaluation of corresponding BPR functions.

Where $f(x) = f\left(\frac{v}{c}\right)$.

There are several efforts in the literature to develop suitable VDFs. BPR (Bureau of Public Roads, 1964), Davidson (Davidson, 1966), Conical (Spiess, 1990), kcelik (Akçelik, 1991) and empirically modified BPR variations (Moses et al., 2013) are the most common variants of VDF.

BPR function:

The BPR (Bureau of Public Roads) function, developed in the late 1960s by fitting data collected in uncongested and uninterrupted freeways, is the most widely used VDF in literature and practice (Branston, 1976; Bureau of Public Roads, 1964). Even though it has several drawbacks, it is employed in studies with less attention to detail and low traffic demand in the network. Its widespread use can also be attributed to its simple formulation and minimal input parameters requirement (Mtoi & Moses, 2014). This common VDF is parabolic in shape, and its formulation is as follows:

$$t(v) = t_0 * [1 + \alpha * (\frac{v}{c})^\beta]$$

Where t_0 free-flow travel time, α and β are coefficients to be estimated by field data, v is the traffic volume and $\frac{v}{c}$ is the volume-to-capacity ratio (degree of saturation). Originally, α and β are recommended to be 0.15 and 4. These two variables are unique to the environment and geometry of the road. Parameter α indicates the ratio of free-flow travel time to the travel time at capacity. Parameter β determines the rate of change of average travel speed from a free-flow regime to a congested condition (Mtoi & Moses, 2014). Higher β values will lead to more sudden onset of congestion (as illustrated in Figure 2.3), whereas higher values of α indicate much worsening conditions with increasing traffic flow (as shown in Figure 2.4). Furthermore, β defines the curvature of the BPR function, with $\beta \leq 1$ indicating concavity and $\beta > 1$ indicating convexity.

The result of BPR might not be accurate in highly congested road segments because this function was developed based on traffic data of an uncongested state of a highway in its early implementation stage (Skabardonis & Dowling, 1997; Spiess, 1990). Hence, since proposing BPR function, local traffic authorities across the world calibrated the BPR curve coefficients by empirical and/or simulated data to suit their local needs.

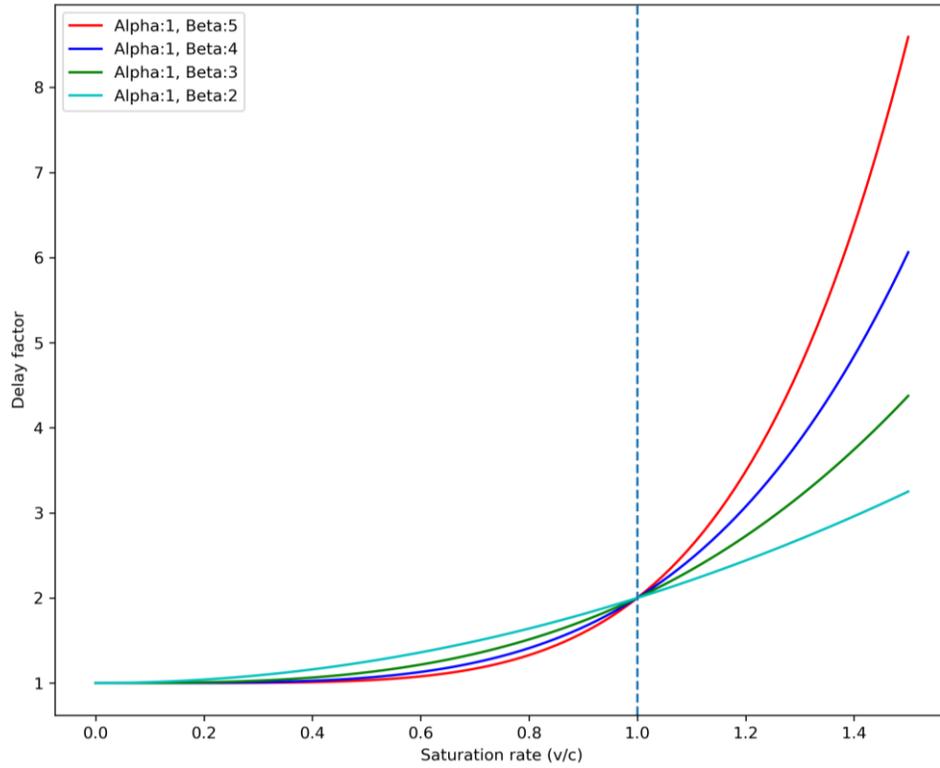


Figure 2.3-Typical BPR function curves with different beta values

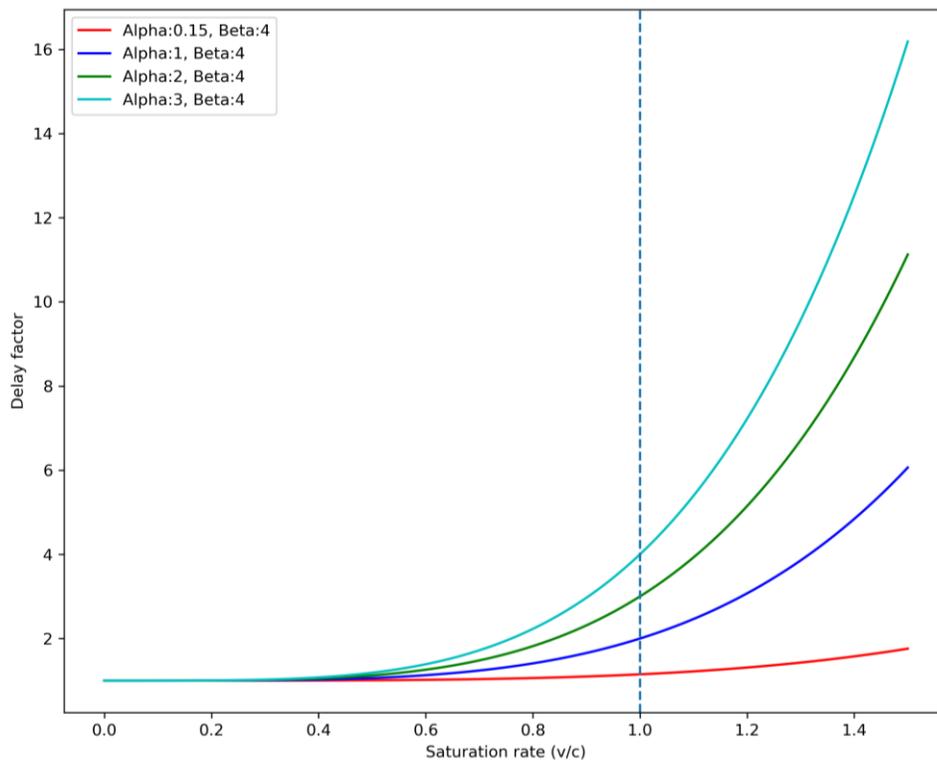


Figure 2.4-Typical BPR function curves with different alpha values

BPR has some drawbacks. Researchers have summarized that the BPR function underestimates travel time in an uncongested regime and overestimates travel time in a congested regime (Akçelik, 1991; Singh & Dowling, 1999). Also, it is noted that the BPR function overestimates traffic volumes while volume exceeds capacity and underestimate traffic volume for free-flow condition (Singh & Dowling, 1999). (Akçelik, 1991) pointed out that by adjusting the parameters of the BPR function, the travel time may become constant in the uncongested regime, whereas it becomes sensitive to volume in a congested regime. Also, it is noted that using the BPR function can lead to extreme values for the link travel time in early assignment iterations. In mixed traffic situations, where several vehicle classes occupy the road in non-lane-based movements, BPR is not suitable (Das & Chilukuri, 2020). Furthermore, different types of traffic facilities (e.g. traffic signals) and traffic operating conditions are not considered in the derivation of the BPR function (Boyce et al., 1981; Skabardonis & Dowling, 1997). In other words, the BPR function is not inspired by queuing principles.

To better capture operational traffic conditions and signalization effects, other VDF formulations have been suggested. Some examples of these functions include Conical, Davidson, and Akçelik. These alternative functions differ by the formulation, input parameters and shape of the curve. The most prevalent metrics in choices are free flow trip time and segment capacity, although other features, such as signal timing parameter, are also taken into account.

Conical function:

The conical VDF is another alternative to the BPR function. This function was created to address issues with the high value of beta parameter in the BPR function. The issue was overloading the link during the first few iterations of an equilibrium assignment (Spiess, 1990). Conical function's name is derived from its hyperbolic conical geometry, which is the intersection of two 2-dimensional planar surfaces and a 3-dimensional cone. This function meet all seven criteria mentioned before for a well-behaved congestion function. This function is developed by Spiess as follows:

$$t(v) = t_0 * \left[2 + \sqrt{\alpha^2 * \left(1 - \frac{v}{c}\right)^2 + \beta^2} - \alpha * \left(1 - \frac{v}{c}\right) - \beta \right]$$

$$\beta = \frac{2\alpha - 1}{2\alpha - 2} \text{ \& } \alpha > 1$$

Both BPR and Conical use similar inputs. α and β are variables consistent with similar parameters in the BPR function. In contrast to the BPR function, only one coefficient is required to be estimated in the conical function. A typical conical function with different α values is presented in Figure 2.5. The conical delay function is based on a mathematical derivation that uses basic geometry and elementary algebra rather than empirical data. The conical volume–delay function, like the BPR function, does not explicitly account for signalization effects.

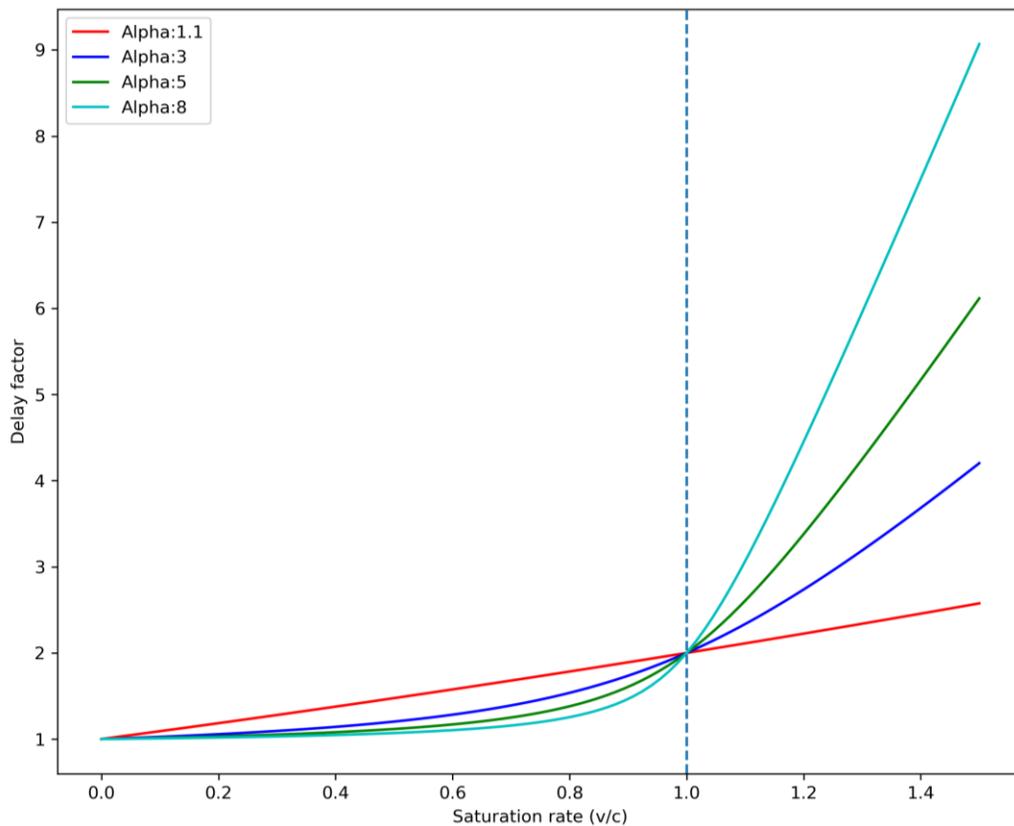


Figure 2.5-Typical Conical function curves

Davidson function:

(Davidson, 1966) brought queueing theory into play and modified the well-known steady-state delay equation, which is as below:

$$t(v) = \frac{1}{c} + \frac{\frac{v}{c}}{c * (1 - \frac{v}{c})}$$

This equation consists of two parts: the first term is for service time; the second represents the queuing delay. Davidson multiplied J_D as a delay factor to the second term. Also, Davidson assumed that service time ($\frac{1}{c}$) equals to free flow speed (t_0). After implementing these changes to the original formula, the Davidson function is proposed as follows:

$$t(v) = t_0 * \frac{1 + J_D * \frac{v}{c}}{1 - \frac{v}{c}}$$

Where J_D is the delay parameter ($1 - J_D$ can represent the quality of service). For Davidson's function, (Taylor, 1977) provided a technique for estimating the parameters and explored the sensitivity and reliability of his method. Compared to the BPR function, there is some inconsistency in Davidson's function (Golding, 1977). For instance, capacity is defined as the inverse of free flow travel time and saturation as volume times free-flow travel time. To cover these inconsistencies, Davidson modified the function and redefined the delay parameter (Davidson, 1978). This modification resulted in another inconsistency: better service quality with increased free-flow travel time (Akçelik, 1991). Besides this, there was one problem with Davidson's delay model: it could not define the travel time when volume exceeds capacity (generating negative travel times). It estimates an infinite travel time when flow equals capacity. Hence, Davidson's delay model was further modified to cover this flaw (Akçelik, 1981). (Taylor, 1984) proposed a network equilibrium assignment that used a modified Davidson's function. (Tisato, 1991) reported that when the delay parameter in the modified Davidson's function gets closer to its limit of one, the slope of the linear extension behaved poorly.

Akçelik function:

On top of the Davidson VDF, Akcelik proposed a time-dependent form of Davidson formulation to encompass the intersection delay (Akçelik, 1991). Akcelik used coordinate transformation to address the issues of inconsistent parameter definition and overestimation of travel time around

the capacity flow caused by Taylor's modification of Davidson's function (Akçelik, 1991). Here is the formulation:

$$t(v) = t_0 + 0.25 * T_f * \left(\frac{v}{c} - 1 + \sqrt{\left(\frac{v}{c} - 1\right)^2 + \frac{8 * J_A * \frac{v}{c}}{c * T_f}} \right)$$

Where T_f is flow period value that shows the period of flow analysis, J_A is a delay parameter that can be estimated by nonlinear regression. Akcelik provides a formula to have a rough estimation of the delay parameter (J_A):

$$J_A = \frac{2c}{T_f} (t_c - t_0)^2$$

Where t_c is the value of travel time at capacity. The rest of the variables have been defined before. Akcelik function can be used for individual road segments and road segments ending at a signalized intersection. To account for the delay in the intersection, T_f has lower values for freeways/coordinated signal roadway and higher values for arterial roads without signal coordination. During the analysis period, it is assumed that the average arrival demand flow rate remains constant and that there is no beginning queue. The Akcelik function has better convergence in the assignment algorithm and generates a more realistic speed in a congested regime. The Akcelik function can handle both free flow and congested regimes by using the Webster formulation for the undersaturated regime and an oversaturated model for the congested regime. Figure 2.6 represents a typical conical function with different parameters.

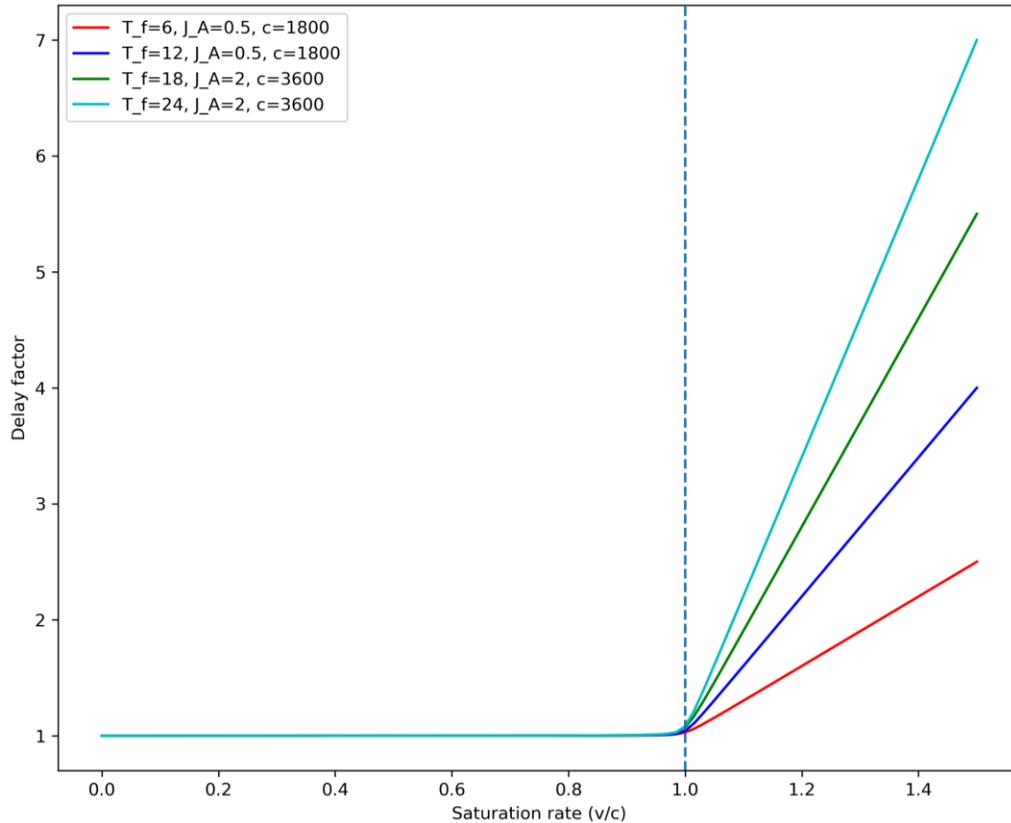


Figure 2.6-Typical Akçelik function curves

Comparison of VDF models:

VDFs can be compared from multiple perspectives, including shape, steepness, predicted speed and underlying theoretical concepts (Saric et al., 2019). Most discussions are on the reliability of these models in the congested regime. The curvature of the VDF curve in congested conditions and the steepness of this curve near the congestion threshold are the two factors that need further examination. According to the traffic flow fundamental diagram, the relationship between volume and speed must be convex. Although VDFs cannot depict the bottom half (congested part) of a volume-speed diagram, some VDFs do reflect the convexity of the fundamental diagram's uncongested part. With default parameters and calibrated values, the BPR and conical functions are defined to be convex. They can, however, have a concave form even if the default values are not used ($\beta \leq 1$ in case of BPR function). Because there is not enough data to calibrate the congested part, the form of VDF when the flow exceeds capacity mostly depends on the modeller's intuition. Another comparative criterion is the steepness of the VDF around the congestion threshold. In certain models, e.g. Akcelik and Conical, a significant decline in speed can be seen

at $v/c = 1$, whereas others show a minor speed change at this point. According to Spiess' comparison of the BPR and Conical functions, Conical is preferable to BPR in every perspective, including computation time, convergence, and flexibility (Spiess, 1990). Field investigation should be done to see how reliable the existing VDFs are.

2.3 Factors Influencing VDFs

VDFs capture the effect of congestion. According to studies, there are additional factors impacting travel time in addition to the number of vehicles traversing the roadway. According to the Federal Highway Administration (FHWA), six sources of congestion are: bottlenecks, poor signal timing, special events, construction zones, severe weather, and traffic accidents. In addition to these congestion sources, physical and operational characteristics of the road environment and road users, such as vehicle composition, road layout, geometry, and land use, may affect the journey time. Some factors influencing the volume-delay relationship for a roadway are discussed below.

Vehicle type:

Generally, VDFs are developed for a relatively single-mode homogenous traffic environment, in which vehicles are mainly motorized (F. Zhao et al., 2020). However, traffic flow in a single road segment may be heterogeneous, and its composition may vary spatially and temporally. Vehicles differ in type, size, engine power, maneuverability, etc. Different vehicle classes can traverse the roadway, including but not limited to motorized two-wheelers, three-wheelers, four-wheelers, bicycles, light commercial vehicles, and heavy commercial vehicles. This difference in physical and operational characteristics of vehicles may result in various inter- and intra-vehicular interactions and non-lane-based movement. This behaviour may result in a wide range of observed speed in the roadway, in which heavy vehicles contribute to the lower range and fast-moving vehicles constitute the higher range. Additionally, compared to an auto, an oversized vehicle, e.g. truck, occupies much more of the road's capacity. Furthermore, due to safety issues, auto drivers may keep a greater distance from a truck than an auto (Yun et al., 2005). Last but not least, compared to the condition where all road users are automobiles, trucks have considerably longer acceleration and deceleration times, which causes far greater deterioration of congested speed (Moridpour et al., 2015; Yun et al., 2005). In order to take into account all of these issues on roadway capacity and speed, (the Highway Capacity Manual, 2000) recommends using a passenger-car-unit (PCU) approach for heavy trucks. Other researchers have proposed modified

functions developed based on real traffic data, theory, and/or microsimulations (Lu et al., 2016; Moridpour et al., 2015; Müller & Schiller, 2015). (Das & Chilukuri, 2020) proposed a link cost function for mixed traffic condition, developed based on the kinematic wave model. They validated the function with the 12 signal cycles from two different signalized intersections. (Yun et al., 2005) developed updated volume delay relationship based on microsimulation results that accounts for the mix of heavy truck traffic. (Šarić & Lovrić, 2021) proposed a function that considers traffic flow heterogeneity and road geometry characteristics. They pointed out that the combined effect of traffic composition and geometric features (longitudinal grade and curvature) significantly impact the roadway's average speed. (Thomas et al., 2012) constructed a multi-regime and class-wise speed—flow relationship for a mixed traffic environment, which captures asymmetric and unequal interactions between different pairs of vehicle classes.

Weather:

Weather conditions may have a significant impact on the operation of traffic systems. (Maze et al., 2006) identified three categories that can be affected by weather conditions: traffic safety, traffic flow relationship and traffic demand. Inclement weather can affect drivers' perceptions in a variety of ways. Road users may reduce their speed due to reduced visibility and pavement friction. On a microscopic level, it is well understood that higher precipitation lengthens trip times due to increased danger (Mashros et al., 2014). (Maze et al., 2006) studied the freeway system and pointed out that adverse weather causes a significant reduction in traffic speed: up to 6% for rain, 13% for snow, and 12% for low visibility. (Ibrahim & Hall, 1994) analyzed the effects of rain and snow on the speed-flow-occupancy relationship on Canadian roads and found that snow has much more effect than rain and can cause up to 50 km/h reduction in free-flow speed. (Li et al., 2021) studied the effects of weather conditions (rain intensity and visibility) on travel-time prediction models using soft set theory. (L. Zhao & I-Jy Chien, 2012) studied the impact of adverse weather condition on traffic speed and travel time reliability and found that under adverse weather, traffic speed dropped, and traffic congestion began to form. (Caceres et al., 2016) considered the impact of various factors, e.g. weather, time of day, accidents, on travel time.

Road incidents:

Road network disruptions, such as roadworks and traffic crashes, can cause a road link's capacity to be reduced or even eliminated, leading to longer travel times.

Road layout, geometry and land use:

Vehicle interactions may become more complicated due to different road layouts, geometry and land use. For instance, the road's horizontal alignment and longitudinal grade and pavement condition might affect vehicle speed. Another element influencing vehicle speed is land use near the road segment (e.g. drivers tend to slow down their speed while passing by schools or leisure centers). The presence of on-road parking and bus lay-bys on the roadway may have an impact on its performance.

2.4 History of VDFs Used for Toronto

For the Emme assignment used in GTHA, an ad-hoc modification to the BPR function known as the tangent function was developed. This function acts exactly like the BPR function up until the volume equals the capacity. When the volume exceeds the capacity, it extrapolates the BPR function using a linear delay function with the BPR function's slope in the $v=c$. Because of its same performance in predicting base case travel time and traffic volumes as BPR, its faster convergence than BPR, and its theoretical backbone (simplified deterministic queuing theory for the oversaturated region), they chose the tangent function instead of BPR for the region.

$$t = \begin{cases} t_0 * \left(1 + \alpha \left(\frac{v}{c}\right)^\beta\right) & \text{if } \frac{v}{c} < 1 \\ t_0 * \left((1 + \alpha - \alpha\beta) + \alpha\beta * \frac{v}{c}\right) & \text{if } \frac{v}{c} > 1 \end{cases}$$

Where α is normally 1 for one-hour assignments. β is 4 and 6 for arterial roads and freeways, respectively.

The findings of a comparison study on six variants of VDF for the Emme assignment used in Toronto revealed that the VDF functions are ranked as follows, from greatest to least CPU time savings: horizontal half-tangent, asymptotic, tangent, conical, and BPR (Cheah et al., 1992).

The VDF concept and functional classes in NCS16¹ are derived from various sources, including NCS11, the Geometric Design Guide, and GGHMv4². The NCS11 and/or GGHMv4 standards were used for lane capacity because the capacity value was not included in the Geometric Design Guide. The tangent function is employed as the VDF in the current GTAmode. In the uncongested regime, this function is BPR. However, when the volume exceeds the capacity, it becomes an inclined line with a slope of the BPR curve in the v/c equals 1. GTAmode applies the posted speed limit as the free flow speed in the tangent function. One exception is freeways, where the posted speed plus 10 km/hr is used in modelling. In the following table, nominal link capacity is shown.

NCS22 was still undergoing final edits at the time of this thesis. Therefore, the analysis of this study relied on the final version of NCS16.

2.5 VDFs and Emerging Technologies

Static traffic assignment is employed in the strategic modelling of new mobility alternatives. Therefore, VDFs may need to be updated to consider these new transportation technologies. To be more precise, connected and autonomous vehicles are one of those significant changes that may alter how we define the capacity of current motorways. The capacity and throughput will both rise with the penetration rate of these vehicles. Researchers compared the microsimulation dataset with the results of the BPR function and concluded that in some circumstances, the BPR parameters should be modified based on the different CAV penetration rates (Qiu et al., 2022).

2.6 Data Sources Required for Calibration

An extensive set of data from single or multiple sources are used in VDF calibration (Hansen & Bertini, 2005; Raw & Munn, 2009; Z. Wang & Liu, 2005). The main inputs for calibration are traffic speed and volume. There are several methods to measure travel time and traffic flow from

¹ The 2016 Network Coding Standard (2016) was developed by the University of Toronto Travel Modelling Group for use by member agencies in the GTHA. See TMG (2017).

² The Greater Golden Horseshoe Model, Version 4, developed by the Ontario Ministry of Transportation.

a traffic stream. Loop detectors, radar, and video surveillance systems are the common traffic measuring devices, most of which are found on limited-access highways. On the other hand, collector and urban arterial roads have fewer traffic measurement equipment. Other than these in-place instruments, data collected via location-based and crowd-sourced mobile sources can be utilized to support the VDF calibration procedure. If no observed data is available, microsimulations could be utilized to replicate the actual traffic situation and generate synthetic data for VDF curve fitting (Dowling & Skabardonis, 1993; Skabardonis & Dowling, 1997).

A traffic counter is a common way to measure traffic flow and volume. On the road, these counters are either permanently or temporarily installed. One of the difficulties is obtaining travel time data in addition to traffic volume from these devices. Single loop detectors are unable to provide any information about speed. Instead of point measurement, one method is to place two loop detectors close together to detect the highway's instantaneous speed, which still does not represent the delay vehicles encounter on the whole road segment but can give an approximate estimate of the link's travel time. Other data sources can be used and fused with the traffic count data in addition to this method. For instance, the traditional license plate-matching method can also be used to find the average travel time by the matched vehicles (Xiong & Davis, 2009). In addition, a GPS logger mounted on a probe vehicle can provide us with the speed profile of a road segment, which can give us an estimate of the travel time. These data sources are well enough for highways, but signal-related inputs are also needed for road segments that end at intersections to obtain a reliable travel time estimate.

Data resolution in terms of time and space is also critical. Most scenarios in travel demand models are built for the morning rush hour. Aggregating speed-flow with a 1-hour interval or finer resolution can aid in the creation of an appropriate volume-delay function. Traffic data from a wide range of roadway functional classes with great geographical coverage can also help in the development of context-specific VDFs by the area and functionality type.

2.7 Data Cleaning

The calibrated traffic parameters of a road segment may be misrepresented as a result of anomalies and outliers. Because of this, these atypical data points need to be cleaned before the dataset is

imported into the calibration step. There are various statistical and machine learning techniques to detect and eliminate the anomalies, some of which need single loop detector data, while others require data from a series of loop detectors.

For a single loop detector, comparing speed, flow and occupancy data points of one time intervals to those of its preceding and subsequent time intervals can help to detect traffic breakdowns. Short observation durations of 1 or 5 minutes are reasonable to notice a breakdown between intervals. A breakdown is a traffic state which occurs when a stable traffic condition becomes unstable. There are multiple ways to determine the breakdown from speed-flow time series data. To begin, the breakdown might be detected using a minimum speed difference. Second, a minimum traffic flow should be specified to avoid the detection of false nighttime breakdowns. The literature recommends a speed difference of 10 km/h and a minimum traffic flow of 600 vehicles/lane/hour (Neuhold & Fellendorf, 2014). Other researchers have compared all data points with one another on a speed-density or density-flow plane and identified irregularities. (Dudek et al., 1974) proposed an approach, named as the standard normal deviate (SND) method, which uses mean value (μ) and some measure of variation (σ) to detect if a data point is below some threshold value of $\mu - c\sigma$. (Kalair & Connaughton, 2021) used similar approach with kernel density estimation to construct a bivariate distribution of density-flow and determined that outside the 95% of the probability curve is strongly correlated to significant congestion events. (Casey et al., 2020) applied a manually tuned DBSCAN method to eliminate outlier points from low-density areas in the speed-density plot. (Yuan et al., 2018) used the convolutional-LSTM model to forecast the time and location of traffic accidents. Other researchers utilized a convolutional neural network for anomaly detection (Zhu et al., 2018, 2019).

For a series of loop detectors, California algorithms and variants are used, which employ basic pattern matching and comparing occupancy values to indicate the congestion event window (Levin & Krause, 1978; Payne et al., 1976). A similar approach is the McMaster algorithm, which section the occupancy-flow diagram into uncongested, bottleneck flow and congested regimes and use particular thresholds to find the traffic breakdown between one station and its upstream and downstream stations (Persaud & Hall, 1989).

2.8 VDF Calibration Method

Calibration is the process of obtaining a mathematical expression from real-world data. Three criteria should be noted in calibrating roadway performance metrics: nonlinearity, heteroskedasticity, and site-specificity. Nonlinearity, in particular, means that the rate of change of travel time delay varies with the level of traffic flow. The term heteroskedasticity refers to the variation of the variance of the independent variable, which is the speed in the context of VDF calibration. With an increase in traffic volume, more variation in vehicles' speed is captured due to the stochastic nature of vehicular interactions. In order to account for site-specificity, the speed-flow data should be calibrated for a sample of road segments that includes all road functional classes with a variety of physical attributes in the region of study.

VDF calibration is done in two parts. The proper and well-behaved formulation should be chosen first, followed by the statistical estimation (regression analysis) for calibrating the function's parameters. In common practice, implementing one of the existing functions and adjusting the default parameters based on national standard guidelines appear to be typical, with little to no attention paid to the resulting speed outputs. However, in limited advanced practices, field data is used to develop, adjust and verify unique VDF.

The main challenge for the calibration of VDFs is the measurement of demand in a congested regime. For the uncongested part of a VDF ($v/c < 1$), we can evaluate and calibrate the speed-flow relationships by empirical observations. However, as conventional traffic count data is not adequate for the calibration of the congested part ($v/c > 1$), a theoretical approach, e.g. arbitrary formulation, should be taken under consideration (Jastrzebski, 2000). To be more precise, bottleneck analysis and queue length estimation can be used to estimate demand beyond capacity for calibration of the congested part of VDFs (Huntsinger & Roupail, 2011). In roads ending at the signalized intersection, demand can be measured from the queue length and discharging flow at the stop line. However, in roads without signalized intersections, e.g. highways, a recurring bottleneck can be found, and demand can be measured on those spatial constraints. It should be noted that different queue scenarios will contribute to different shapes and slopes of VDF in congested part.

Formulation-specific coefficients and road characteristics are the two sets of parameters that can be estimated in the calibration stage. VDF calibration can be done using a variety of optimization

techniques. The least-square approach is traditionally used to calibrate the travel time-flow data points. This approach minimizes the sum of square errors, and coefficients are estimated. The weighted least squares approach is one step ahead and can accommodate the variation in the variance of the speed values. One researcher proposed a genetic algorithm method for VDF parameter calibration (Cetin et al., 2012). Regarding the input variables in the minimization, a partial or complete set of coefficients can be imported in the calibration step. Calibration can also be done with a base model with default coefficients. There is also the option of not calibrating the model and instead relying on traditional lookup tables for coefficients and parameters. In the optimization problem, in order to force the volume-delay curve to be convex, some constraint needs to be imposed in the calibration stage.

As mentioned, the main challenge in link travel time estimation models is mapping speed-flow data from the observable congested regime (B) to the unobservable congested regime (C), illustrated in Figure 2.7. There are five major representations of volume-to-capacity ratios in the VDF estimation. In the formulations below, $\frac{v}{c}$ is represented as y .

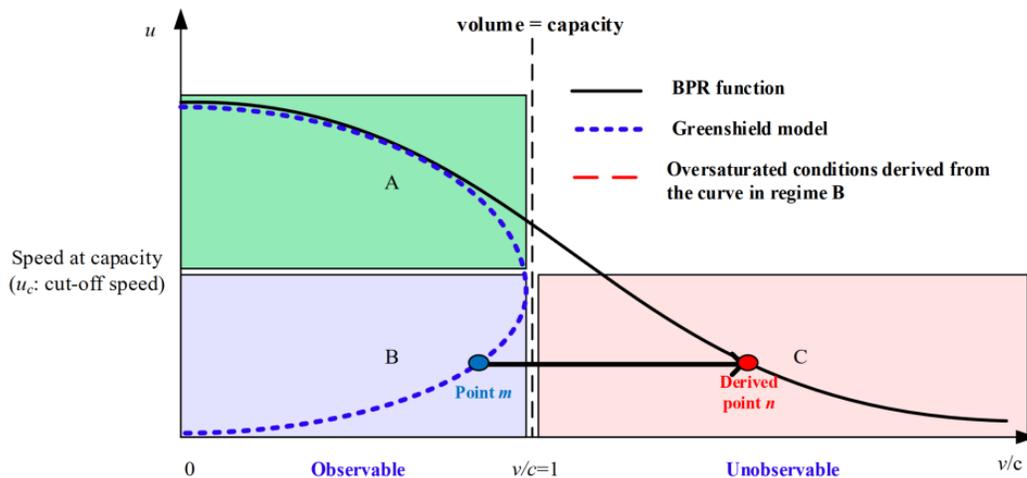


Figure 2.7-Three regimes in speed vs. v/c ratio coordinate plane, Source: (Wu et al., 2020)

1. Volume-based method (VBM):

The VBM is taken from the Florida Standard Urban Transportation Modeling Structure (FSUTMS) (Moses et al., 2013). In the congestion regime, this method mirrors the speed-flow data point by $v/c=1$ line.

$$y = \begin{cases} \frac{q}{c}, & \text{if } t \in \text{uncongested time period} \\ \frac{c + (c - q)}{c}, & \text{if } t \in \text{congested time period} \end{cases}$$

2. Density-based method (DBM):

In DBM, the travel time-flow calibration of VDF turns into travel time-density estimation. This has been done because the speed-density relationship is monotonic in congested cases, which resembles the same form as the delay-volume curve. (Kucharski & Drabicki, 2017) used the hydrodynamic relation of the fundamental diagram to calculate quasi-density from time-mean speeds and flows. Hence the transformation of volume over the capacity ratio to the quasi-density ratio will be as follows:

$$y = \frac{k}{k_{(q_{max})}} = \frac{\frac{q}{s}}{k_{(q_{max})}}$$

Where k is quasi-density, $k_{(q_{max})}$ is density-at-capacity, in which these two parameters are the substitution for flow and capacity at any time, respectively. In addition, q is observed traffic flow, and s is observed traffic speed. In order to find $k_{(q_{max})}$ from the observed speed-flow data, the author recommended considering capacity as the 95th percentile of 15-minute flow and calculating the corresponding density. Also, there is an option to consider $k_{(q_{max})}$ as a variable to be estimated in the optimization problem.

3. Queue demand-based method (QBM)

(Wu et al., 2020) proposed peak-period-based calibrating framework based on the freeway bottleneck modelling perspective.

$$y = \begin{cases} \frac{D_h}{c}, & \text{if } u'_{min} > u_c \\ \frac{D}{c}, & \text{if } u'_{min} \leq u_c \end{cases}$$

Where D_h is the volume within the peak hour, D is the volume within the congested period, c is the ultimate capacity, u'_{min} is highest speed within the peak hour, and u_c is the cut-off speed. Figure 2.8 illustrate different periods proposed in this method.

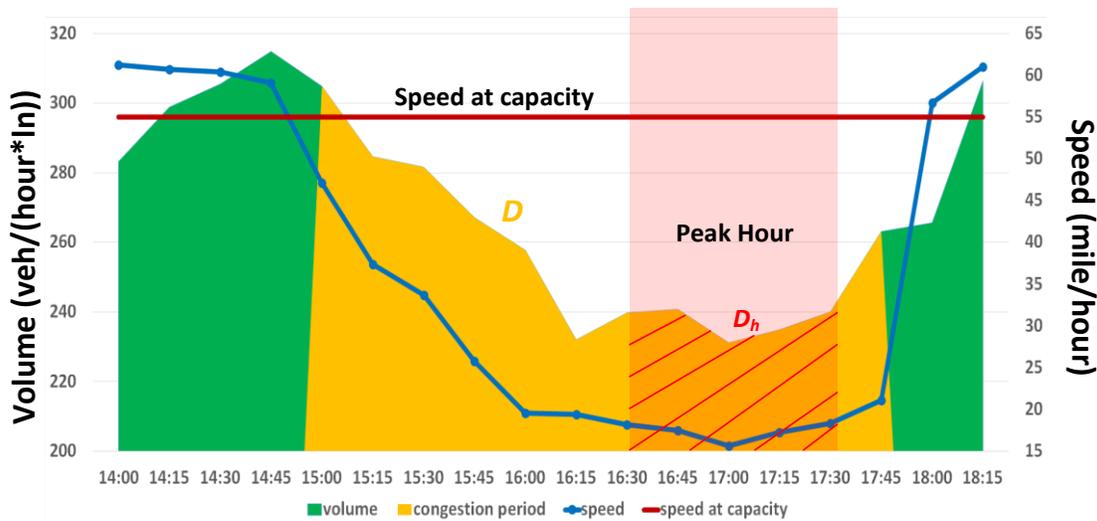


Figure 2.8-Illustration of peak hour, congestion periods and peak period, Source: (Wu et al., 2020)

4. Time-dependent queue-based D/c model (TDQDC)

(Huntsinger & Rouphail, 2011) proposed a time-dependent queue-based D/c model (TDQDC) to estimate the demand beyond capacity as follows:

$$y = \begin{cases} \frac{q}{c}, & \text{if } t \in \text{uncongested period} \\ \frac{D(t)}{c}, & \text{if } t \in \text{congested period} \end{cases}$$

$$D(t) = c + \text{queue}(t)$$

Where $D(t)$ is an average measure of queued demand and capacity at time t , c is the ultimate capacity and $\text{queue}(t)$ is the queue at time t . $\text{queue}(t)$ is equal to the density at time interval t for a detector multiplied by the influence area of the detector.

5. Time-dependent cumulative demand flow-based travel time method (TDCD)

TDCD method is a rolling horizon model which generates one sample point for each interval. Its formulation is as follows:

$$y = \begin{cases} \frac{q}{c}, & \text{if } t \in \text{uncongested time period} \\ \frac{D(\tau)}{\mu(\tau)}, & \text{if } \tau \in \text{congested time period} \end{cases}$$

Where q is flow rate, c is capacity, $D(\tau)$ is the cumulative demand during time interval τ , and $\mu(\tau)$ is the discharge rate during time interval τ . This author analyzed the correlation between flow of bottleneck, its upstream and downstream to generate accurate discharge rates for different traffic conditions.

2.9 Inaccuracies in VDF Calibration

There are four reasons for inaccuracies in traffic flow parameters calibration: sample selection bias, functional form of the model, the variance of observed data, and singular value estimation.

Unbalanced data points may result in an inaccurate calibrated function. Sample selection bias is the term that is used for this error. It is expected that most data points will have low flows and near free-flow speeds outside the peak hour. Hence, data points in an uncongested regime have more contribution to the final calibrated curve, which will result in misestimating the congested regime. There are two strategies in the literature to overcome this issue. First, a low volume data sampling approach may be utilized. The two methods that may be applied in this strategy are interval-based sampling and threshold-based sampling. In interval-based sampling, the axis of the dependent variable is divided into equal bins, and in each bin, an equal number of the sample will be randomly chosen (Zhang et al., 2018). In threshold-based sampling, the axis of the dependent variable is divided into two parts based on a threshold line, and an equal number of samples will be selected in each part. The second option is the weighting strategy. Weights are established using the weight determination method. The observations are ranked based on the dependent variable and are weighted based on their own value and their preceding and following observations (Qu et al., 2015).

In order to minimize the inaccuracy due to the functional form of the traffic flow relationship, all possible mathematical forms should be tested on the data, and the best one with the acceptable goodness of fit should be chosen.

As mentioned before, most data sources may report the average speed for each observation interval. Limited data sources may report each vehicle's speed or the variance of observed speed in each time interval. The speed variance can be imported as the weights in the weighted least square method (WLSM). In this way, the heteroskedasticity of speed is considered in the calibration procedure. The objective function is as follows:

$$\min Z = \sum_{i=1}^P \bar{w}_i (v_i - \hat{v}(k_i, m))^2$$

Where v_i and $\hat{v}(k_i, m)$ are observed and estimated speeds corresponding to density k_i , and w_i is the weights based on the inverse of the variance of speed for each observed data.

In the literature, most traffic flow parameter calibration is done only based on one dependent and one independent variable. If we assume that the noise of the observations is independent of the traffic flow and speed, we can use speed-flow-density joint estimation to calibrate the parameters (Cheng et al., 2021). The objective of the joint estimation is the minimization of the sum of square errors in both speed-density and flow-density relationships as follows:

$$\min Z = \sum_{i=1}^P [(v_i - \hat{v}(k_i, m))^2 + \delta * (q_i - \hat{q}(k_i, m))^2]$$

$$\delta = \frac{\sigma_v^2}{\sigma_q^2}$$

Where v_i and $\hat{v}(k_i, m)$ are observed and estimated speeds corresponding to density k_i , q_i and $\hat{q}(k_i, m)$ are observed, and estimated flows corresponding to density k_i , σ_v^2 and σ_q^2 are the variances in the speed and flow measurements, respectively. σ_v^2 and σ_q^2 can be calculated based on the observed data, or they can be estimated in the optimization problem as additional parameters.

2.10 Stochasticity and Uncertainty

Transport models are subjected to uncertainty since they describe complex systems. This uncertainty has the potential to affect all model components, as well as the model outputs. There might be two sources of uncertainty. On the one hand, the uncertainty might emerge from the quality of the traffic data or the model structure itself, known as epistemic uncertainty. On the

other hand, the variability and randomness of traffic behaviour might cause another type of uncertainty in output parameters known as aleatory/ontological uncertainty. By having more accurate data or employing alternative models, the former uncertainty can be reduced. The latter uncertainty can be addressed by statistical analysis, which considers stochastic behaviour. Driver behaviour as an internal factor, time of the day, weather condition, and link characteristics as external factors can lead to stochastic behaviour of the true parameters of VDF. In order to quantify the ontological uncertainty, model results should be given as a distribution rather than a single-point estimate (Manzo, 2013).

Within traffic assignment models, free flow speed and capacity are the two link-specific parameters that are present in VDFs. These parameters can be expressed as the point value or distribution of values. These coefficients are estimated from observable data or derived using guidelines. Free flow speed can be measured directly, while capacity is more difficult to calculate because of its dynamic nature. (Neuhold & Fellendorf, 2014) considered that the probability of traffic breakdowns determines the stochastic distribution of capacity. In the literature, a Weibull distribution is considered to be a good representation of stochastic behaviour of road capacity (Brilon et al., 2005). Weibull probability density and distribution functions are as follows:

$$f(x) = \frac{a}{b^a} * x^{a-1} * e^{-\left(\frac{x}{b}\right)^a}$$

$$F(x) = 1 - e^{-\left(\frac{x}{b}\right)^a}$$

Where a and b are the form and scale parameters. Maximum likelihood estimation should be used to determine the parameters. Large sample size is required to estimate the distribution of capacity, which might be a disadvantage of this approach (Brilon et al., 2005).

2.11 Existing Application of Machine-Learning Methods

There are multiple studies that used unsupervised machine learning methods to cluster traffic conditions. For instance, (Azimi & Zhang, 2010) applied three pattern recognition methods (K-means, fuzzy C-means, and CLARA (clustering large applications)) to speed-flow data and compared the resulting classes with the Highway Capacity Manual level of service criteria. The K-mean clustering algorithm turned out to perform better and appears to be consistent with HCM-defined LOS. (Yang & Qiao, 1998) employed a neural-network pattern recognition method for

clustering traffic flow condition with the application in Chinese highway traffic. (Sun & Zhou, 2005) proposed a K-mean algorithm for clustering speed-density data points to two or three clusters, which helps to determine the breakpoints in multi-regime traffic models. (Xia & Chen, 2007a) developed an agglomerative clustering method and used Bayesian Information Criterion (BIC) and dispersion measurement techniques to define an optimum number of clusters for the freeway operating condition. (Xia & Chen, 2007b) applied a data clustering method on the freeway speed-density data and pointed out that density had a more significant effect on clustering results compared to speed. (Kianfar & Edara, 2013) investigated three types of clustering algorithms to partition traffic flow data to congested and free flow regimes. They concluded that K-means and hierarchical clustering outperform general mixture model (GMM) clustering. (Xia et al., 2012) suggested a clustering approach for online identification of traffic state. They used modified agglomerative clustering, which stores statistical features of data instead of storing all the historical traffic data. (Gu et al., 2018) proposed a big data-driven two-stage framework for clustering link fundamental diagrams and road functional classes. They tested multiple connectivity-based and centroid-based approaches for clustering and concluded that the connectivity-based approach performs better. (D. Wang et al., 2020) performed mixture model-based clustering to calibrate fundamental diagram for the application in network level simulation. They pointed out that heuristic clustering techniques like K-means and hierarchical clustering may not work properly when used on data that contain overlaps across clusters.

Chapter 3

3 Study Area and Dataset

3.1 Study Area

The Greater Toronto and Hamilton Area (GTHA), Canada's most populous urban region, is the subject of this research. The GTHA is made up of six municipalities: regional municipalities of Durham, Halton, Peel, and York, as well as the Cities of Toronto and Hamilton (Figure 3.1). In 2021, the GTHA's overall population was 51.19 percent of Ontario's and 19.68 percent of Canada's total population. The GTHA is not only a dense area with an average population density of 883.49 persons/km², but it also has a rapid population growth of about 4.7 percent between 2016 and 2021 (Statistics Canada, 2021). More statistics regarding population are shown in Table 3.1.

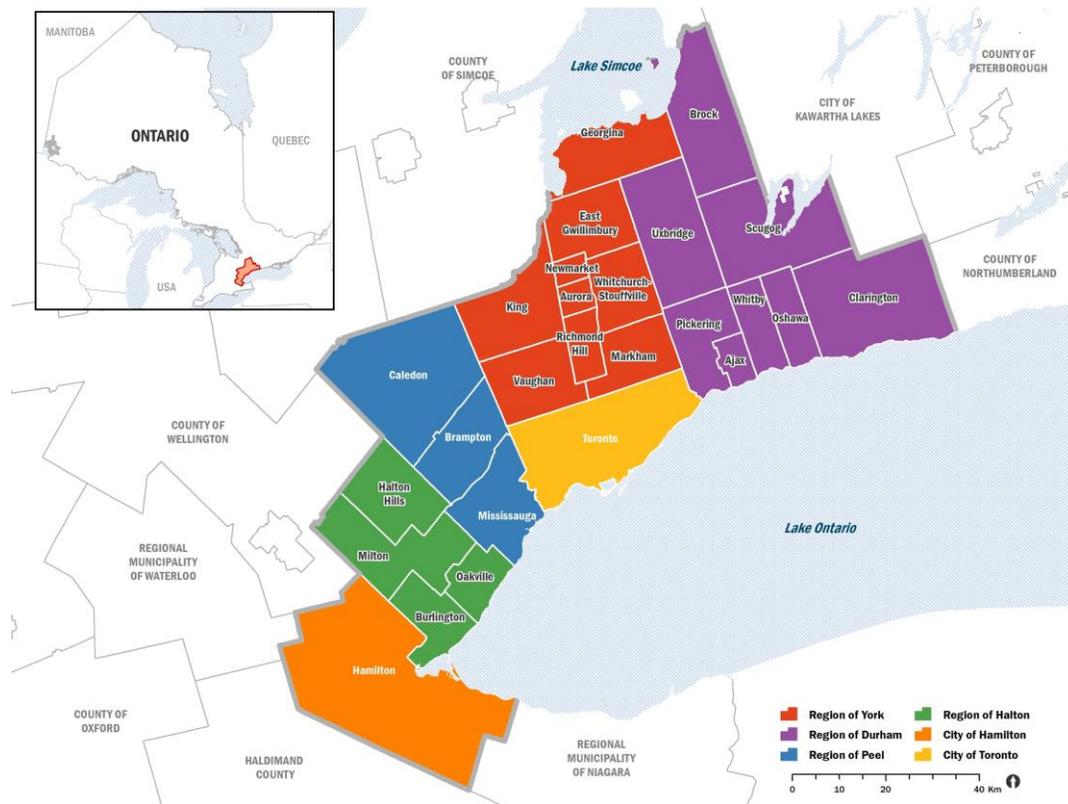


Figure 3.1-GTHA Municipalities, Source: (Government of Canada, 2022)

Table 3.1-Population of the Greater Toronto and Hamilton Area (GTHA)

	Population, 2021	Population, 2016	Population percentage change, 2016 to 2021	Total private dwellings	Private dwellings occupied by usual residents	Population density per square kilometre	Land area in square kilometres
Toronto, City	2,794,356	2,731,571	2.3	1,253,238	1,160,892	4,427.80	631.1
Hamilton, City	569,353	536,917	6	233,564	222,807	509.1	1,118.31
York, Regional municipality	1,173,334	1,109,909	5.7	405,863	391,034	667.3	1,758.27
Peel, Regional municipality	1,451,022	1,381,739	5	467,970	450,746	1,163.20	1,247.45
Halton, Regional municipality	596,637	548,435	8.8	214,322	208,601	617.8	965.71
Durham, Regional municipality	696,992	645,862	7.9	250,559	243,048	276.5	2,521.11
Total	7,281,694	6,954,433	4.706	2,825,516	2,677,128	883.492	8,242

When it comes to the GTHA's transportation networks, there are a number of expressways that link the regions, including Highways 400 and 404 in the north-south direction and Highway 401, the tolled Highway 407, the Queen Elizabeth Way (QEW), and the Gardiner Expressway in the west-east direction. The GTHA network has 27951 kilometres of roadway, of which 2824 kilometres are highways that serve a large proportion of mobility needs of the region. Figure 3.2 illustrates the GTHA road network, which is used in the GTAModel.

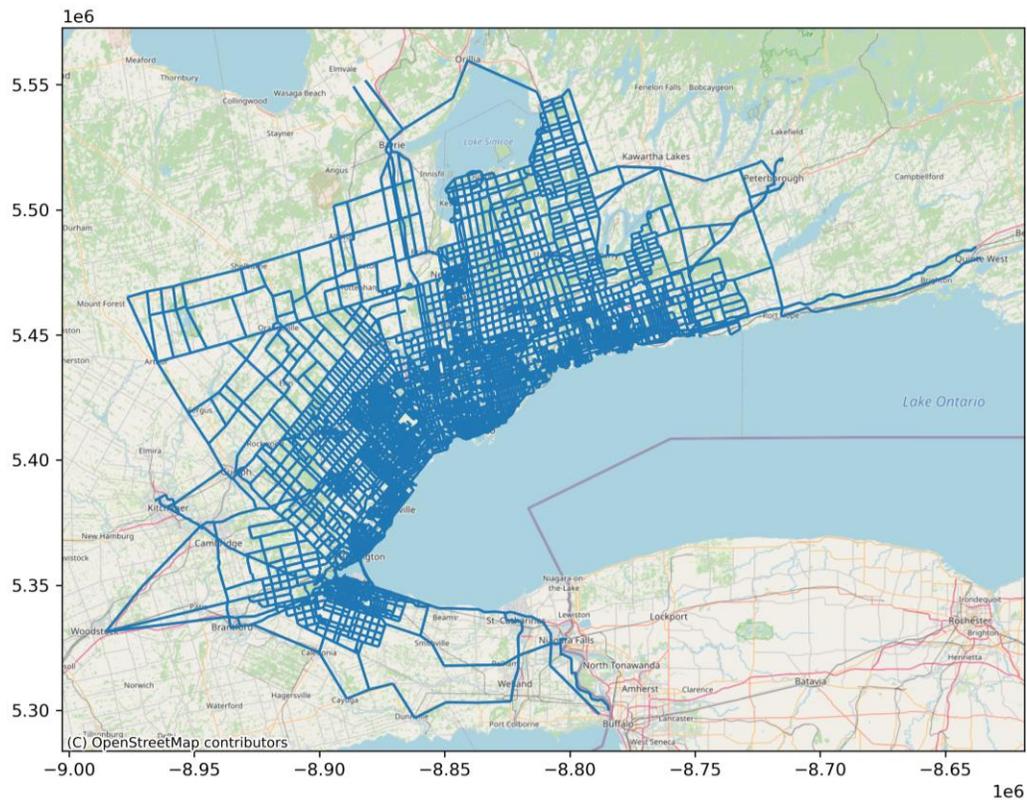


Figure 3.2-GTHA road network, Source: (UofT Travel Modelling Group (TMG), 2017)

3.2 Data Sources

The primary data source used in this study is the speed-flow data exported from the StreetLight Data platform. Besides, GTAModel V4.0 travel demand model output and DMTI Land-use data are also used in the study.

3.2.1 StreetLight Data

As a result of recent breakthroughs in information and communication technology, the use of real-time crowd-sourced data feeds in transportation modelling has increased. These novel data sources have wider spatial coverage compared to conventional traffic data sources. In this research project, UTTRI has the academic license of StreetLight Data, which is an on-demand mobility analytics platform. This location intelligence company process anonymized location-based services (LBS) data collected by smartphones and navigation devices to produce various transportation measures. In this study, Streetlight was used to provide 15-minute vehicular traffic counts and average speed. Having data for one year from 791 road segments in the GTHA can give us an excellent grasp of the temporal variations and trends exhibited in the region's road infrastructure.

In terms of data acquisition, the user may use the Streetlight platform to specify a road segment as an input, and transportation metrics can be generated for that section. To match the streetlight results to the GTAModel network, road segments in the streetlight platform are manually drawn and labelled the same as GTAModel. Users can select the mode of travel from these four options: all vehicles, truck, bicycle, and pedestrian. As Streetlight aggregate the data on a daily basis for the selected period of time, the data is exported for each week from this platform. The data includes information on the day of the week and time of day. However, there is no date index in the data. To get date information, the required data was extracted per week, and the file name was labelled with the week's start and end dates. The dataset also includes trucks' relative trip activity (by weight class) and speed/travel time distribution, but this feature is not available for all segments or times of the day. Depending on the road functional classifications, data quality and availability may vary. The reason for this might be that on some roadways, just a few or no cars traversing a section of the road may be recorded.

GTAModel network contains information about the number of lanes. However, this information is not consistent with reality, especially for those sections that contain additional lanes for the

on/off-ramp. Therefore, this information is updated by checking the satellite view of the road segment on Google Maps. In the case that the road segment has a varying number of lanes, the larger section is considered as the number of lanes.

A sample of the data exported from StreetLight Data's platform is provided in Table 3.2:

Table 3.2-Sample dataset exported from StreetLight Data

Vehicle Type	Zone Name	Zone Is Pass-Through	Zone Direction (degrees)	Zone is Bi-Direction	Day Type	Day Part	Average Daily Segment Traffic (StL Volume)	Avg Segment Speed (kph)
Personal	11856-15128	yes	32	no	1: Monday (M-M)	05: 1:00am-1:15am	281	108
Personal	11856-15128	yes	32	no	1: Monday (M-M)	06: 1:15am-1:30am	421	105
Personal	11856-15128	yes	32	no	1: Monday (M-M)	07: 1:30am-1:45am	140	104
Personal	11856-15128	yes	32	no	1: Monday (M-M)	08: 1:45am-2:00am	702	100
Personal	11856-15128	yes	32	no	1: Monday (M-M)	10: 2:15am-2:30am	281	111
Personal	11856-15128	yes	32	no	1: Monday (M-M)	11: 2:30am-2:45am	140	120
Personal	11856-15128	yes	32	no	1: Monday (M-M)	12: 2:45am-3:00am	140	122
Personal	11856-15128	yes	32	no	1: Monday (M-M)	13: 3:00am-3:15am	140	84
Personal	11856-15128	yes	32	no	1: Monday (M-M)	14: 3:15am-3:30am	140	112
Personal	11856-15128	yes	32	no	1: Monday (M-M)	15: 3:30am-3:45am	140	124
Personal	11856-15128	yes	32	no	1: Monday (M-M)	17: 4:00am-4:15am	140	102
Personal	11856-15128	yes	32	no	1: Monday (M-M)	18: 4:15am-4:30am	421	118
Personal	11856-15128	yes	32	no	1: Monday (M-M)	20: 4:45am-5:00am	140	116

A thorough overview of the data fields in Table 3.2 is provided below:

Vehicle Type: Mode of transportation (personal, truck, pedestrian, bicycle)

Zone Name: A specific link I.D., encoded by Travel Modeling Group (TMG) at the University of Toronto, that can be connected to GTAmoel's road network

Zone is Pass-Through: This means that the road segment drawn in the StreetLight platform can calculate the traffic volume which passes through its start, middle and end gates.

Zone is Bi-Directional: This implies that the total volume of traffic on the road in both directions is measured.

Day Type: This field represents the day of the week. As an attribute, it also includes "All Days," which is an aggregated metric for all seven days of the week.

Day Part: This field displays the time of day in local time. Users can have the data subdivided into 1-hour or 15-minute epochs all through the day.

Average Daily Segment Traffic (StL Volume): This traffic count metric is generated based on the processed GPS data by StreetLight Data's machine learning algorithm, expressed in number of vehicles per time epoch.

Average Segment Speed (kph): Average speed is the arithmetic mean of all speed observations on a specific road segment throughout a specific time epoch, expressed in kilometres per hour.

3.2.2 GTAModel V4.0 Outputs

GTAModel V4.0 is the latest operational version of the integrated activity-based travel demand model that is used to study and anticipate travel patterns in the GTHA. The Travel Modelling Group (TMG) created this model, which is continually being updated (UofT Travel Modelling Group (TMG), 2022).

GTAModel contains all 24-hour travel patterns, among which the AM peak period (6:00-9:00) is the focus of this study. From various outputs of this travel demand model, the roadway saturation rate (volume over capacity ratio) is employed in this study. This measure is calculated using GTAModel-based EMME road and transit assignment software.

3.2.3 Road Network and Street Attributes

The GTHA 2016 EMME network was developed by TMG and its partners. The information regarding the attributes of elements in this network is written in the Network Coding Standard (NCS), which is updated every five years when the region's Transportation Tomorrow Survey (TTS) is undertaken. The EMME network contains nodes, links and transit lines, from which the links are the focus of the study. Some attributes are assigned to each link, including link, length, number of lanes, functional class, speed, lane capacity, and lane type. It should be noted that for each functional class a specific volume delay function is defined. The functional classes in NCS16 are derived from the combination of multiple sources: NCS11, geometric design guide and GGHMv4 VDF definitions. Table 3.3 shows the link functional classes and VDF definitions in the NCS16. The GTAModel Emme network has 50691 road segments that are clustered into 18 classes.

Table 3.3-Link functional class and VDF definitions, Source: (UofT Travel Modelling Group (TMG), 2017)

Area	Class	Subclass	Land use	Other Factors	Speed Range	Lane Capacity	VDF
N/A	Freeways	Freeway				1800	11
		Expressway				1800	12
		Freeway Ramp				1400	13
		Toll highway				1800	14
		Toll highway ramps				1400	15
		Freeway/expressway HOV				1800	16
		Freeway/expressway HOV ramp				1400	17
		Freeway/expressway truck only					
Rural	Arterials	Long Distance Arterials		Unsignalized long distance arterials	70-80	1400	20
		Major Country Roads		Major roads with a greater number of signals	60-90	1000	21
	Collector	Collector Road	Main Street or Collector Roads		40-60	500	22
Suburban	Arterials	Principal urban arterials	Low density residential/commercial development with no direct accesses	Long signal spacing and good signal coordination/progression	60-90	1000	30
	Collector	Suburban Collector Roads			40-60	500	31
Urban	Arterials	Major urban arterials	Low/medium density residential or commercial with some accesses	Longer signal spacing, good level of signal coordination and green-time allocation	50-80	800	40
		Major urban arterial HOV	Low/medium density residential or commercial with some accesses	Longer signal spacing, good level of signal coordination and green-time allocation	50-80	800	41
		Minor urban arterials	Low/medium density residential or commercial with direct accesses	Closer signal spacing, occasional illegal parking causing interference	50-80	700	42
	Collector	Downtown/city centre roads	Roads in high density office/commercial (CBD) with high pedestrian activity, parking, etc.	Presence of street cars and cyclists	40-60	600	50
		Collector Roads	Roads providing access to local streets	All-way stops, traffic calming measures	40-60	500	51
N/A	Local	Centroid Connectors	Local Streets		40	9999	90

3.2.4 DMTI Land-use Data

This analysis makes use of the Desktop Mapping Technologies Inc. (DMTI) land use dataset, which contains historical data on land use patterns at the dissemination area (DA) scale. The land-use percentage within a DA is divided into the following categories: residential, commercial, institutional, industrial, parks, open space, and water. This dataset contains 2001, 2006, 2011, and 2016. The latest year is used for the analysis.

The DMTI lands use data is spatially joined on the GTAm model road network. When a road segment crosses several land-use areas, the land use allocated to that road segment is averaged over the intersecting region.

Chapter 4

4 Methodology

4.1 Calibration Framework Structure and Specifications

To replicate the true potential of transportation infrastructure supply, the analysis should be done for a typical day that has normal traffic operating conditions. In order to account for long-term planned and unexpected short-term incidents, a multi-step machine learning-based framework is proposed to clean, classify and calibrate traffic flow fundamental diagrams of various roadway functional classes in the GTHA. This framework is formed by three clustering algorithms that detect typical days, normal traffic conditions, and road functional classes in 3 consecutive stages. In the last stage of the framework, the cleaned and clustered data is utilized for VDF calibration. Figure 4.1 present the framework mentioned above.

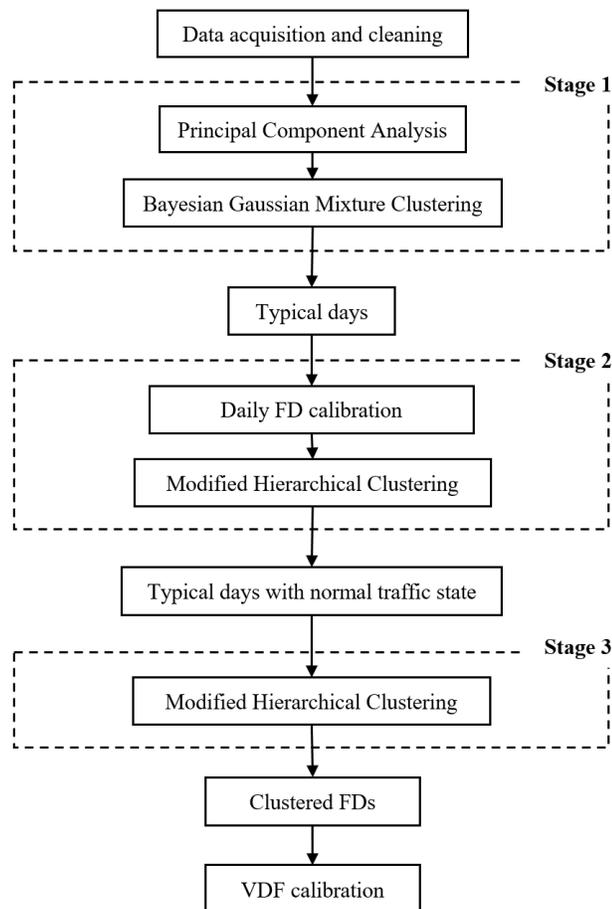


Figure 4.1-Three-stage clustering and calibration framework

Data acquisition and data cleaning are the preliminary steps of this framework. One morning peak hour run of the existing GTAModel-based Emme assignment is used to identify the road segments for data acquisition from StreetLight Data. Although the GTAModel's existing volume delay function is not well calibrated, the traffic assignment that uses the existing VDFs can provide an initial suggestion for selecting a congested road segment for further VDF calibration. Congested road segments are chosen based on a volume-to-capacity ratio. 791 road segments are selected from the 3642 road segments that have V/C greater than 0.9 in the network, which is shown in Figure 4.2. The reason for selecting the existing congested roadway is that we need to find the road segments that contain a wide range of data points of congested and uncongested traffic conditions across each day, which is necessary for avoiding sample selection bias in calibrating the VDF parameters.

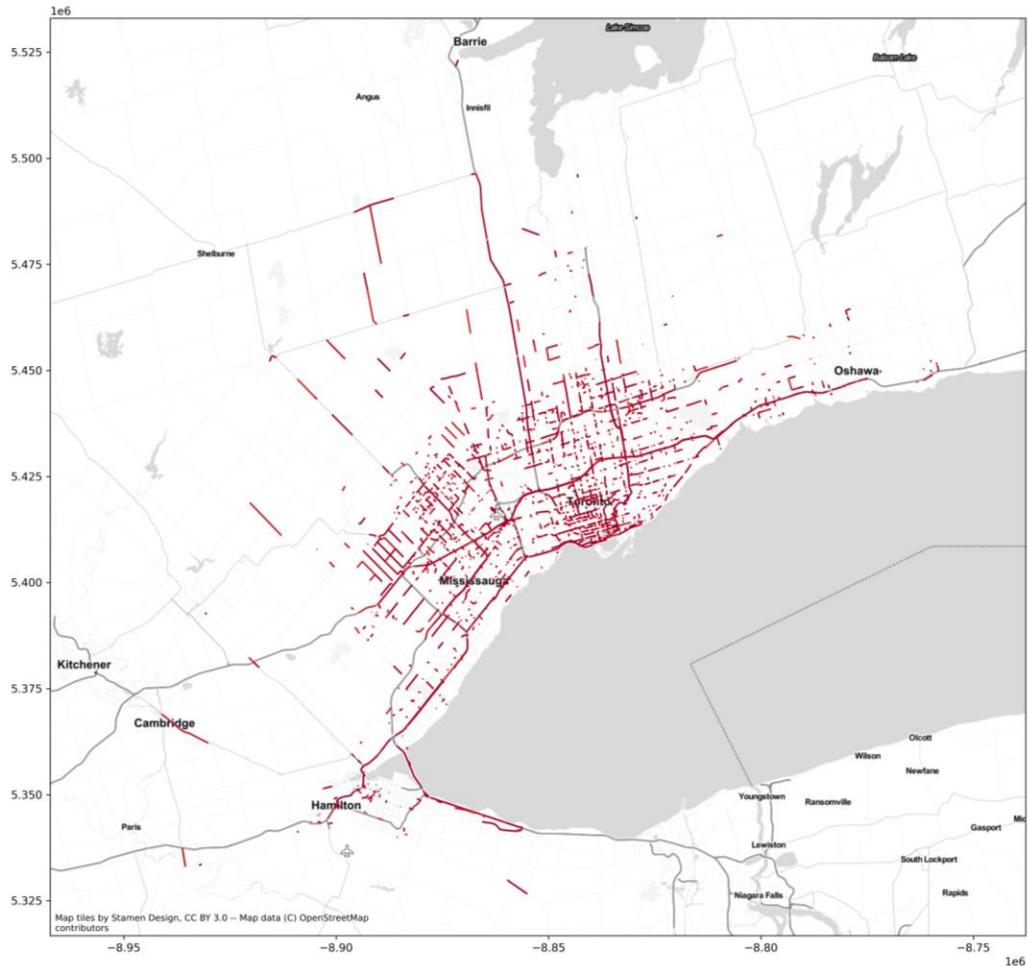


Figure 4.2-GTAModel Emme-based Road segments with V/C >0.9

Before getting to the next step in the framework, some features should be engineered for the purpose of this study. Because GTAModel is being run for the morning peak hour, the data used to build the VDF function should be consistent with this time frame. The data extracted from the StreetLight Data database is based on 15-minute intervals. On the data, a rolling window of one hour is conducted. The number of observations in this window must be four. For each window, the speed is averaged and the flows are summed. After applying the rolling window, a maximum of 93 pairs of 1-hour-aggregated speed-flow pairs for each day are generated. After applying a one-hour aggregation, the data is smoothed, accumulated, and less sparse, which facilitates the curve fitting. The 1-hour-aggregated density (veh/km/lane) is calculated based on the speed, flow and number of lanes. The lane-based density allows us to compare different road segments with different numbers of lanes.

After retrieving the typical days and appropriate features, we can continue to the second clustering stage. This clustering step aims to remove days with unusual traffic patterns. Weather conditions or other short-term unexpected incidents, e.g. accidents, might produce this irregularity and between-day variability. The dissimilarity between the daily calibrated fundamental curves can be used to detect this issue and cluster the data to normal and abnormal traffic states. Hence, before applying the clustering algorithm to the data, calibration of traffic flow fundamental equations should be done on a daily basis. For the daily calibration, we have chosen the modified two-regime Greenshield model. This model consists of two parts: constant speed for the free-flow traffic condition and the modified Greenshield model for the congested traffic condition. As this model has a continuous form, it does not consider capacity drop, which helps to have less calibrated parameters to deal with in the next steps. The calibrated curve is drawn in Figure 4.5, and has the following mathematical formulation:

$$\begin{cases} v = u_f = v_{int} \times \left(1 - \frac{k_b}{k_j}\right)^\alpha & (0 \leq k \leq k_b) \\ v = v_{int} \times \left(1 - \frac{k_b}{k}\right)^\alpha & (k_b \leq k \leq k_j) \end{cases}$$

Where v = speed, u_f = free – flow speed, k_b = breaking – point density, k = density, v_{int} = speed intercept, k_j = jam density, and α = power term.

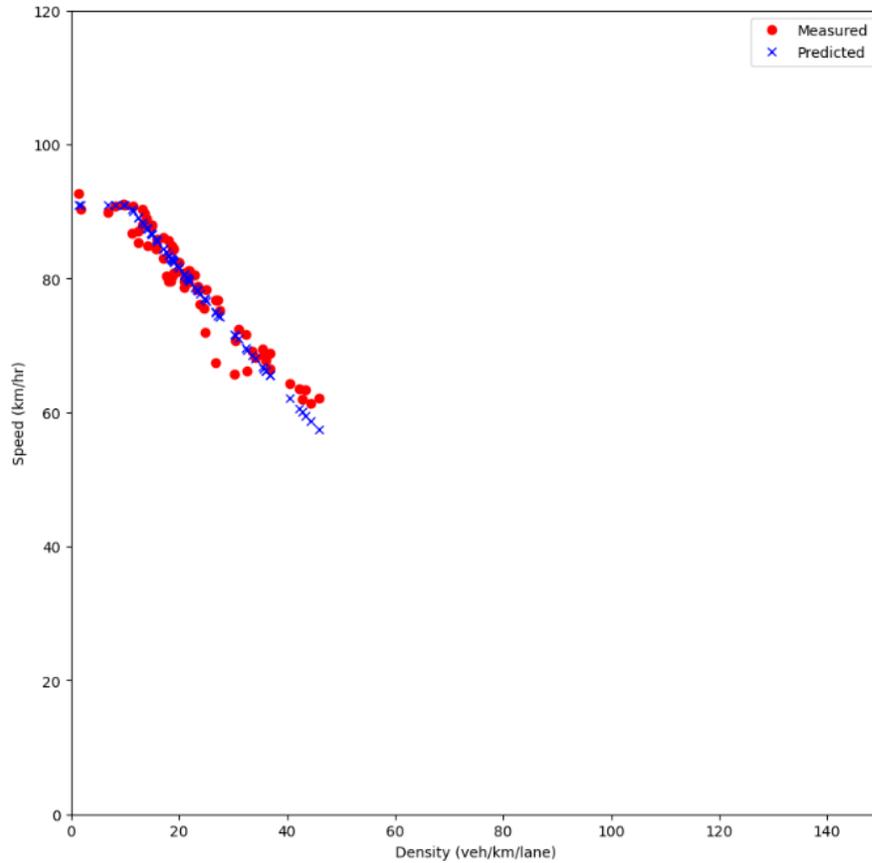


Figure 4.5-modified two-regime Greenshield traffic flow model calibrated with empirical data

Among the three traffic fundamental equations, we have chosen the speed-density equation because it has a one-to-one mapping function which is suitable for the calibration. The speed-density equation can be easily transformed into speed-flow and flow-density equations by using the fundamental traffic flow equation. The modified Greenshield model has four parameters that can be calibrated with empirical data. The jam density is considered to be constant and equal to 144 veh/km/lane, which is based on the assumption that the distance between vehicles in a traffic jam is 7 meters. Therefore, free flow speed, breaking point density and power term are the three variables for calibration. For each day, the parameters of the speed-density equation are calibrated. The nonlinear least square method is used for the curve fitting procedure. The mathematical formulation of the least squares method (LSM) is expressed as follows:

$$\min Z = \sum_{i=1}^P (v_i - \hat{v}(k_i, m))^2$$

Where v_i and $\hat{v}(k_i, m)$ are observed and estimated speeds corresponding to density k_i . The optimization has been done with 'trust-constr' method in Scipy library in python, which minimizes a scalar function subjected to constraints. Optimization settings are defined in Table 4.1 and Table 4.2:

Table 4.1-Parameter optimization setting for highways

Parameter	Initial value	Optimization bound
v_{int}	100	≥ 0
k_b	0	[0,25]
α	2	≥ 1.5

Table 4.2-Parameter optimization setting for all roadways excluding highways

Parameter	Initial value	Optimization bound
v_{int}	100	≥ 0
k_b	0	= 0
α	2	≥ 1.5

Since we calibrate on a daily basis and may not have enough data in the congested regime, the approach does not account for sample selection bias. At the end of the calibration process, we have a matrix with $N \times 3$ arrays, in which N is the number of rows (date of typical days) and columns are the three calibrated parameters. The clustering technique should be applied once the daily fundamental diagram calibration for the typical days has been completed. On the calibrated parameters, we have applied various connectivity-/centroid-/density-based clustering, and finally, hierarchical clustering with a modified affinity matrix is employed. The area between two calibrated curves is used as a measure of similarity between the traffic dynamics of different days. The area between curves is determined using 10 points from each daily calibrated curve to decrease computing costs. Other similarity measures can be employed to build the affinity matrix that is utilized in hierarchical clustering. For example, (Gu et al., 2018) employed Frechet distance. In the modified hierarchical clustering, the number of clusters is assumed to be 10. Any clusters that have a sample size lower than 30 are eliminated, which are assumed as a cluster of abnormal traffic state. In this clustering stage, another between-day variability is captured, which is unexpected

short-term incidents. Towards the end of this procedure, the cluster of typical days with normal traffic condition is extracted, as shown in Figure 4.6. For this cluster, the mean values of calibrated parameters are calculated for being used in the last stage of clustering.

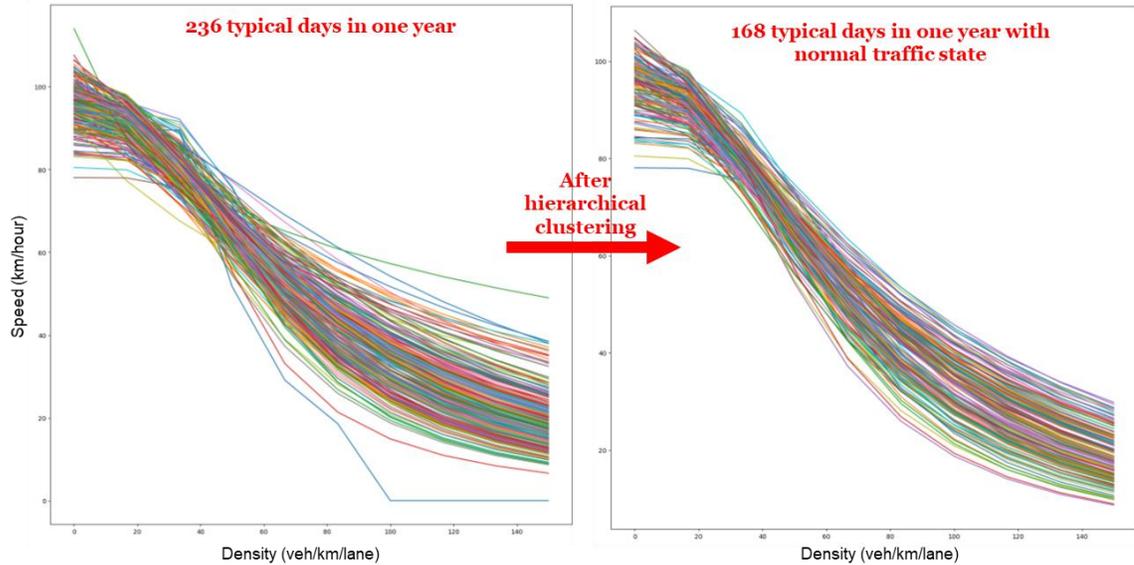


Figure 4.6-Daily calibrated speed-density diagram for one road segment before and after applying modified hierarchical clustering

For each road segment, the first and second stages of clustering are conducted, and the mean calibrated curve is computed. By having one representative of each road segment's traffic performance, the next step is to find the clusters of roadway functional classes. All road segments with their mean calibrated parameters are imported in the modified hierarchical clustering. The area between the calibrated curves of two road segments is the similarity measure in this clustering. The number of clusters is manually tuned depending on user assessment. Figure 4.7 depicts a speed-density diagram with 7 clusters after applying modified hierarchical clustering.

After finding the road functional classes, we should go through the last stage of the framework, which is VDF calibration. Data points for all road segments within one cluster are used for VDF calibration. Various VDF formulations are tested, and the one with the highest goodness of fit is chosen.

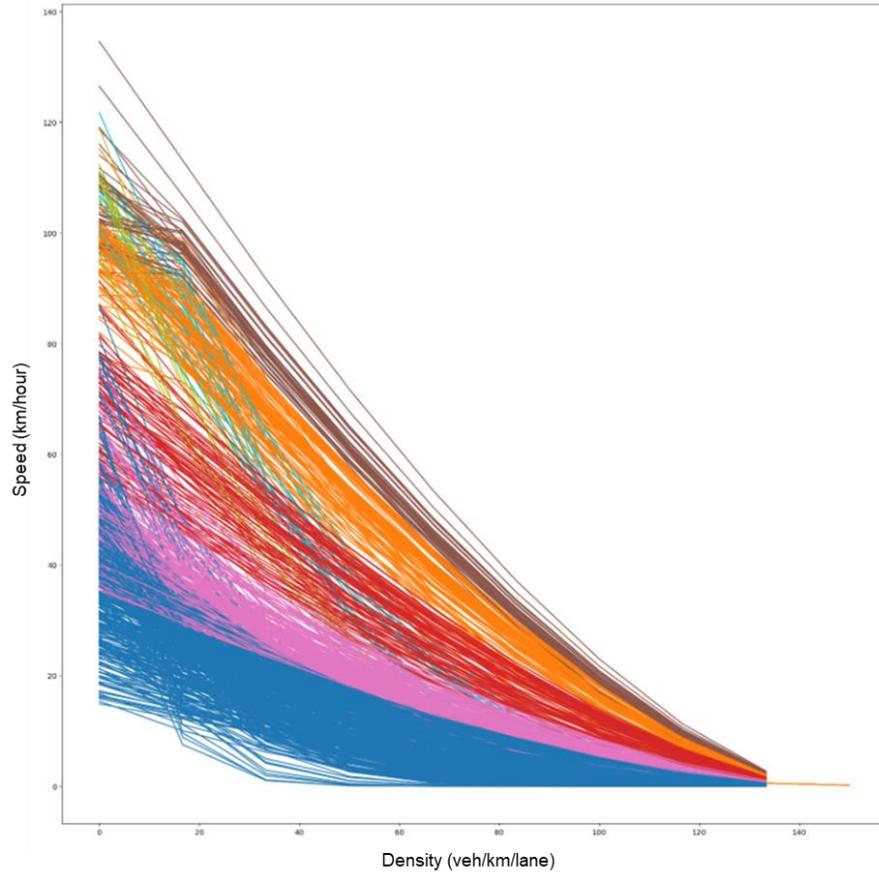


Figure 4.7-Speed-density diagram for all road segments after applying modified hierarchical clustering with $n_cluster = 7$

Chapter 5

5 Results and Discussion

5.1 Result

In this chapter, we put our proposed framework into practice using observed data to assess the performance of the region's various road functional classes. The calibrated parameters of the traffic flow fundamental diagram are shown. Using these calibrated diagrams as a basis, the free flow speed and capacity are computed and compared against the NCS16's reported measures. Some adjustments are suggested so that the new version of NCS reflects the newly calibrated performance measures. To date, NCS22 is under final edit, and the result of this thesis can be reflected in NCS22.

5.1.1 Calibrated Parameters for all Sampled Road Functional Classes

This section presents the existing and newly-calibrated free flow speed and capacity for all sampled road functional classes. As shown in Figure 5.1, the existing free-flow speed in NCS16 for roadways in the City of Toronto's downtown core is mostly 40 to 60 km/hr, compared to GTHA borders which are mostly 70 to 90 km/hr. Based on NCS16, highways have a free flow speed of 100 to 110 km/hr. The calibrated free-flow speed based on observed data has the same pattern as the NCS16 metrics.

The reported capacity in NCS16 ranges from 400 to 2000 veh/hr/lane, although the calibrated capacity based on empirical data reveals another story. The calculated capacity ranges between 400 to 3500 veh/hr/lane, where this upper bound exceeds maximum reported capacities in previous works in literature (see Figure 5.2). The possible reason for this inconsistency is discussed in Section 6.2.

The difference between existing and calculated values of free flow speed and capacity is illustrated in Figure 5.3. This diagram shows that most free flow speed values in NCS16 are overestimated compared to reality. Furthermore, if we exclude highways, capacity values are well-estimated compared to NCS16. Regarding the highways, the capacity is underestimated in NCS16.

Figure 5.4 can help us compare the values of observed data with existing values in NCS16 in more detail. Roughly speaking, the data points are mostly near the 45-degree line in both capacity and

free flow speed curve. Most free flow data points are below this inclined line, which means the NCS16 is overestimating this metric. In regards to the capacity, if we exclude the high capacity values, most of the capacity values are overestimated in NCS16. Real-world data shows that for high capacity values, which are mostly highways, NCS16 underestimates the capacity.

Boxplots of Figure 5.5 demonstrates how free flow speed and capacity values vary over road segments with a similar number of lanes. Road segments with 1 or 2 lanes have lower values of free flow speed and capacity values compared to the road segments that have 3 or more lanes. Links with 3 lanes have the widest range of free flow speed and capacity compared to other links.

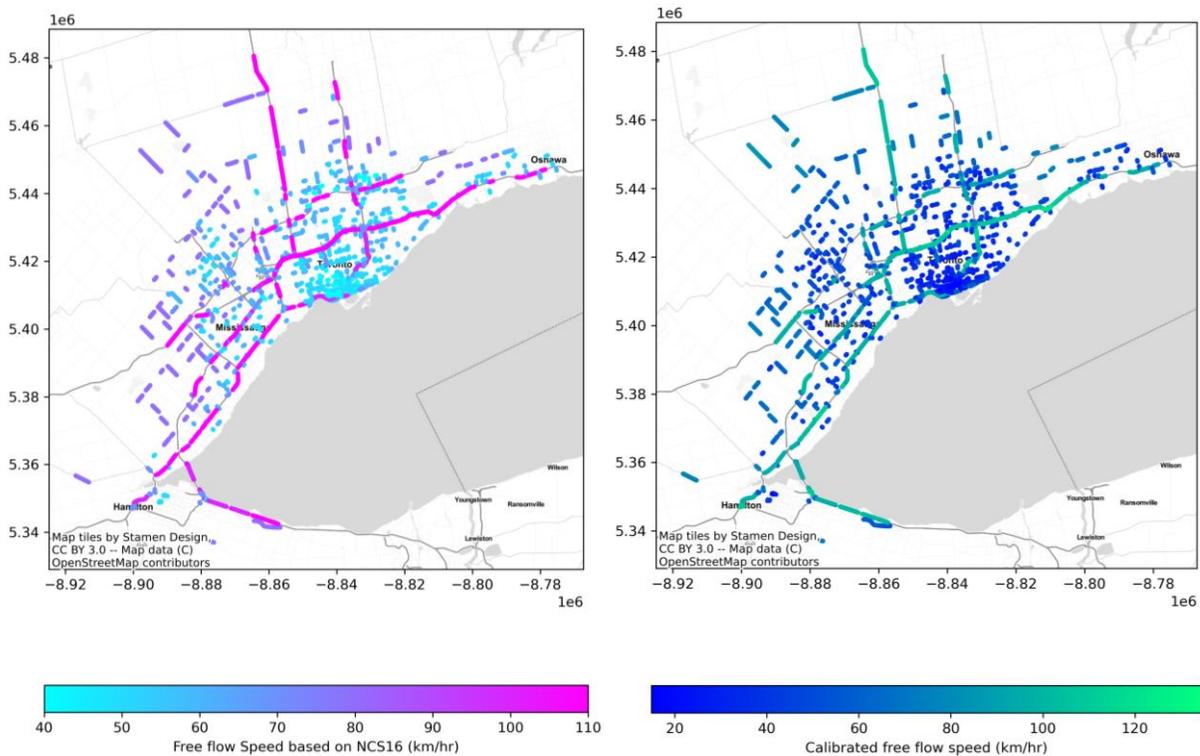


Figure 5.1- Free flow speed based on NCS16 (left) and calibrated data (right), for all VDF codes

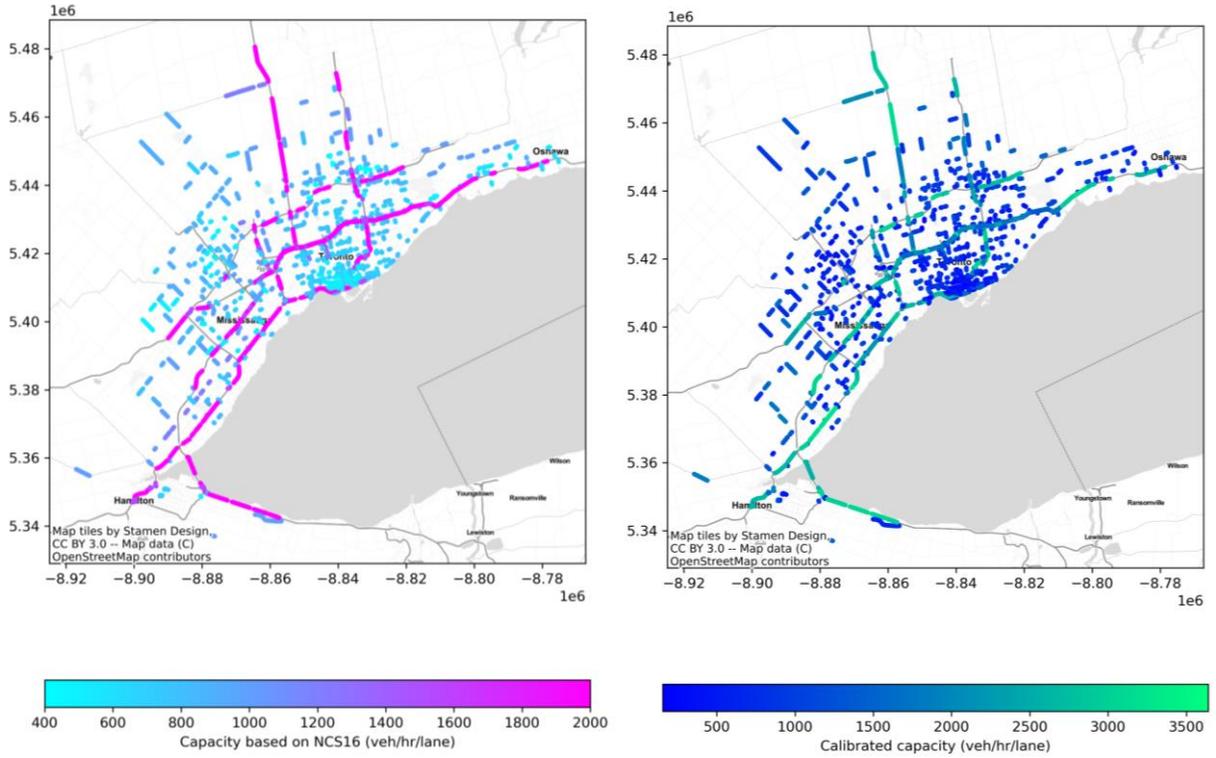


Figure 5.2-Capacity based on NCS16 (left) and calibrated data (right), for all VDF codes

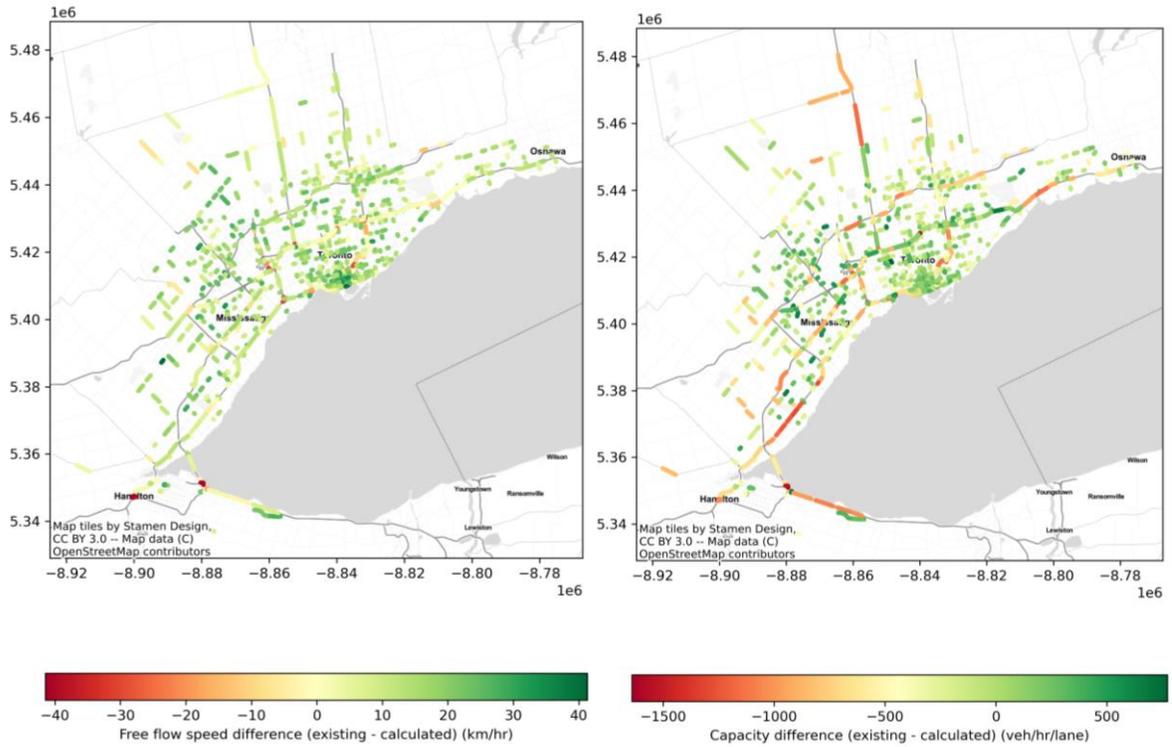


Figure 5.3- Existing and calculated values difference of free flow speed and capacity, for all VDF codes

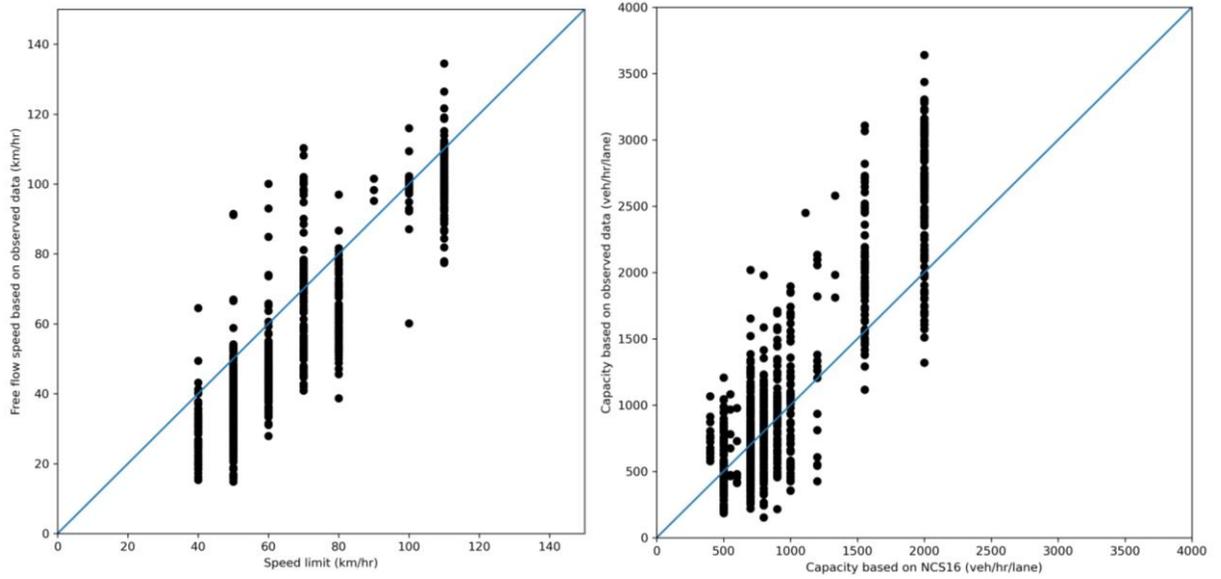


Figure 5.4-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , for all VDF codes

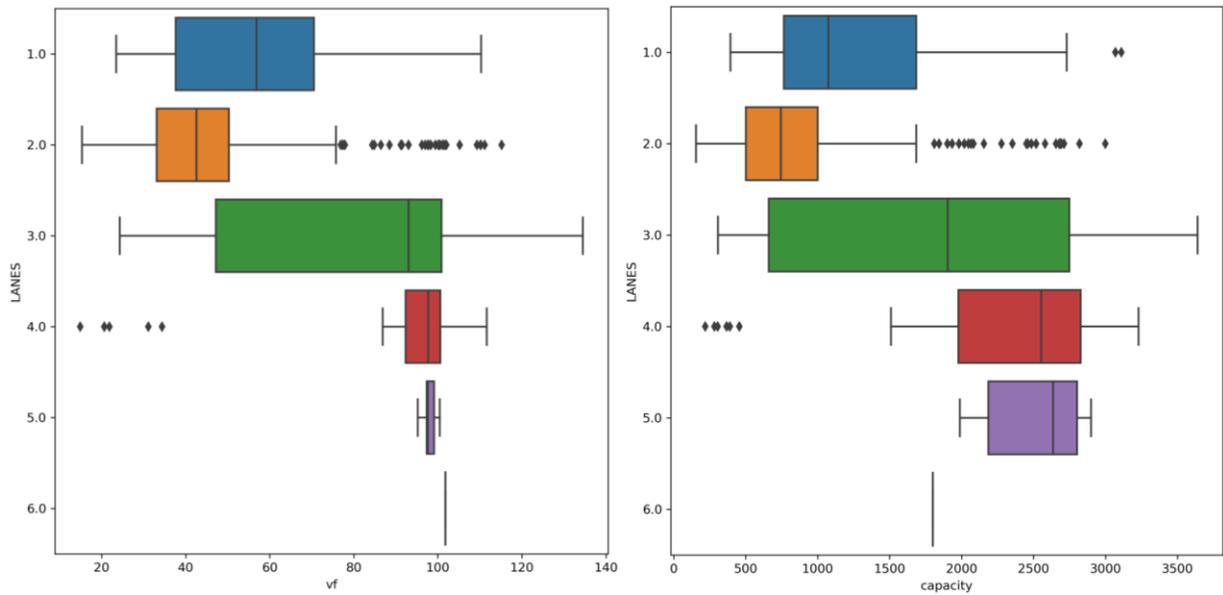


Figure 5.5-Box plot of calibrated free flow speed (left) and capacity (right) classified by the number of lane, for all VDF codes

Two of the VDF codes are covered in greater detail in the remaining parts of this section. In Appendix 1, there are diagrams for other VDF codes.

Highway Ramps (VDF code: 13)

The majority of the free-flow speeds posted in NCS15 for the highway ramps are 70 km/hr, with a range of 50 to 110 km/hr (Figure 5.6). Highway ramps in NCS16 can currently accommodate 800 to 2000 vehicles per hour per lane, with 1555 being the average (Figure 5.7). The number of lanes ranges from 1 to 3 (Figure 5.8). The calibrated free flow speed ranges from 50 to 110 km/hr. The highway ramps south of the financial district have lower free-flow speeds compared to other locations. As opposed to highway-to-highway ramps, which connect two fast channels of traffic, these downtown core highway ramps connect highways to arterials, which run at slower speeds. As shown in Figure 5.9, the NCS16 underestimates the capacity and free-flow speed for highway-to-highway ramps, which could be handled similarly to highway segments.

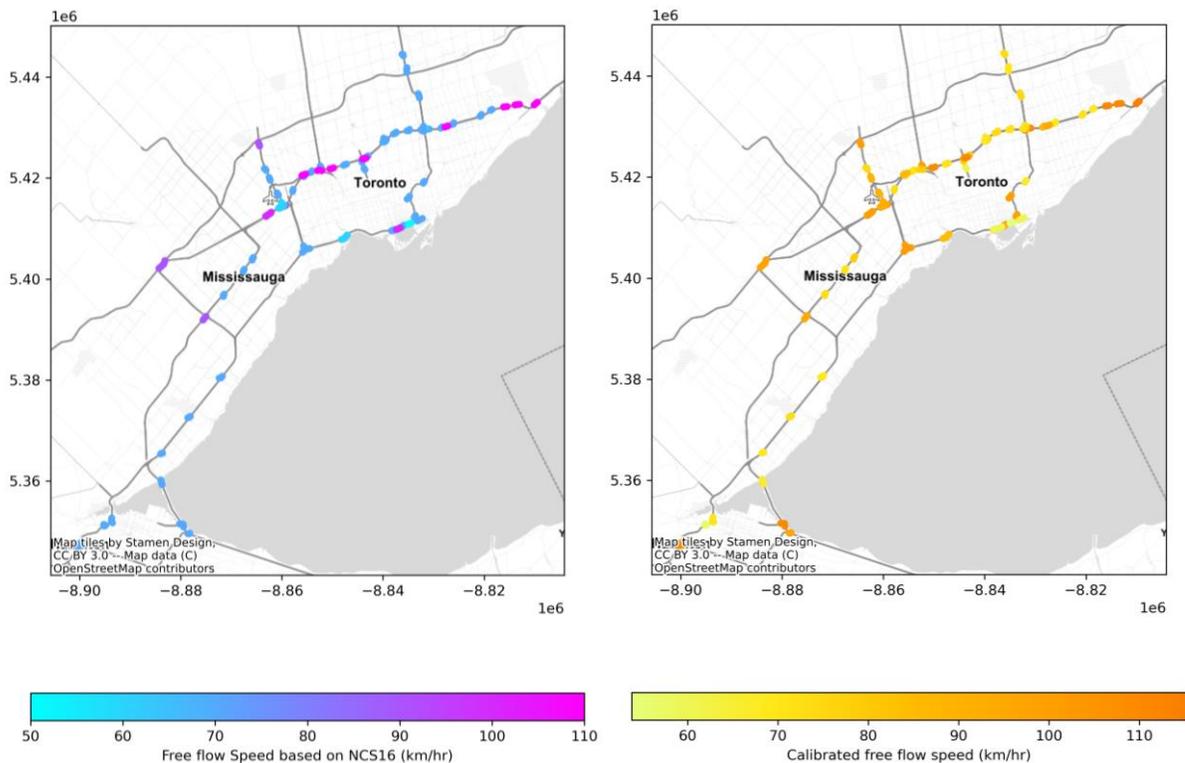


Figure 5.6-Free flow speed based on NCS16 (left) and calibrated data (right), VDF code: 13

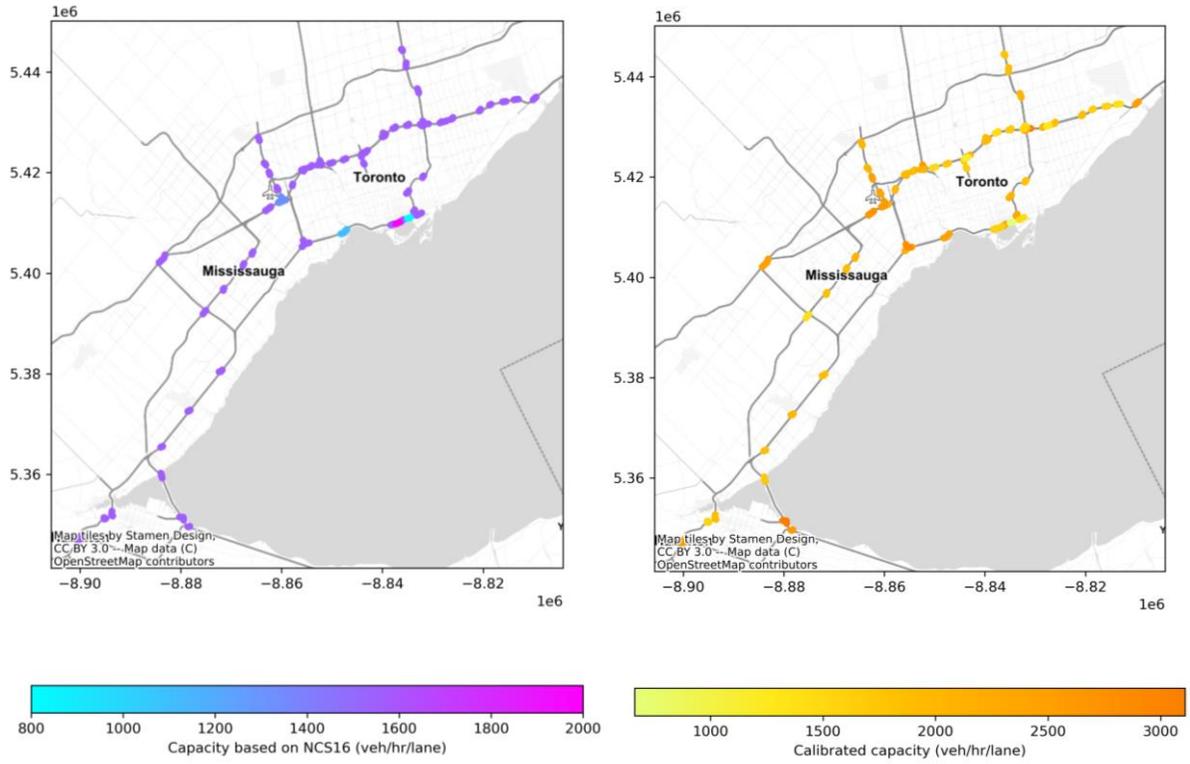


Figure 5.7-Capacity based on NCS16 (left) and calibrated data (right), VDF code: 13

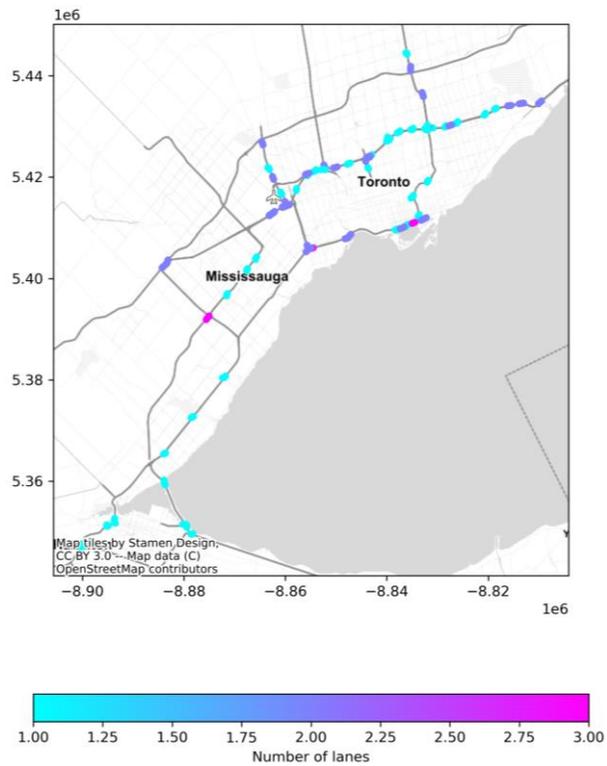


Figure 5.8-Number of lanes, VDF code: 13

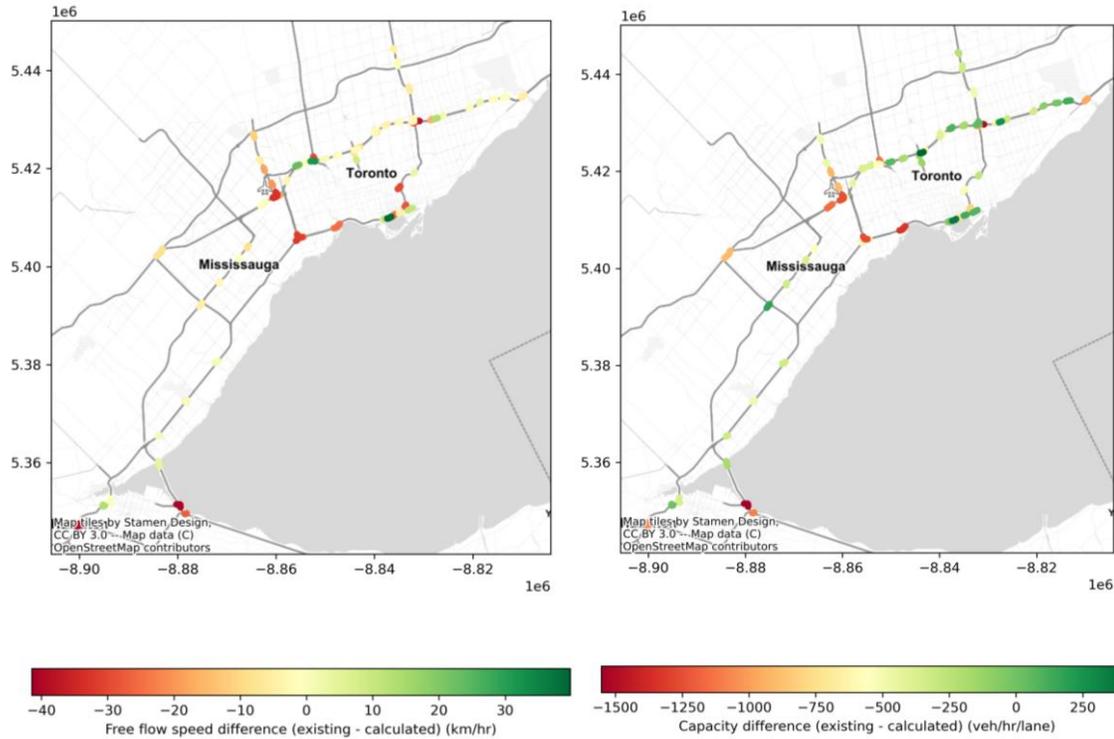


Figure 5.9- Existing and calculated values difference of free flow speed and capacity, VDF code: 13

Downtown City Center Arterials (VDF code: 50)

As can be seen from Figure 5.10 and Figure 5.11, the existing free-flow speed and capacity in NCS16 for the downtown city center arterial roads are in the range of 40 to 50 km/hr and 500 to 700 veh/km/lane, respectively. This region has road segments with a varying number of lanes, most of which have 2 lanes per direction (Figure 5.12). The calibrated free flow speed ranges between 20 to 55 km/hr in this region. Road segments with slower travel speeds are found in the downtown area close to the financial district. In this region, the intersections are spaced closely together, making it difficult for drivers to accelerate or decelerate rapidly, potentially lowering the free-flow speed. In addition, the slower free-flow speed in this area may have been caused by a pedestrian crossing the roadway or by streetcars, which may have blocked a part of the street.

The stated free-flow speed in NCS16 is inconsistent with reality (see Figure 5.13), particularly for those reported on Dundas Street, where there is a significant discrepancy. For this specific street, observed free-flow speeds are in the range of 20 to 30 km/hr, but they are stated to have 50 km/hr free-flow speed in NCS16.

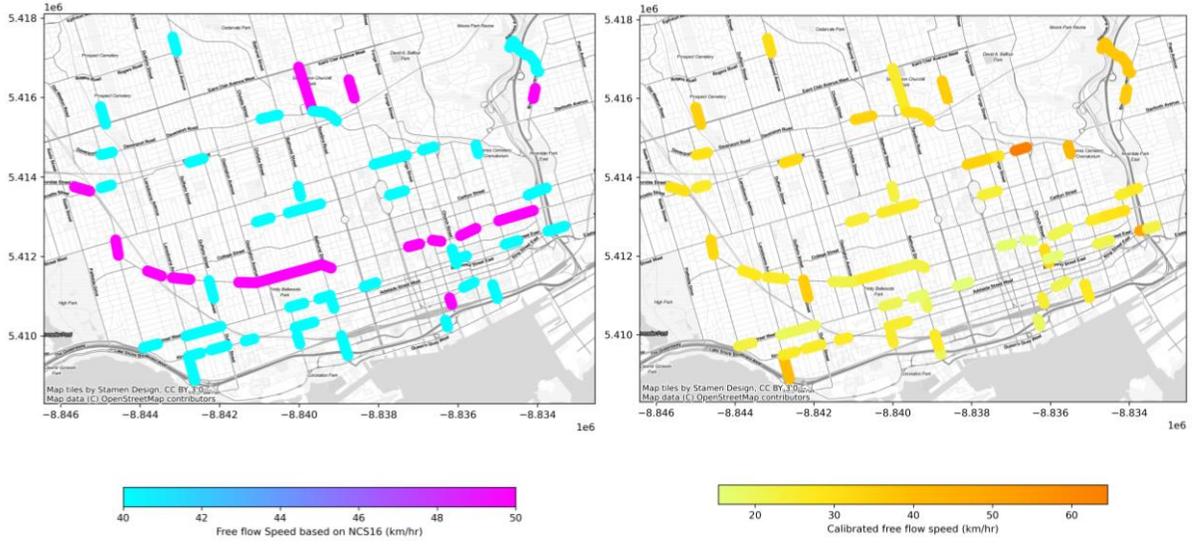


Figure 5.10-Free flow speed based on NCS16 (left) and calibrated data (right), VDF code: 50

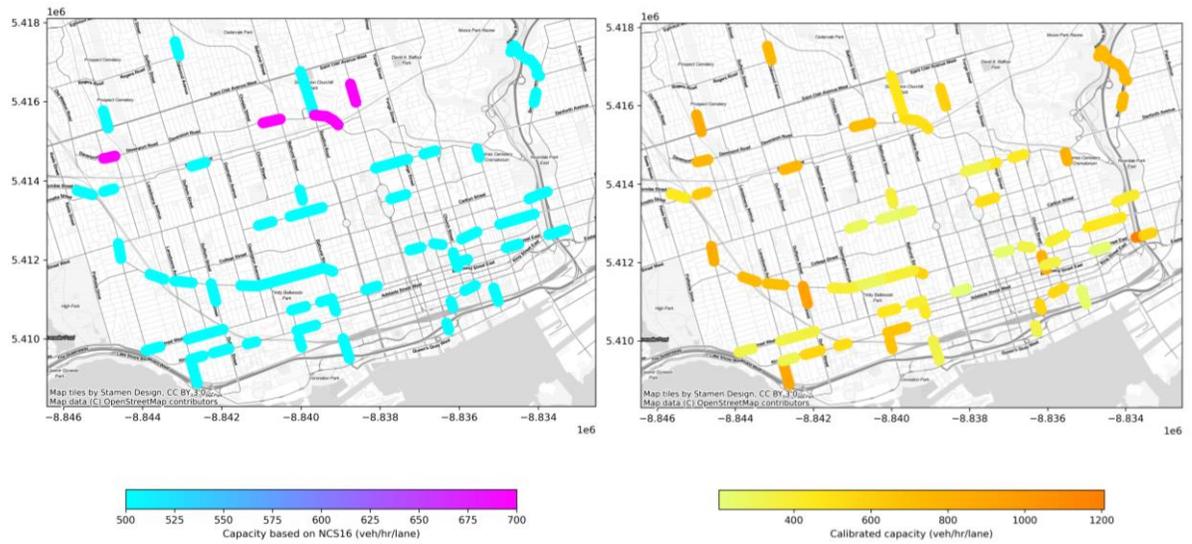


Figure 5.11- Capacity based on NCS16 (left) and calibrated data (right), VDF code: 50

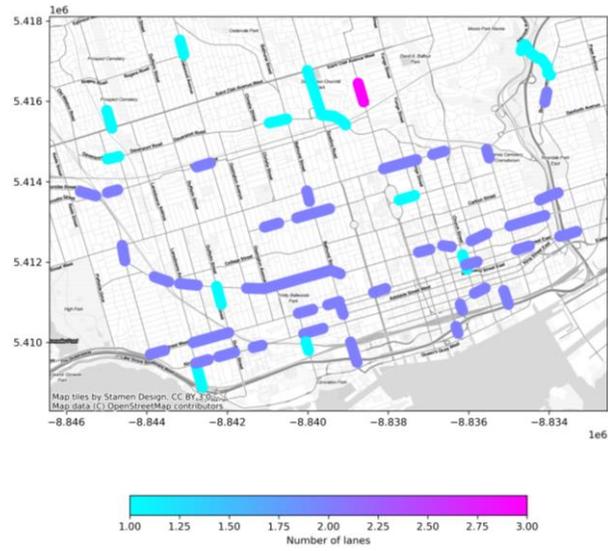


Figure 5.12-Number of lanes, VDF code: 50

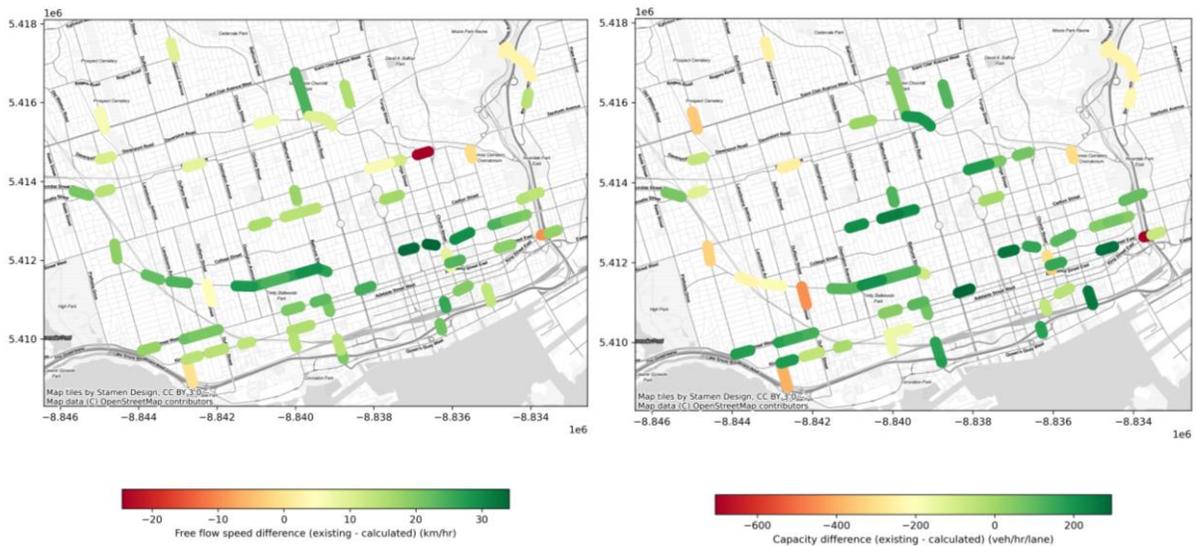


Figure 5.13- Existing and calculated values difference of free flow speed and capacity, VDF code: 50

5.1.2 Road Physical Attributes Distribution among Each Cluster

The traffic flow calibrated parameters are imported in the third stage of clustering, and new road functional classes are exported. A various number of clusters are tested, and 4 is chosen among them. Figure 5.14 shows the clustered road segments after applying modified hierarchical clustering, and Figure 5.15 demonstrates the spatial distribution of those segments. Clusters are named 'Urban', 'Suburban', 'Rural', and 'Freeway'. As illustrated in Figure 5.15, 'Urban' cluster is primarily associated with the City of Toronto's downtown core. 'Suburban' and 'Rural' clusters are mostly on the GTHA's periphery. Highway segments are mostly found in 'Freeway' cluster. Figure 5.16 shows the distribution of the number of lanes within each cluster. The majority of 'Freeway' cluster is assigned to roadways with three or more lanes. In contrast, the major part of 'Rural' cluster is allocated to roadways with 2 or 1 lane/s. 'Urban' and 'Suburban' clusters are nearly identical in terms of the proportion of lanes, with both having a great number of road segments with two lanes. 'Freeway', 'Rural', 'Suburban', and 'Urban' clusters have the highest to the lowest proportion of higher-posted-speed road segments, respectively (Figure 5.17). The proportion of land use adjacent to the roadway is demonstrated in Figure 5.18. Land use near the roadway in 'Urban' and 'Rural' clusters is mostly residential and open area, respectively.

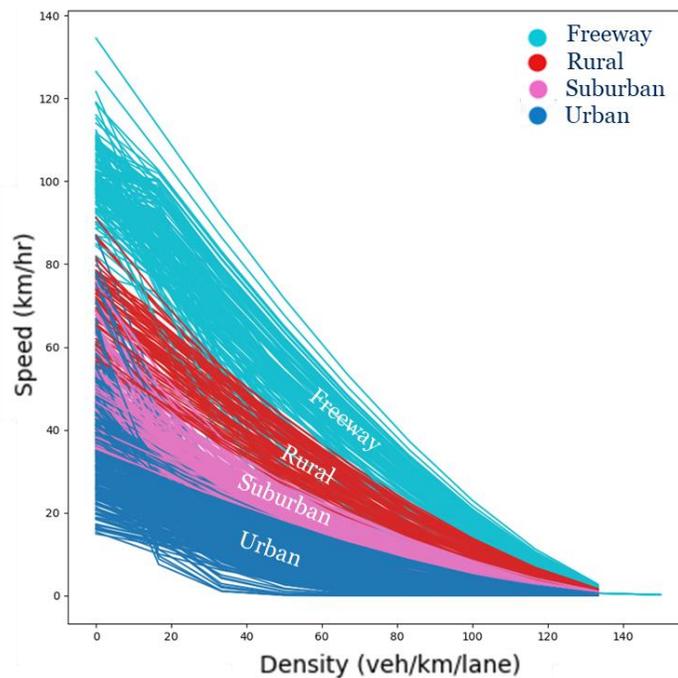


Figure 5.14-Traffic flow fundamental diagrams of all sampled road segments after applying modified hierarchical clustering with $n_cluster = 4$

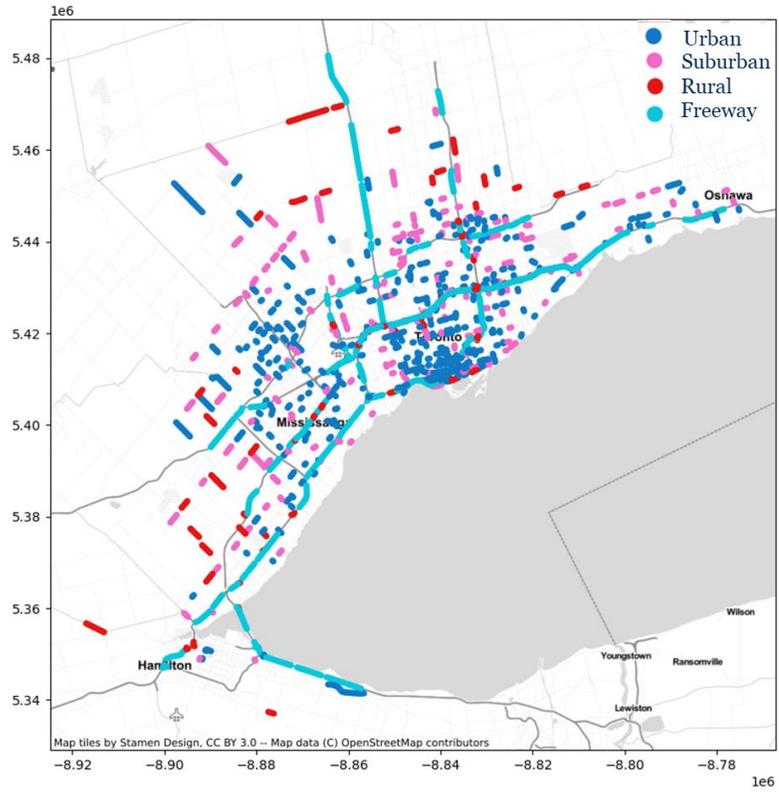


Figure 5.15-Clustered road segments, $n_cluster = 4$

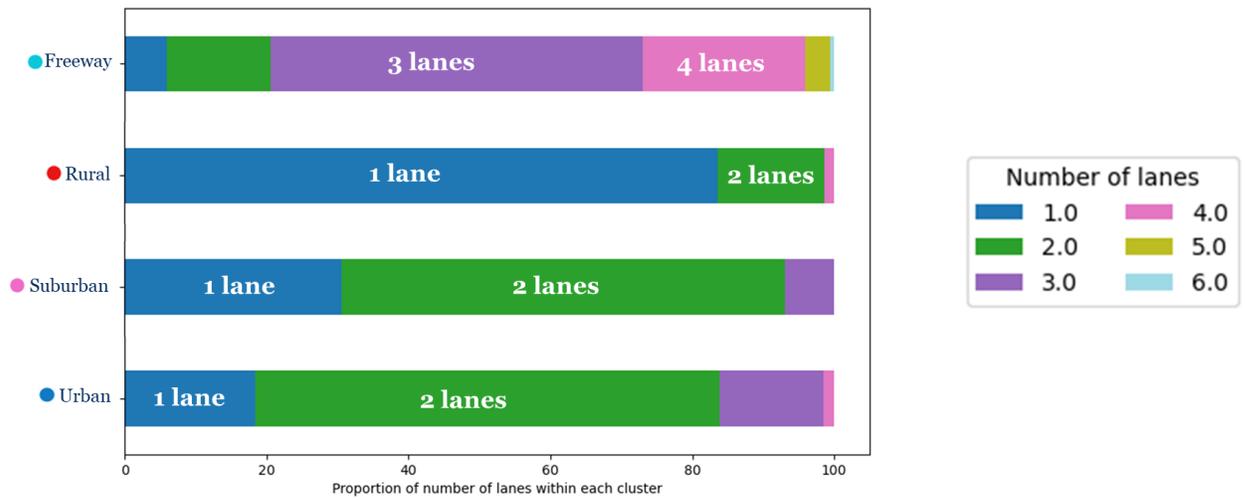


Figure 5.16-Distribution of 'number of lanes' within each cluster

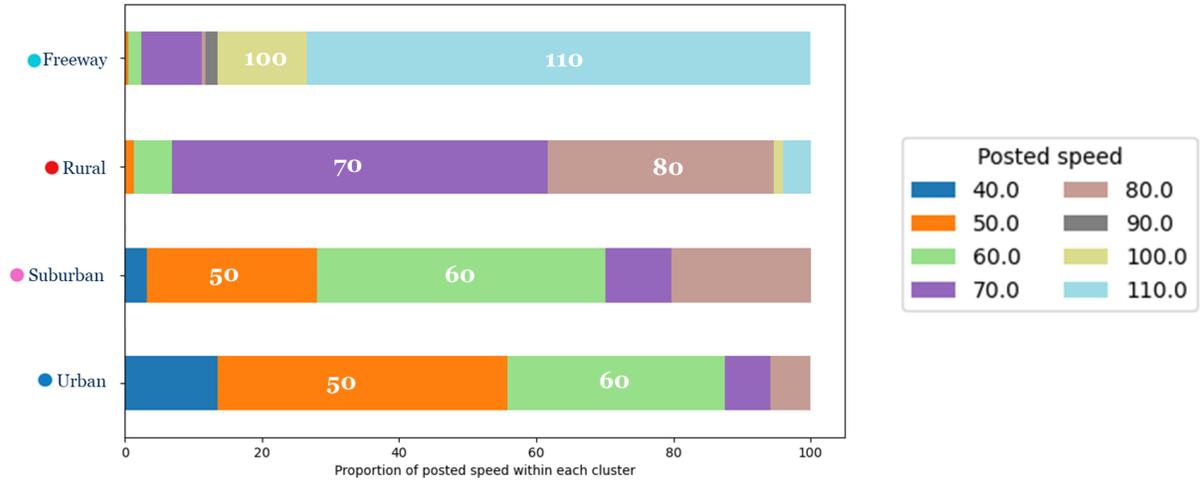


Figure 5.17-Distribution of posted speed within each cluster

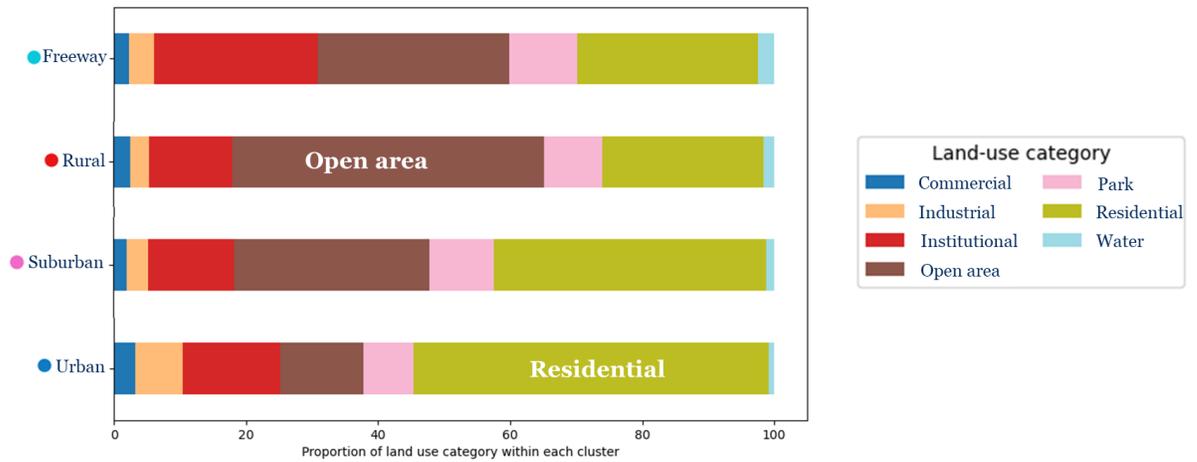


Figure 5.18-Distribution of land use category within each cluster

Chapter 6

6 Conclusions, Limitations, and Future Opportunities

6.1 Conclusions

The existing roadway network should be prepared to accommodate future mobility challenges, solutions, and technologies. It will be easier to plan appropriately for the future of transportation and urban development if we have a clear understanding of what we presently have on the ground. The performance of the current roadways is one of these transportation-related issues that should be taken under investigation. Hence, the purpose of this study was to analyze and compare the existing roadway performance metrics with the newly-calibrated one by using the emerging data sources. This study has allowed us to reproduce realistic measures of infrastructural supply in the static traffic assignment and gave us better insight of real-world context-specific congestion behaviour.

This research provides details on the existing roadway performance metrics in NCS16. For instance, the calibrated fundamental diagram for the sampled roadways shows that the existing free-flow speed in NCS16 is overestimated compared to real-world calibrated data. The reason for this could be that the free flow speed used in GTAModel-based Emme traffic assignment is the posted speed of the roadway. This number may be lower or higher than the observed free-flow speed. If highways are excluded, the capacity value appears to be well estimated. However, the capacity of highways is underestimated in NCS16.

6.2 Limitations of the Study

Investigating the reliability and usefulness of StreetLight Data as a data source for the calibration of the supply side of the travel demand models was one of the goals of this project. It is innovative to have a platform where traffic-related data is extracted from mobile crowd-sourced sensors. A broad spatial and temporal resolution of Streetlight data gives a unique opportunity to investigate any selected road segment or traffic analysis zone in the network. However, after digging deeper into the data, it appears that StreetLight Data generated metrics are not precise enough to model traffic behaviour. In some cases, the observed traffic volume data exceeds the capacity limits of the roadway. Although the StreetLight Data platform offers the calibration option for data, loop detector data was not available throughout the entire year of this study.

Average daily traffic (ADT) for each week of the year was required for this purpose. By importing average annual daily traffic (AADT) as a calibration value to StreetLight Data, it gives similar values for all weeks through the entire year for the segment under study, which is not the data required for our proposed framework. Overall, the only option available for use in this study was StreetLight Data volume estimates, which were not reliable, particularly for some highway sections.

The traditional performance model used in static macroscopic traffic models is not able to portray the inherent dynamic nature of congestion, including but not limited to queue accumulation and discharge, spillback, wave propagation, and capacity drop (Gentile et al., 2005). Hence, the dynamic traffic assignment was proposed to capture the complex traffic behaviour which was simplified in the conventional models. In DTA, route and departure time choices are modelled in a time-varying environment. The substantial data gathering, development, and calibration necessary for a dynamic model for the entire city is regarded to be excessively time-consuming and costly. Nevertheless, in several big urban regions, such as the GTHA, dynamic models are still under development. The roadway performance models in the dynamic models can be classified and calibrated using the framework developed in this study. The only difference is that speed-density data points should be used for calibration in comparison to travel time-flow data points used in VDF calibration. This procedure can all be done using the same dataset. Many additional supply and demand characteristics need to be calibrated in addition to this performance model in the dynamics models, which are out of the scope of this study.

6.3 Future Research Opportunities

Avenues for further research can be classified into these topics:

First, this study was conducted using processed location-based service (LBS) data, a new source of big data which is generated by GPS-equipped devices. The reported speed and flow values of this data source may not match/replicate the accurate traffic measure of the field. The reason for this is that sampled GPS-enabled probe vehicle data and machine learning techniques are used to estimate the speed and volume values. Therefore, one could use loop detector data along with Google Map API or video-based measurement as initial data to be imported into the framework suggested in this study.

Second, in regards to model spatial transferability, the proposed framework is not region-specific. It can be used anywhere else in the world. This processing pipeline can be run by any source of traffic data that has speed and volume values. It should be noted that the temporal resolution and availability of data must be sufficient to produce accurate estimates. To be more specific, intervals of data should be at most 1 hour, and the data should be available at least for 3 months. As machine learning methods are data-hungry, more days and finer temporal resolution of the data may lead to better performance.

Third, StreetLight Data, the data source used in this study, contains truck data for the road segments by the weighting class. However, this measurement is not a precise count of trucks; rather, it only depicts relative truck activity on the road segment. To evaluate mixed traffic behaviour, one can scale up Streetlight's relative truck activity data by using the ground-truth number of trucks for Ontario's highways. This data source was not available at the time of this study.

Fourth, a GTAModel-based Emme assignment run laid the groundwork for the selection of the road segment sample in this study. This could be a good first guess for identifying the road segments that are congested at some times of the day. However, there are some other reliable methods that can be used for this. For instance, one could use historical (typical) traffic data from Google Maps, which is accessible via their website/API. In this Google Maps feature, four different colours are used, with the red colour indicating the congested road segment.

Fifth, the congested regime of VDFs, for which we lack empirical data, can be calibrated using observed route choice data. One may determine which VDF curves could replicate the same congestion pattern throughout the transportation system by comparing the static traffic assignment's assigned route with the observed route choice behaviour. The historical route choice behaviour is available via the StreetLight Data platform. There were two issues in using this data for this thesis. First, the StreetLight's observed route between the origin and the destination contained some discontinuities, some of which, but not all, could be manually corrected. Second, the routes between each pair of origin and destination are not reported by the Emme static traffic assignment, which is required as a component in the defined optimization problem. Utilizing Aimsun software, which produces the OD route for static traffic assignment, may be the answer to

this problem. At the time of this study, GTAModel an Aimsun-based traffic assignment model for the GTHA was under construction.

Sixth, the phenomenon of the capacity drop is not taken into account for the calibration of the roadway fundamental diagram. One may calibrate the fundamental diagram using a discontinuous two regime traffic flow fundamental equation, which considers this phenomenon. Also, the jam density is considered to be 144 veh/km/lane in this study, which should be investigated in more detail.

Seventh, since this clustering framework is built using a multistage unsupervised machine learning method, there is no tool to assess its performance. One may use the calendar and weather data to label this unlabeled speed-flow data and create a confusion matrix to see how accurate this approach is at recognizing holidays and weather events from normal days.

Eight, sensitivity analysis tests can be done on the model output based on using the distributions of parameters and input variables. This can be accomplished using a variety of techniques, such as the random resampling technique or Monte Carlo simulation. At least three metrics can be used to examine the model's output: vehicle kilometers on a single link and throughout the entire network, travel resistance within the network, and average speed at the link level (Manzo et al., 2014). Furthermore, the variability of uncertainty within different road functional classes can be assessed (Manzo, 2013).

Ninth, one may test different VDFs in EMME traffic assignment software and compare the results with the travel time studies of the Ontario Ministry of Transportation.

Finally, building on this study, one can develop area-based cost flow functions, known as macroscopic cost flow (MCF) functions to explore network topological effects on network traffic performance (Wong & Wong, 2016).

References

- Akçelik, R. (1981). Traffic signals: Capacity and timing analysis. *Australian Road Research Board Research Report 123 (7th Print: 1998)*. [https://doi.org/10.1016/0191-2607\(81\)90135-7](https://doi.org/10.1016/0191-2607(81)90135-7)
- Akçelik, R. (1991). Travel Time Functions for Transport Planning Purposes: Davidson's Function, Its Time Dependent Form and Alternative Travel Time Function. *Australian Road Research*, 21, 44–59.
- Azimi, M., & Zhang, Y. (2010). Categorizing freeway flow conditions by using clustering methods. *Transportation Research Record*, 2173, 105–114. <https://doi.org/10.3141/2173-13>
- Bliemer, M. C. J., Raadsen, M. P. H., Brederode, L. J. N., Bell, M. G. H., Wismans, L. J. J., & Smith, M. J. (2017). Genetics of traffic assignment models for strategic transport planning. *Transport Reviews*, 37(1), 56–78. <https://doi.org/10.1080/01441647.2016.1207211>
- Boyce, D. E., Janson, B. N., & Eash, R. W. (1981). the Effect on Equilibrium Assignment of Different Congestion Functions. *Transportation Research Part A: General*, 15(3), 223–232.
- Branston, D. (1976). *Link Capacity Functions: a Review*. 10, 223–236.
- Brilon, W., Geistefeldt, J., & Regler, M. (2005). Reliability of Freeway Traffic Flow. *Transportation and Traffic Theory*, July, 125–144. <https://doi.org/10.1016/b978-008044680-6/50009-x>
- Bureau of Public Roads. (1964). Traffic Assignment Manual. *Washington: US Department of Commerce, Urban Planning Division*.
- Caceres, H., Hwang, H., & He, Q. (2016). Estimating freeway route travel time distributions with consideration to time-of-day, inclement weather, and traffic incidents. *Journal of Advanced Transportation*. <https://doi.org/10.1002/atr>
- Casey, G., Zhao, B., Kumar, K., & Soga, K. (2020). Context-specific volume-delay curves by combining crowd-sourced traffic data with automated traffic counters: A case study for London. *Data-Centric Engineering*, 1(3). <https://doi.org/10.1017/dce.2020.18>
- Cetin, M., Foytik, P., S. Son, A., Khattak, J., Robinson, R. M., & Lee, and J. (2012). Calibration of Volume-Delay Functions for Traffic Assignment in Travel Demand Models. *Paper Presented at the 91st Annual Meeting of Transportation Research Board. Washington, DC*.
- Cheah, L., Dalton, P., & Hariri, O. (1992). Alternative Approaches to Volume-Delay-Function formulation. *7th Annual EMME/2 International User's Conference*, 1–6.
- Cheng, Q., Liu, Z., Lin, Y., & Zhou, X. (Simon). (2021). An s-shaped three-parameter (S3) traffic stream model with consistent car following relationship. *Transportation Research*

- Part B: Methodological*, 153(October), 246–271. <https://doi.org/10.1016/j.trb.2021.09.004>
- Das, A. K., & Chilukuri, B. R. (2020). Link cost function and link capacity for mixed traffic networks. *Transportation Research Record*, 2674(9), 38–50. <https://doi.org/10.1177/0361198120926454>
- Davidson, K. B. (1966). A flow travel time relationship for use in transportation planning. *Australian Road Research Board (ARRB) Conference*, 3(1), 183–194. <http://trid.trb.org/view/1209266>
- Davidson, K. B. (1978). Theoretical Basis of a Flow-Travel Time Relationship for Use in Transportation Planning. *Australian Road Research*, 8(1), 32–35.
- Dowling, R., & Skabardonis, A. (1993). Improving Average Travel Speeds Estimated by Planning Models. *Transportation Research Record*, 1366, 68–74.
- Dudek, C. L., Messer, C. J., & Nuckles, N. B. (1974). Incident detection on urban freeways. *Transportation Research Record*, 495, 12–24. <http://dx.doi.org/>
- Fukushima, M. (1984). A modified Frank-Wolfe algorithm for solving the traffic assignment problem. *Transportation Research Part B*, 18(2), 169–177. [https://doi.org/10.1016/0191-2615\(84\)90029-8](https://doi.org/10.1016/0191-2615(84)90029-8)
- Gentile, G., Meschini, L., & Papola, N. (2005). Macroscopic arc performance models with capacity constraints for within-day dynamic traffic assignment. *Transportation Research Part B: Methodological*, 39(4), 319–338. <https://doi.org/10.1016/j.trb.2004.04.005>
- Gentile, G., & Noekel, K. (2009). Linear user cost equilibrium: the new algorithm for traffic assignment in VISUM. In *Proceedings of the European Transport Conference, Noordwijkerhout, The Netherlands*. <https://doi.org/10.1080/18128602.2012.691911>
- Golding, S. (1977). On Davidson's Flow/Travel Time Relationship. *Australian Road Research*, 7, 36–37.
- Gomes, G., & Horowitz, R. (2009). Automatic Calibration of the Fundamental Diagram and Empirical Observations on Capacity. *Transportation Research Board, January*, 1–14.
- Government of Canada. (2022). *Open Government Municipal Boundaries dataset*. <https://jonathancritchley.ca/gtha.html>
- Gu, Z., Saberi, M., Sarvi, M., & Liu, Z. (2018). A big data approach for clustering and calibration of link fundamental diagrams for large-scale network simulation applications. *Transportation Research Part C: Emerging Technologies*, 94, 151–171. <https://doi.org/10.1016/j.trc.2017.08.012>
- Hansen, S., & Bertini, R. L. (2005). Using Archived ITS Data to Improve Travel Demand Forecasting. *Department of Civil and Environmental Engineering and Nohad A. Touloukian School of Urban Studies and Planning, Portland State University, Oregon*.

- Highway Capacity Manual. (2000). Highway Capacity Manual. *Washington: Transportation Research Board, National Research Council.*
- Horowitz, A. (1991). Delay-Volume Relations for Travel Forecasting: Based on the 1985 Highway Capacity Manual. *FHWA, U.S. Department of Transportation*, 64. https://www.fhwa.dot.gov/planning/tmip/publications/other_reports/delay_volume_relations/ch09.cfm
- Huntsinger, L. F., & Roupail, N. M. (2011). Bottleneck and queuing analysis: Calibrating volume-delay functions of travel demand models. *Transportation Research Record*, 2255, 117–124. <https://doi.org/10.3141/2255-13>
- Ibrahim, A. T., & Hall, F. L. (1994). Effect of adverse weather conditions on speed-flow-occupancy relationships. *Transportation Research Record*, 1457, 184–191.
- Irwin, N. A., Dodd, N., & Von Cube, G. H. (1961). Capacity Restraint in Assignment Programs. *Highway Research Board Bulletin, 40th Annual Meeting of the Highway Research Board*, 297, 109–127.
- Jastrzebski, W. (2000). Volume delay functions. *In Proceedings of the 15th International EMME/2 Users Group Conference, Vancouver, Canada.*
- Kalair, K., & Connaughton, C. (2021). Anomaly detection and classification in traffic flow data from fluctuations in the flow–density relationship. *Transportation Research Part C: Emerging Technologies*, 127(April). <https://doi.org/10.1016/j.trc.2021.103178>
- Kaplan, E. L., & Meier, P. (1958). Nonparametric Estimation from Incomplete Observations
Author (s): E . L . Kaplan and Paul Meier Source : Journal of the American Statistical Association , Vol . 53 , No . 282 (Jun . , 1958) , pp . 457- Published by : American Statistical Association Sta. *American Statistical Association*, 53(282), 457–481.
- Kianfar, J., & Edara, P. (2013). A Data Mining Approach to Creating Fundamental Traffic Flow Diagram. *Procedia - Social and Behavioral Sciences*, 104, 430–439. <https://doi.org/10.1016/j.sbspro.2013.11.136>
- Kucharski, R., & Drabicki, A. (2017). Estimating macroscopic volume delay functions with the traffic density derived from measured speeds and flows. *Journal of Advanced Transportation*, 32, 1–10. <https://doi.org/10.1155/2017/4629792>
- Levin, M., & Krause, G. M. (1978). Incident detection: A Bayesian Approach. *Transportation Research Record*, 682, 52–58.
- Li, H., Wang, Q., & Xiong, W. (2021). New Model of Travel-Time Prediction Considering Weather Conditions: Case Study of Urban Expressway. *Journal of Transportation Engineering, Part A: Systems*, 147(3), 1–8. <https://doi.org/10.1061/jtepbs.0000491>
- Lu, Z., Meng, Q., & Gomes, G. (2016). Estimating link travel time functions for heterogeneous traffic flows on freeways. *Journal of Advanced Transportation*, 50(8), 1683–1698. <https://doi.org/10.1002/atr.1423>

- Malatest. (2018). *TTS 2016: 2016, 2011, 2006, 1996 and 1986 travel summaries for the Greater Toronto & Hamilton Area*. 102.
http://www.dmg.utoronto.ca/pdf/tts/2016/2016TTS_Summaries_GTHA.pdf
- Manzo, S. (2013). Investigating uncertainty in BPR formula parameters: a case study. *DTU Transport, Technical University of Denmark*.
www.ectri.org/YRS13/Documents/Papers/.../YRS13_Session1b_Manzo_DTU-Paper.p
- Manzo, S., Nielsen, O. A., & Prato, C. G. (2014). Effects of uncertainty in speed-flow curve parameters on a large-scale model: Case study of the danish national model. *Transportation Research Record*, 2429, 30–37. <https://doi.org/10.3141/2429-04>
- Mashros, N., Ben-Edigbe, J., Alhassan, H. M., & Hassan, S. A. (2014). Investigating the impact of rainfall on travel speed. *Jurnal Teknologi*, 71(3), 33–38.
<https://doi.org/10.11113/jt.v71.3756>
- Maze, T. H., Agarwal, M., & Burchett, G. (2006). Whether weather matters to traffic demand, traffic safety, and traffic operations and flow. *Transportation Research Record*, 1948, 170–176. <https://doi.org/10.3141/1948-19>
- Moridpour, S., Mazloumi, E., & Mesbah, M. (2015). Survey and empirical evaluation of nonhomogeneous arrival process models with taxi data. *Journal of Advanced Transportation*, 49(4), 535–552. <https://doi.org/10.1002/atr>
- Moses, R., Mtoi, E., McBean, H., & Ruegg, S. (2013). Development of Speed Models for Improving Travel Forecasting and Highway Performance Evaluation. *Final Report Project No. BDK83- 977-14, Florida Department of Transportation*.
http://www.fsutmsonline.net/images/uploads/reports/Speed_Modeling_Final_Report.pdf
- Mosher Jr, W. W. (1963). A capacity-restraint algorithm for assigning flow to a transport network. *Highway Research Record*, 6, 41–70.
- Mtoi, E. T., & Moses, R. (2014). Calibration and Evaluation of Link Congestion Functions: Applying Intrinsic Sensitivity of Link Speed as a Practical Consideration to Heterogeneous Facility Types within Urban Network. *Journal of Transportation Technologies*, 04(02), 141–149. <https://doi.org/10.4236/jtts.2014.42014>
- Müller, S., & Schiller, C. (2015). Improvement of the volume-delay function by incorporating the impact of trucks on traffic flow. *Transportation Planning and Technology*, 38(8), 878–888. <https://doi.org/10.1080/03081060.2015.1079388>
- Neuhold, R., & Fellendorf, M. (2014). Volume Delay Functions Based on Stochastic Capacity. *Transportation Research Record: Journal of the Transportation Research Board*, 2421(1), 93–102. <https://doi.org/10.3141/2421-11>
- Nielsen, O. A., & Jørgensen, R. M. (2008). Estimation of speed-flow and flow-density relations on the motorway network in the greater Copenhagen region. *IET Intelligent Transport Systems*, 2(2), 120–131. <https://doi.org/10.1049/iet-its:20070024>

- Overgaard, K. R. (1967). Urban Transportation Planning: Traffic Estimation. *Traffic Quarterly*, 21, 197–218.
- Patriksson, P. (1994). The Traffic Assignment Problem: Models and Methods. In *Utrecht: VSP*.
- Payne, H., Helfenbein, E. D., & Knobel, H. C. (1976). Development and Testing of Incident Detection Algorithms. In *Research Methodology and Detailed Results. Tech. Rep. FHWA-RD-76-20, Federal Highway Administration, Washington D.C.*
<https://rosap.ntl.bts.gov/view/dot/744>
- Persaud, B. N., & Hall, F. L. (1989). Catastrophe theory and patterns in 30-second freeway traffic data- Implications for incident detection. *Transportation Research Part A: General*, 23(2), 103–113. [https://doi.org/10.1016/0191-2607\(89\)90071-X](https://doi.org/10.1016/0191-2607(89)90071-X)
- Qiu, E., Viridi, N., Grzybowska, H., & Waller, T. (2022). Recalibration of the BPR function for the strategic modelling of connected and autonomous vehicles. *Transportmetrica B*, 10(1), 779–800. <https://doi.org/10.1080/21680566.2022.2040063>
- Qu, X., Wang, S., & Zhang, J. (2015). On the fundamental diagram for freeway traffic: A novel calibration approach for single-regime models. *Transportation Research Part B: Methodological*, 73, 91–102. <https://doi.org/10.1016/j.trb.2015.01.001>
- Rakha, H., & Crowther, B. (2002). Comparison of Greenshields, Pipes, and Van Aerde car-following and traffic stream models. *Transportation Research Record*, 1802, 248–262. <https://doi.org/10.3141/1802-28>
- Raw, J., & Munn, D. (2009). Investigation of Speed–Flow Relations and Estimation of Akcelik Volume Delay Functions for Travel Demand Models in Virginia. *Presented at 12th TRB Conference on the Application of Transportation Planning Methods, Transportation Research Board of the National Academies, Washington, D.C.*
- Saric, A., Albinovic, S., Dzebo, S., & Pozder, M. (2019). Volume-Delay Functions : A Review. *Proceedings of the International Symposium on Innovative and Interdisciplinary Applications of Advanced Technologies (IAT), Volume 2*, 3–12.
- Šarić, A., & Lovrić, I. (2021). Improved Volume–Delay Function for Two-Lane Rural Highways with Impact of Road Geometry and Traffic-Flow Heterogeneity. *Journal of Transportation Engineering, Part A: Systems*, 147(10), 1–13. <https://doi.org/10.1061/jtepbs.0000575>
- Saw, K., Katti, B. K., Joshi, G., Dqg, S., Lwk, D., Frqjhwlrq, W. F., Rq, D., Urdg, W. K. H., Ri, G., Kdv, F., Dqg, L., Wlph, W., Olqn, R. Q., Qrw, G. R., Zlwk, Y. D. U., Olqn, W. K. H., Wuls, Å., Kdvh, P., Nqrzohgjh, S., ... Dvvljqphqw, W. F. (2015). Review Paper: Literature Review of Traffic Assignment: Static and Dynamic. *International Journal of Transportation Engineering*, 2(4), 339–347.
- Sheffi, Y. (1985). Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods. In *Englewood Cliffs, NJ: Prentice-Hall*.
- Singh, R., & Dowling, R. (1999). Improved speed-flow relationship: application to transportation

planning models. *Proceedings of the Seventh TRB Conference on the Application of Transportation Planning Methods*, 340–349.
<http://trid.trb.org/view.aspx?id=734743%5Cnhttp://docs.trb.org/00939750.pdf>

Skabardonis, A., & Dowling, R. (1997). Improved Speed-Flow Relationships for Planning Applications. *Transportation Research Record: Journal of the Transportation Research Board*, 1572(970996), 18–23.

Smock, R. (1962). an Iterative Assignment Approach To Capacity Restraint on Arterial Networks. *Highway Research Board Bulletin*, 347, pp 60-66.
<https://trid.trb.org/view/120762>

Spiess, H. (1990). Technical Note — Conical Volume-Delay Functions. *Transportation Science*, 24, No. 2, 153–158.

Statistics Canada. (2021). *Population and Dwelling Count Highlight Tables, 2021 Census*.
<https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/index.cfm?Lang=E>

Sun, L., & Zhou, J. (2005). Development of multiregime speed-density relationships by cluster analysis. *Transportation Research Record*, 65(1934), 64–71. <https://doi.org/10.3141/1934-07>

Taylor, M. A. P. (1977). Parameter Estimation and Sensitivity of Parameter Values in a Flow-rate/Travel time Relation. *Transportation Science*, 11, 275–292.

Taylor, M. A. P. (1984). a Note on Using Davidson’s Function in Equilibrium Assignment. *Transportation Research Part B: Methodological*, 18, 181–199.

Thomas, J., Srinivasan, K. K., & Arasan, V. T. (2012). Vehicle class wise speed-volume models for heterogeneous traffic. *Transport*, 27(2), 206–217.
<https://doi.org/10.3846/16484142.2012.697442>

Tisato, P. (1991). Suggestions for an improved Davidson travel time function. *Australian Road Research*, 21(2), 85–100.

UofT Travel Modelling Group (TMG). (2017). *GTHA 2016 EMME Network Coding Standard*. 10. https://tmg.utoronto.ca/Account/Memos/2016_Network_Coding_Standard.pdf

UofT Travel Modelling Group (TMG). (2022). *GTAmodeL V4.0*.
<https://tmg.utoronto.ca/doc/1.4/gtamodel/index.html>

Wang, D., Ozbay, K., & Bian, Z. (2020). A Mixture Model-based Clustering Method for Fundamental Diagram Calibration Applied in Large Network Simulation. *2020 IEEE 23rd International Conference on Intelligent Transportation Systems, ITSC 2020*.
<https://doi.org/10.1109/ITSC45102.2020.9294346>

Wang, Z., & Liu, C. (2005). An empirical evaluation of the loop detector method for travel time delay estimation. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 9(4), 161–174. <https://doi.org/10.1080/15472450500237254>

- Wong, W., & Wong, S. C. (2016). Network topological effects on the macroscopic Bureau of Public Roads function. *Transportmetrica A: Transport Science*, 12(3), 272–296. <https://doi.org/10.1080/23249935.2015.1129650>
- Wu, X., Dutta, A., Zhang, W., Zhu, H., Livshits, V., & Zhou, X. (2020). Characterization and calibration of volume-to-capacity ratio in volume- delay functions on freeways based on a queue analysis approach. *Proceedings of the 100th Annual Meeting of Transportation Research Board*.
- Xia, J., & Chen, M. (2007a). A nested clustering technique for freeway operating condition classification. *Computer-Aided Civil and Infrastructure Engineering*, 22(6), 430–437. <https://doi.org/10.1111/j.1467-8667.2007.00498.x>
- Xia, J., & Chen, M. (2007b). Defining traffic flow phases using intelligent transportation systems-generated data. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 11(1), 15–24. <https://doi.org/10.1080/15472450601122322>
- Xia, J., Huang, W., & Guo, J. (2012). A clustering approach to online freeway traffic state identification using ITS data. *KSCE Journal of Civil Engineering*, 16(3), 426–432. <https://doi.org/10.1007/s12205-012-1233-1>
- Xiong, H., & Davis, G. A. (2009). Field evaluation of model-based estimation of arterial link travel times. *Transportation Research Record*, 2130, 149–157. <https://doi.org/10.3141/2130-18>
- Yang, H., & Qiao, F. (1998). Neural network approach to classification of traffic flow states. *Journal of Transportation Engineering*, 124(6), 521–525. [https://doi.org/10.1061/\(ASCE\)0733-947X\(1998\)124:6\(521\)](https://doi.org/10.1061/(ASCE)0733-947X(1998)124:6(521))
- Yuan, Z., Zhou, X., & Yang, T. (2018). Hetero-ConvLSTM: A deep learning approach to traffic accident prediction on heterogeneous spatio-temporal data. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 984–992. <https://doi.org/10.1145/3219819.3219922>
- Yun, S., White, W. W., Lamb, D. R., & Wu, Y. (2005). Accounting for the impact of heavy truck traffic in volume-delay functions in transportation planning models. *Transportation Research Record*, 1931, 8–17. <https://doi.org/10.3141/1931-02>
- Zhang, J., Qu, X., & Wang, S. (2018). Reproducible generation of experimental data sample for calibrating traffic flow fundamental diagram. *Transportation Research Part A: Policy and Practice*, 111(March), 41–52. <https://doi.org/10.1016/j.tra.2018.03.006>
- Zhao, F., Fu, L., Zhong, M., Liu, S., Wang, X., Huang, J., & Ma, X. (2020). Development and Validation of Improved Impedance Functions for Roads with Mixed Traffic Using Taxi GPS Trajectory Data and Simulation. *Journal of Advanced Transportation*, 2020. <https://doi.org/10.1155/2020/7523423>
- Zhao, L., & I-Jy Chien, S. (2012). Analysis of Weather Impact on Travel Speed and Travel Time Reliability. *The Twelfth COTA International Conference of Transportation Professionals*.

Zhu, L., Guo, F., Krishnan, R., & Polak, J. W. (2018). A Deep Learning Approach for Traffic Incident Detection in Urban Networks. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, 2018-Novem*, 1011–1016.
<https://doi.org/10.1109/ITSC.2018.8569402>

Zhu, L., Krishnan, R., Sivakumar, A., Guo, F., & Polak, J. W. (2019). Traffic Monitoring and Anomaly Detection based on Simulation of Luxembourg Road Network. *2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019*, 382–387.
<https://doi.org/10.1109/ITSC.2019.8917015>

Appendix 1

Highway (VDF code: 11)

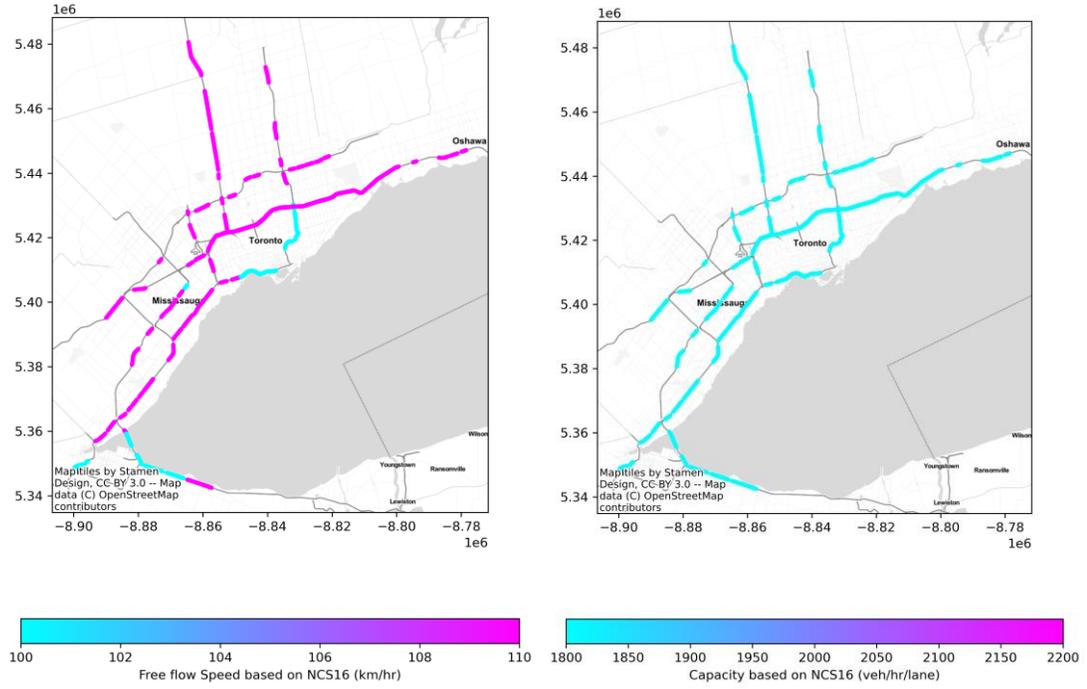


Figure 6.1-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 11

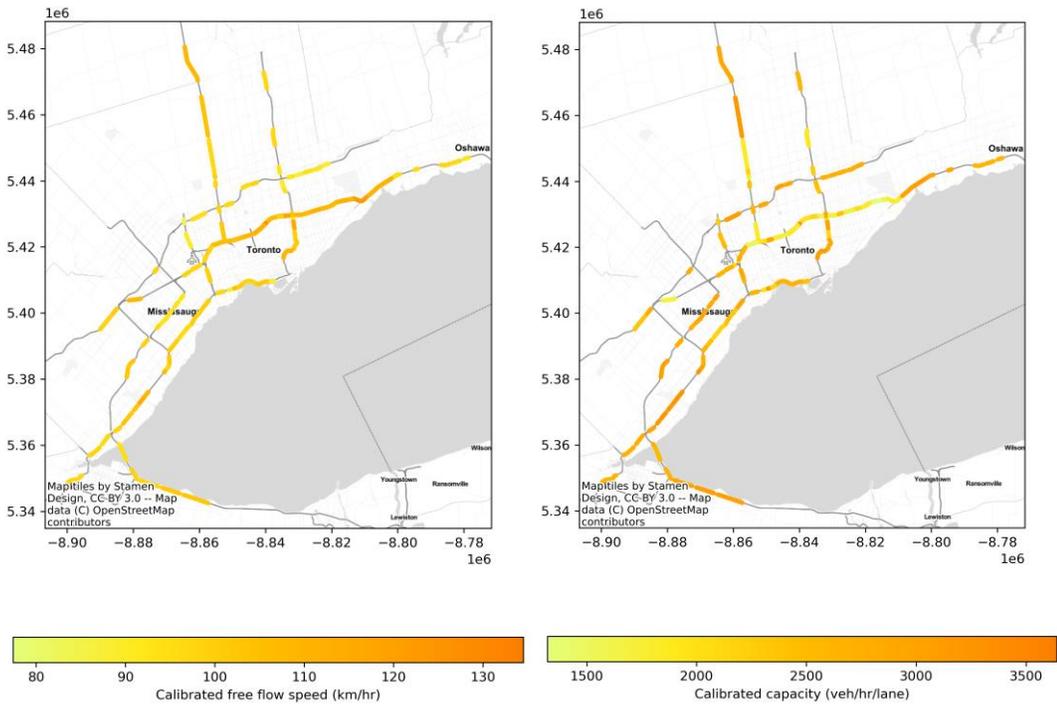


Figure 6.2-Calibrated capacity (right) and free flow speed (left), VDF code: 11

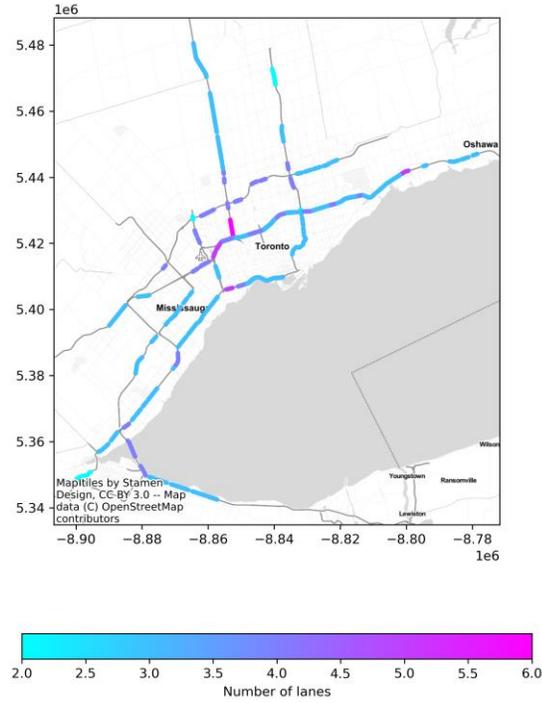


Figure 6.3-Number of lanes, VDF code: 11

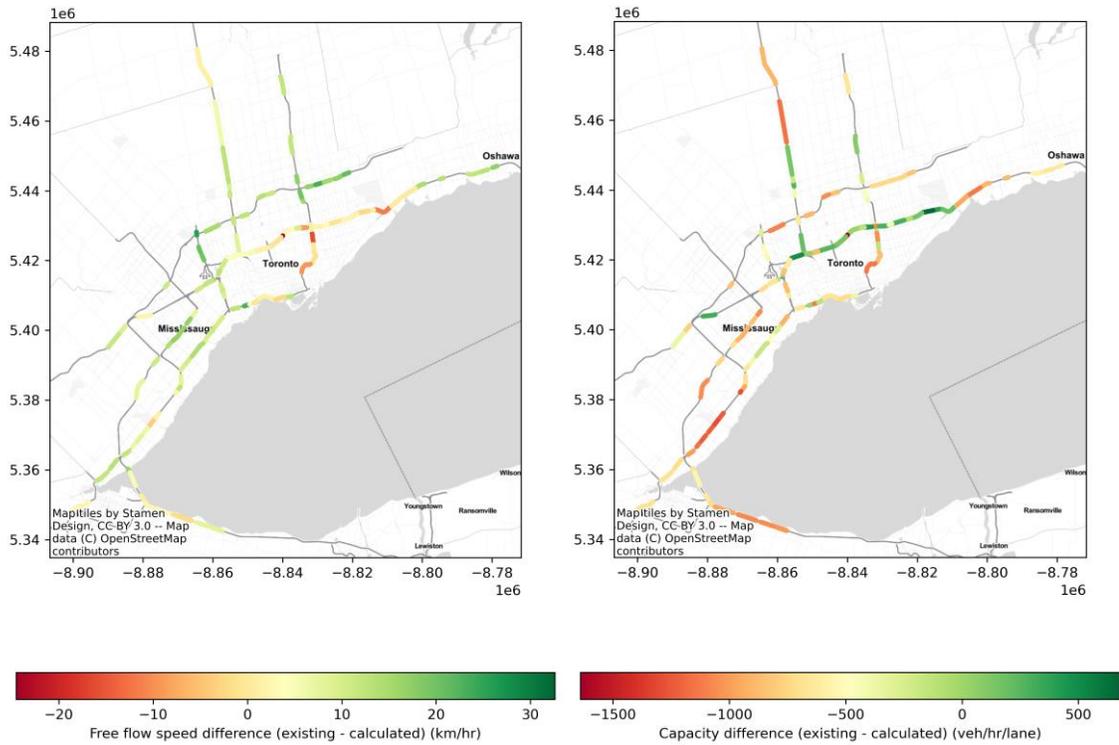


Figure 6.4-Existing and calculated values difference of free flow speed and capacity, VDF code: 11

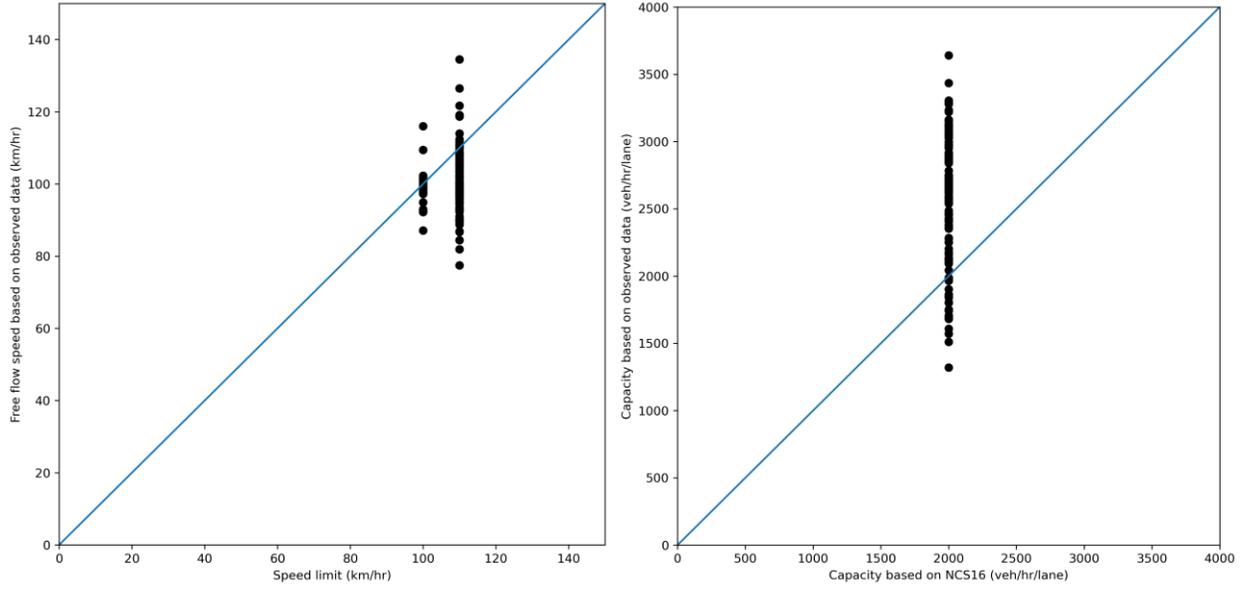


Figure 6.5-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 11

Highway Ramps (VDF code: 13)

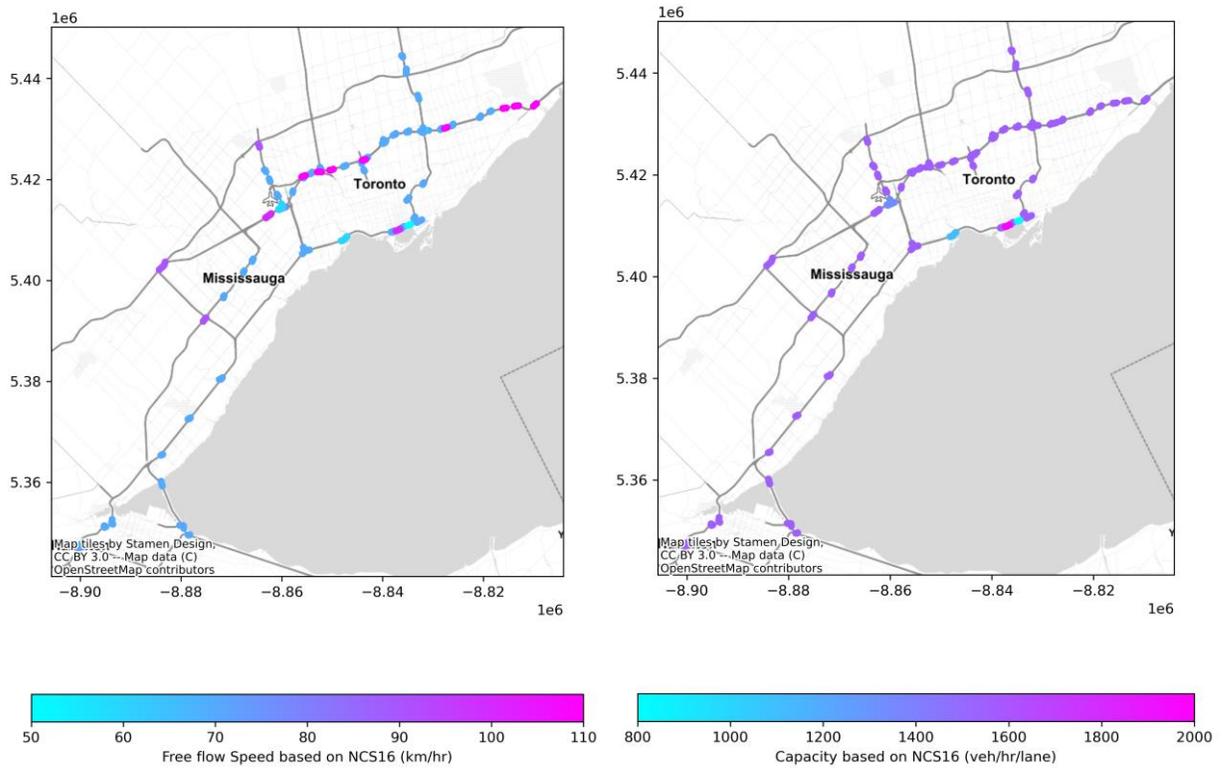


Figure 6.6-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 13

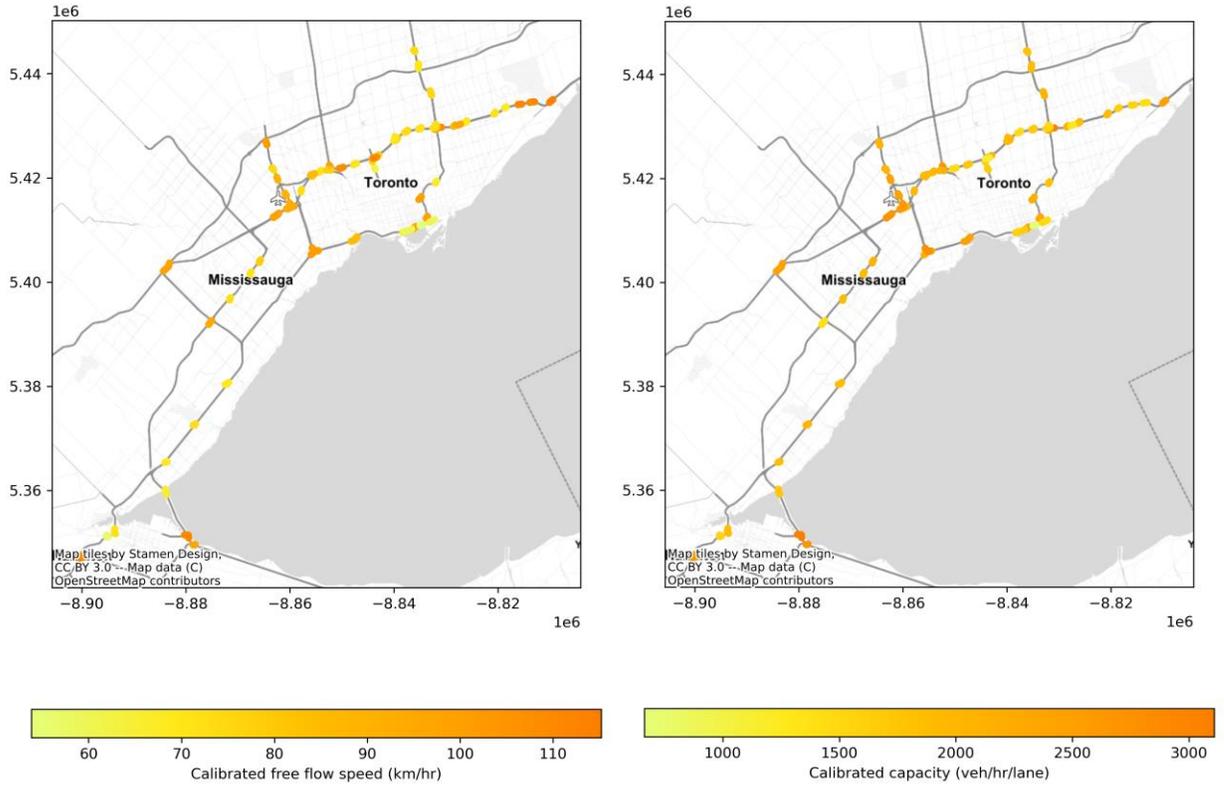


Figure 6.7-Calibrated capacity (right) and free flow speed (left), VDF code: 13

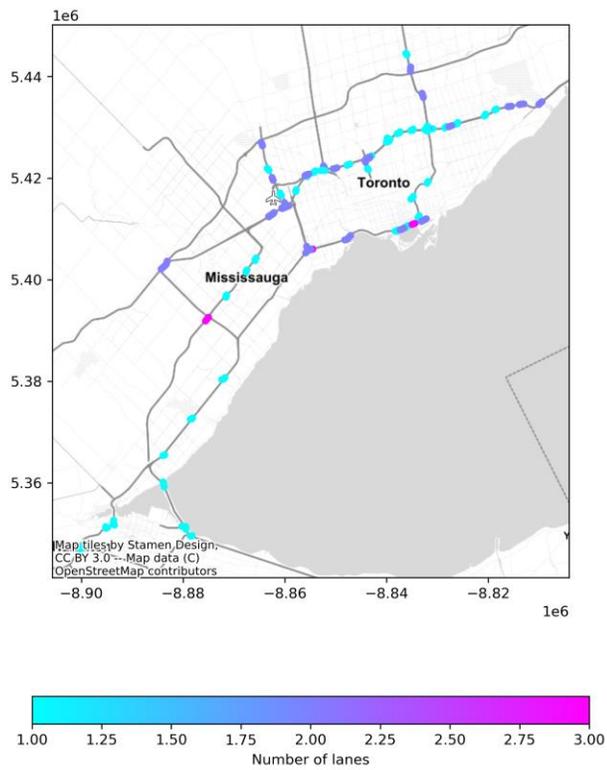


Figure 6.8-Number of lanes, VDF code: 13

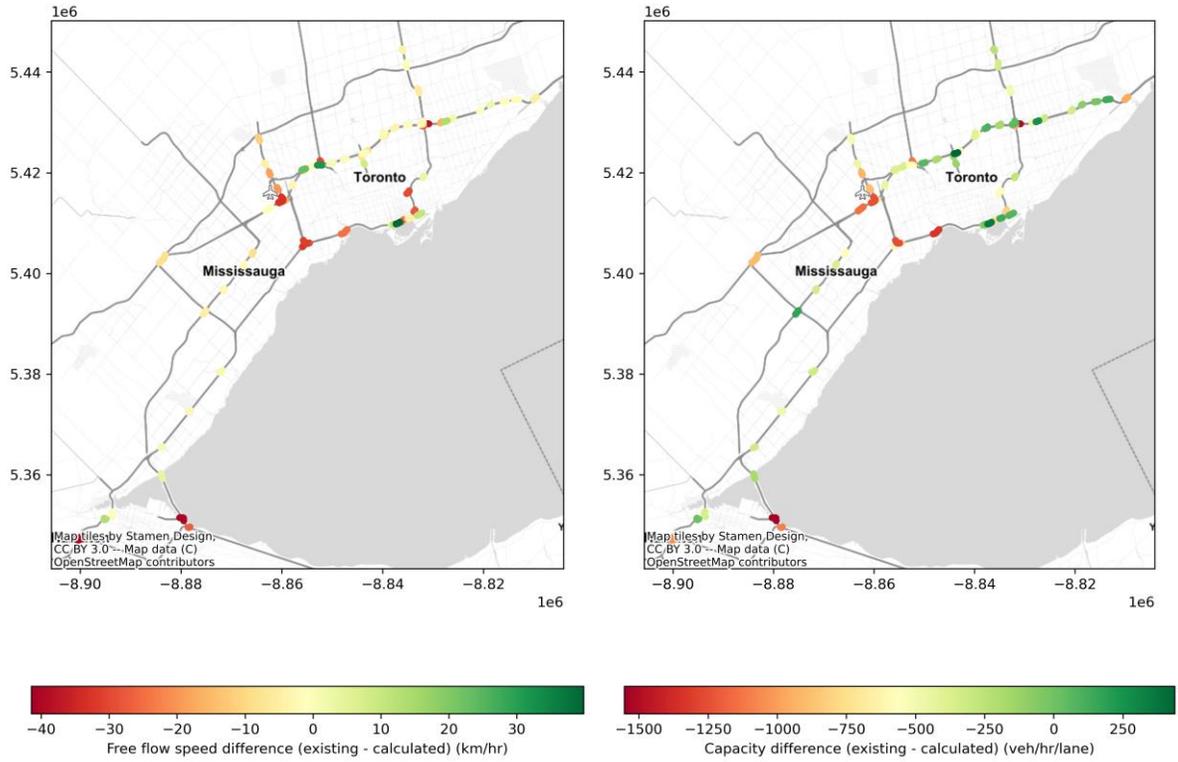


Figure 6.9--Existing and calculated values difference of free flow speed and capacity, VDF code: 13

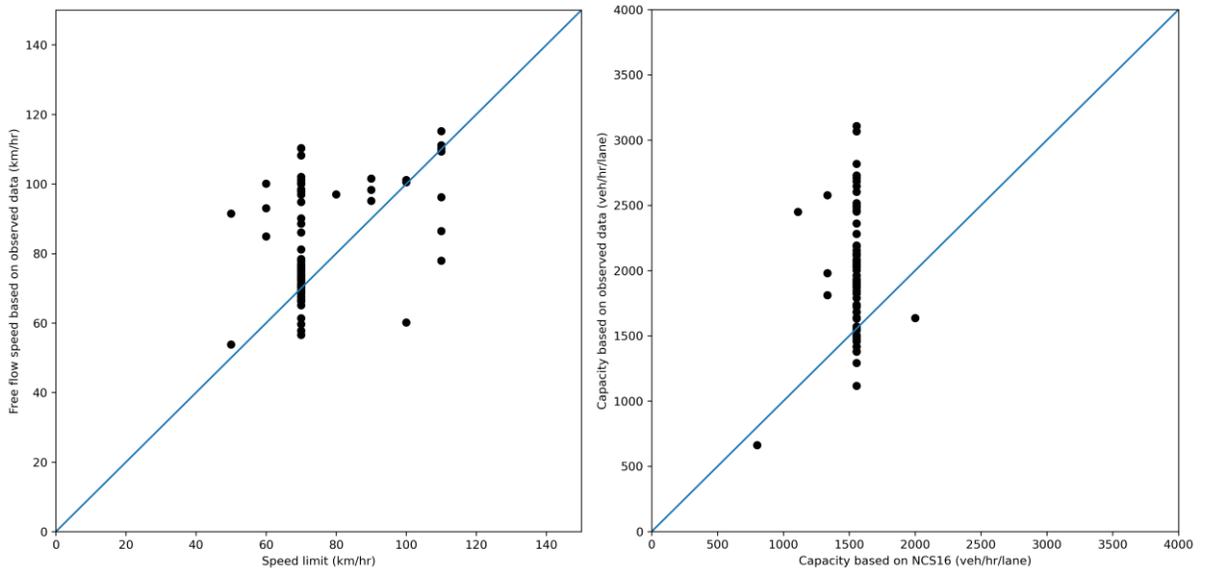


Figure 6.10-Comparison curve between existing and calibrated free flow speed (left) and capacity (right), VDF code: 13

Rural Long Distance Arterials (VDF code: 20)

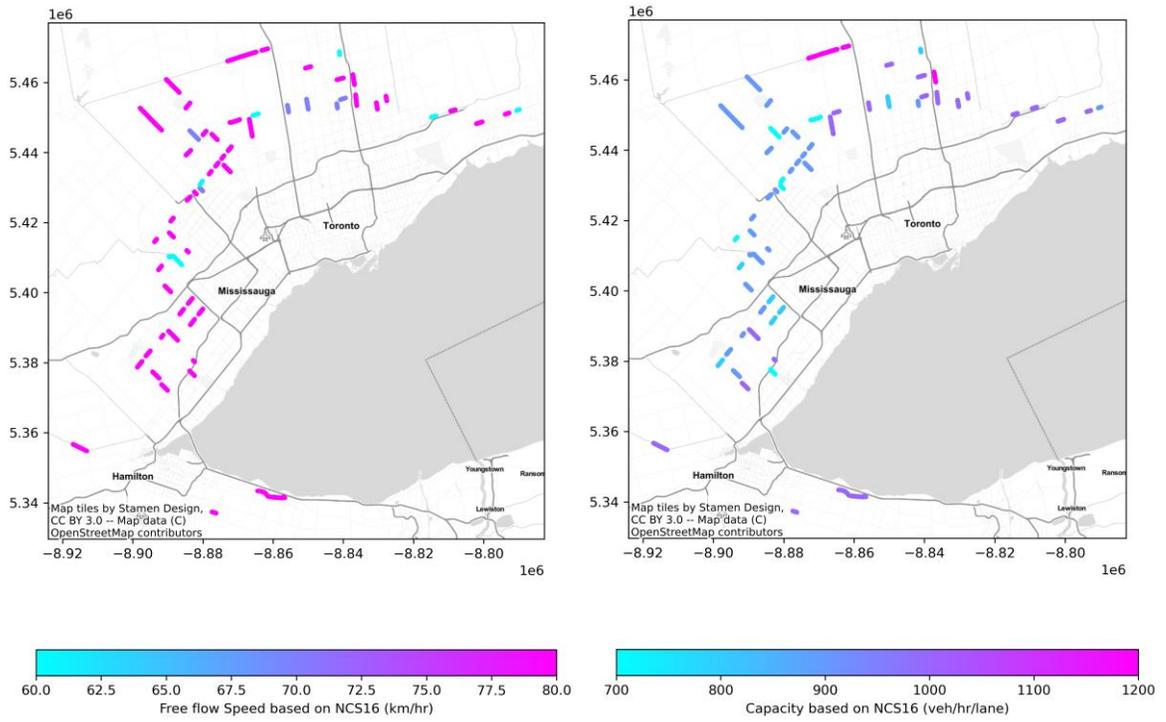


Figure 6.11-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 20

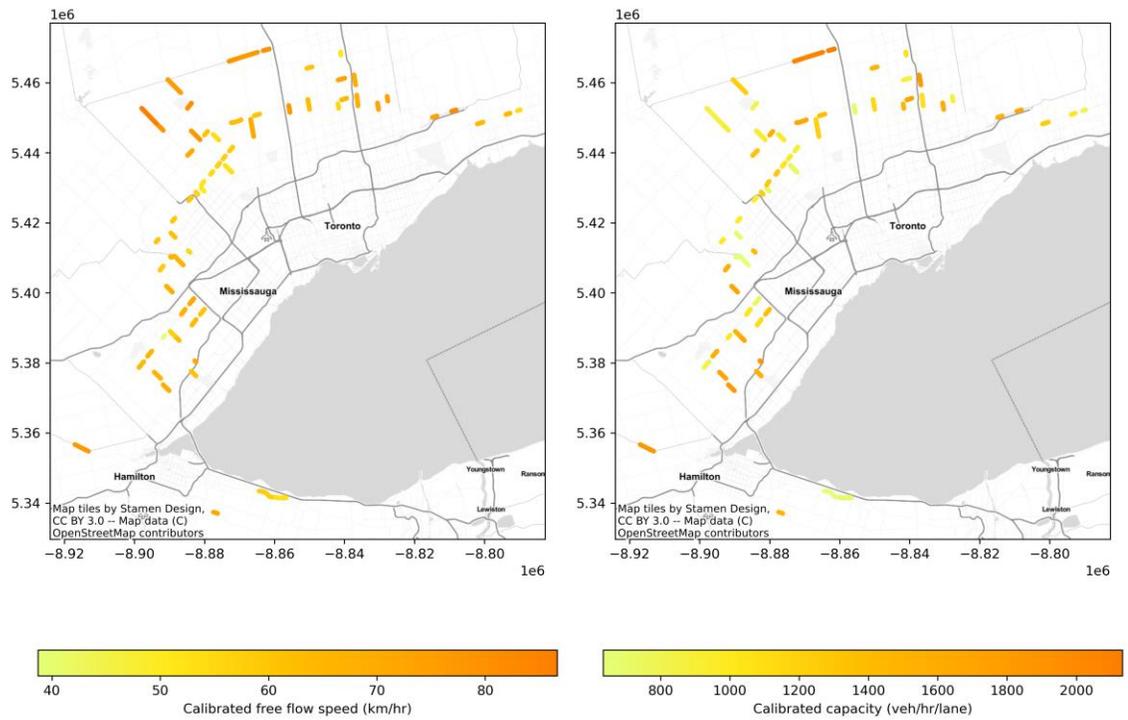


Figure 6.12-Calibrated capacity (right) and free flow speed (left), VDF code: 20

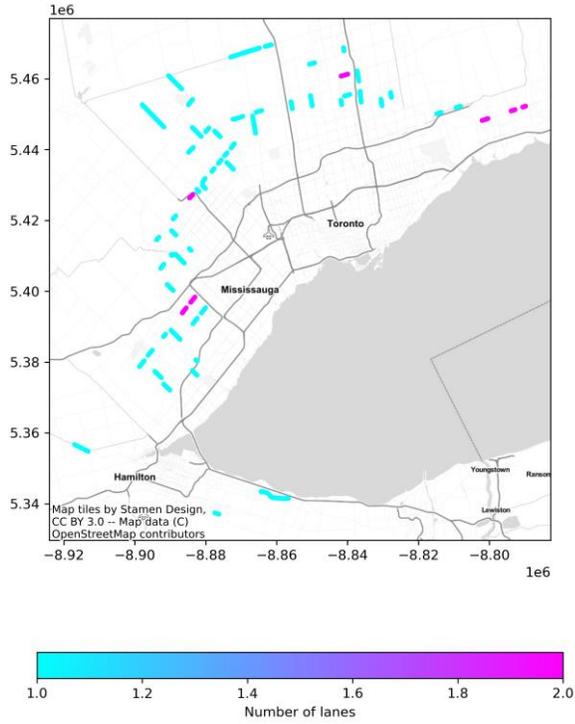


Figure 6.13-Number of lanes, VDF code: 20

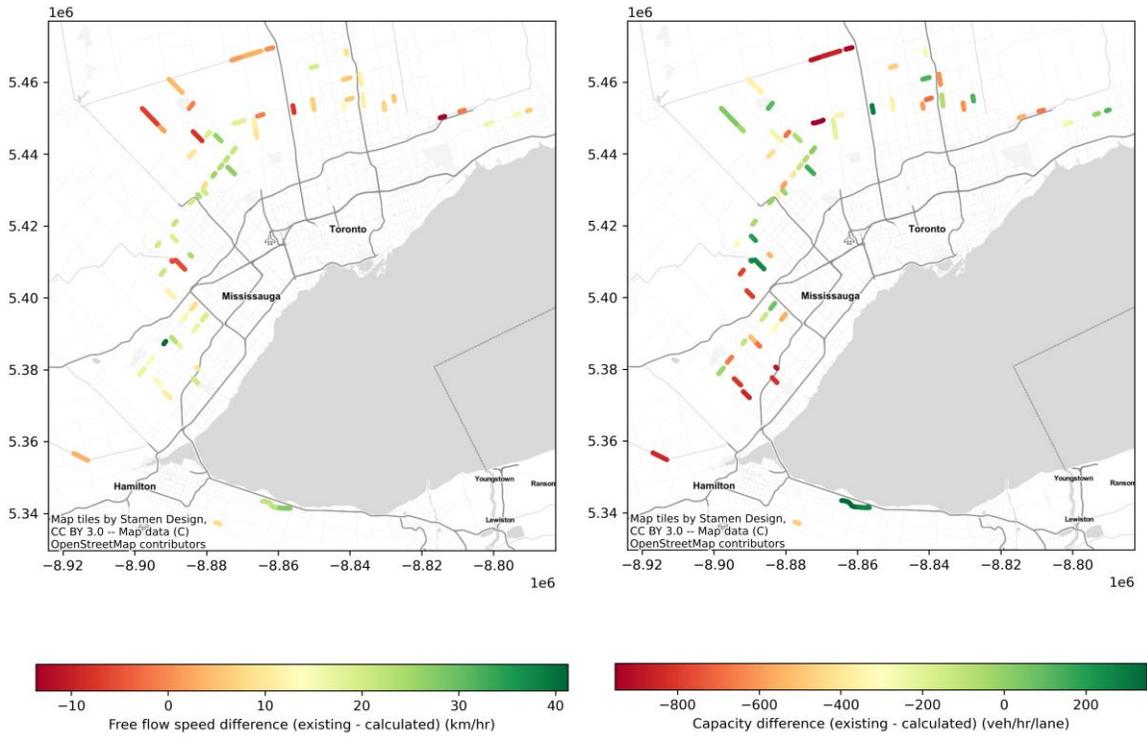


Figure 6.14-Existing and calculated values difference of free flow speed and capacity, VDF code: 20

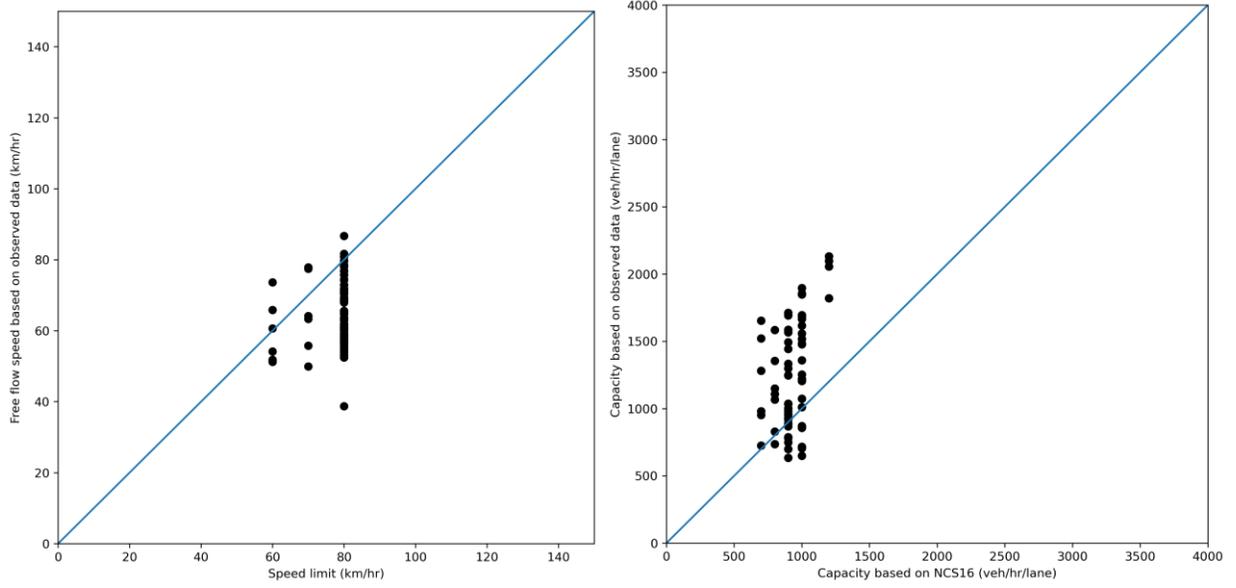


Figure 6.15-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 20

Rural Major Country Road (VDF code: 21)

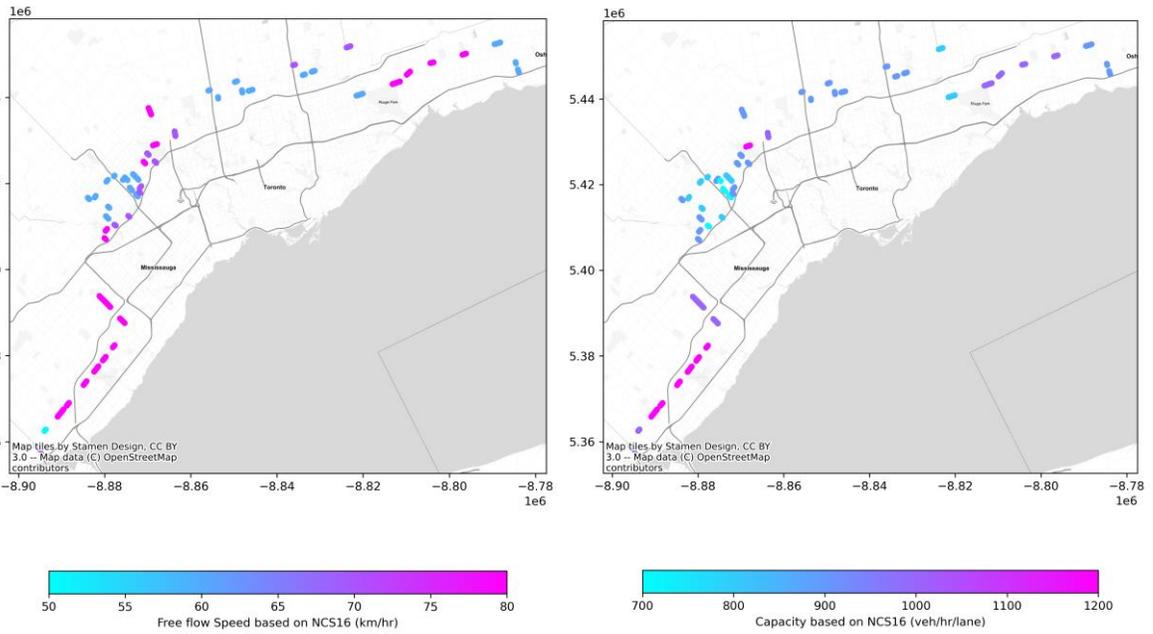


Figure 6.16-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 21

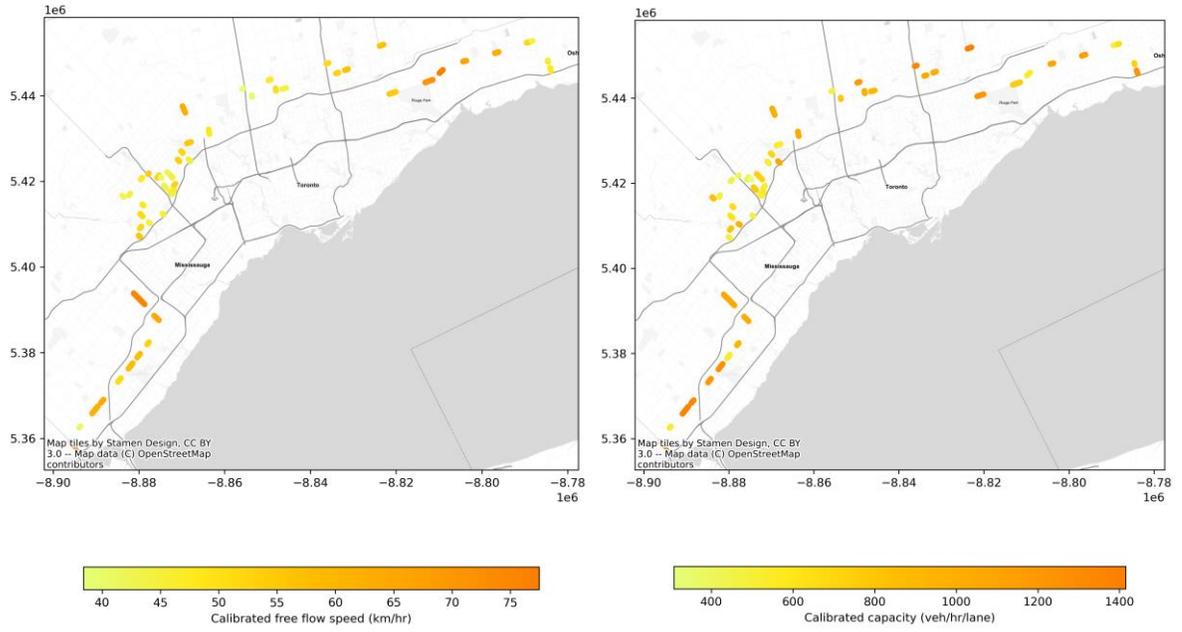


Figure 6.17-Calibrated capacity (right) and free flow speed (left), VDF code: 21

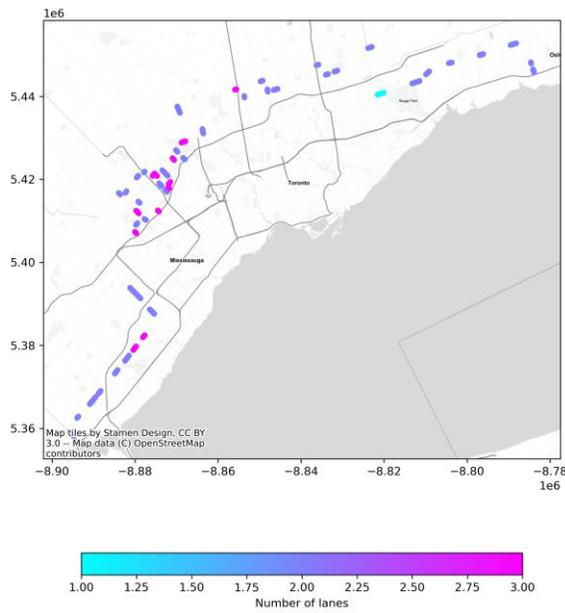


Figure 6.18-Number of lanes, VDF code: 21

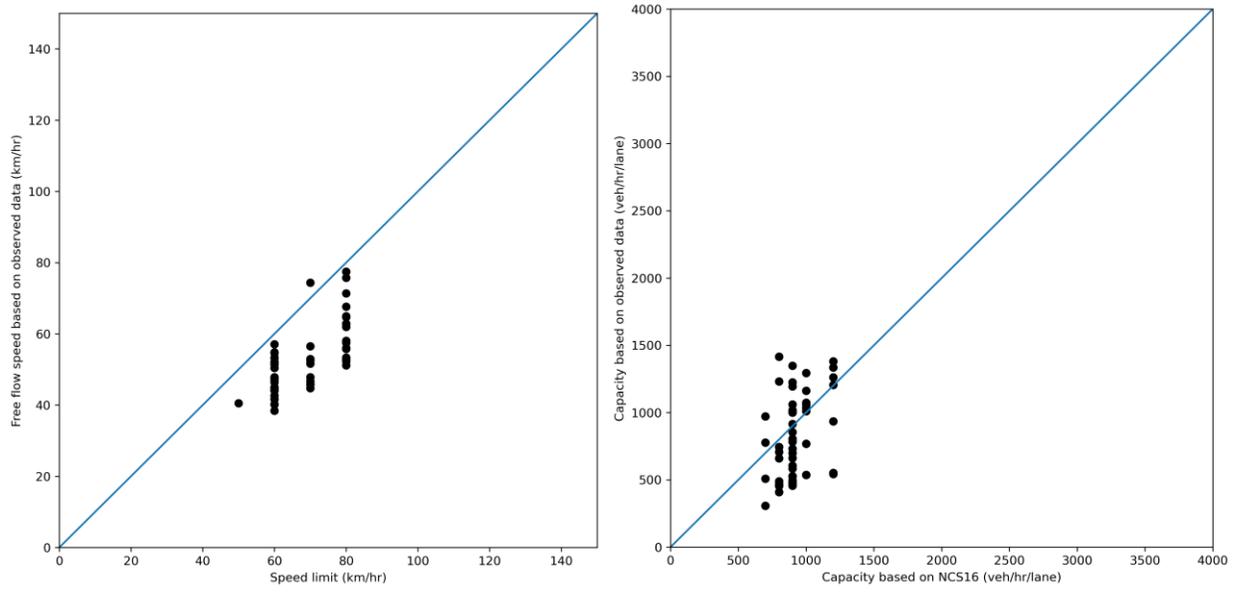


Figure 6.19-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 21

Rural Collector Road (VDF code: 22)

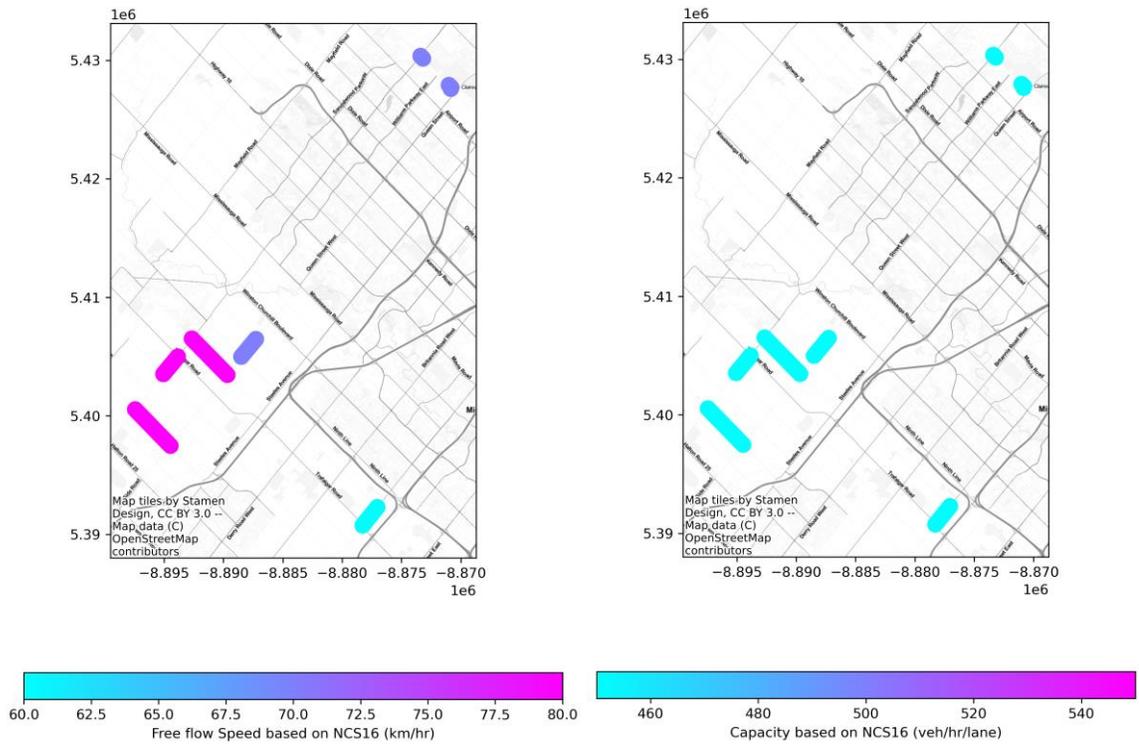


Figure 6.20-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 22

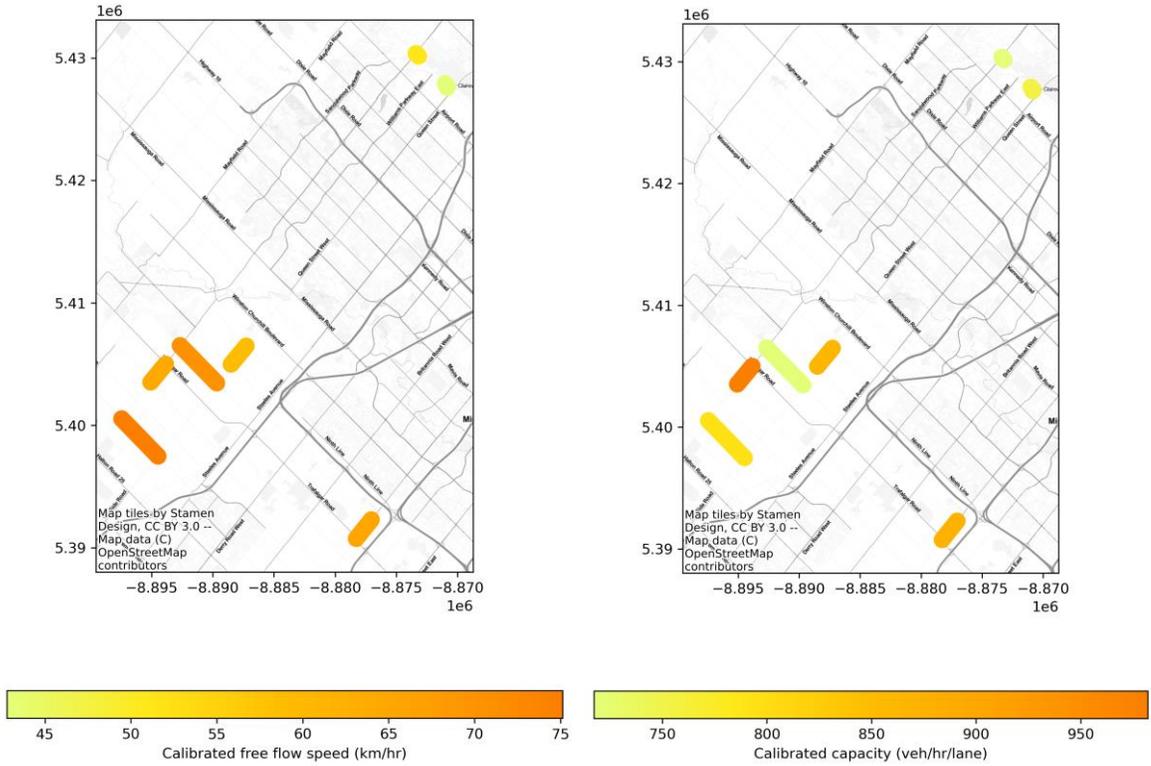


Figure 6.21-Calibrated capacity (right) and free flow speed (left), VDF code: 22

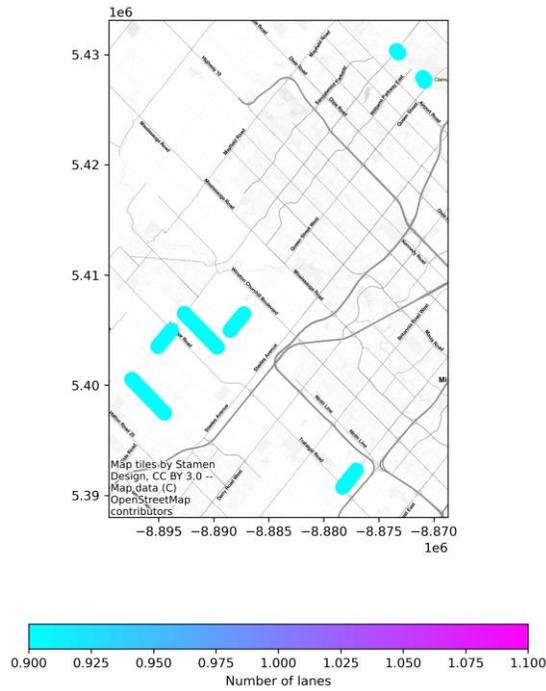


Figure 6.22-Number of lanes, VDF code: 22

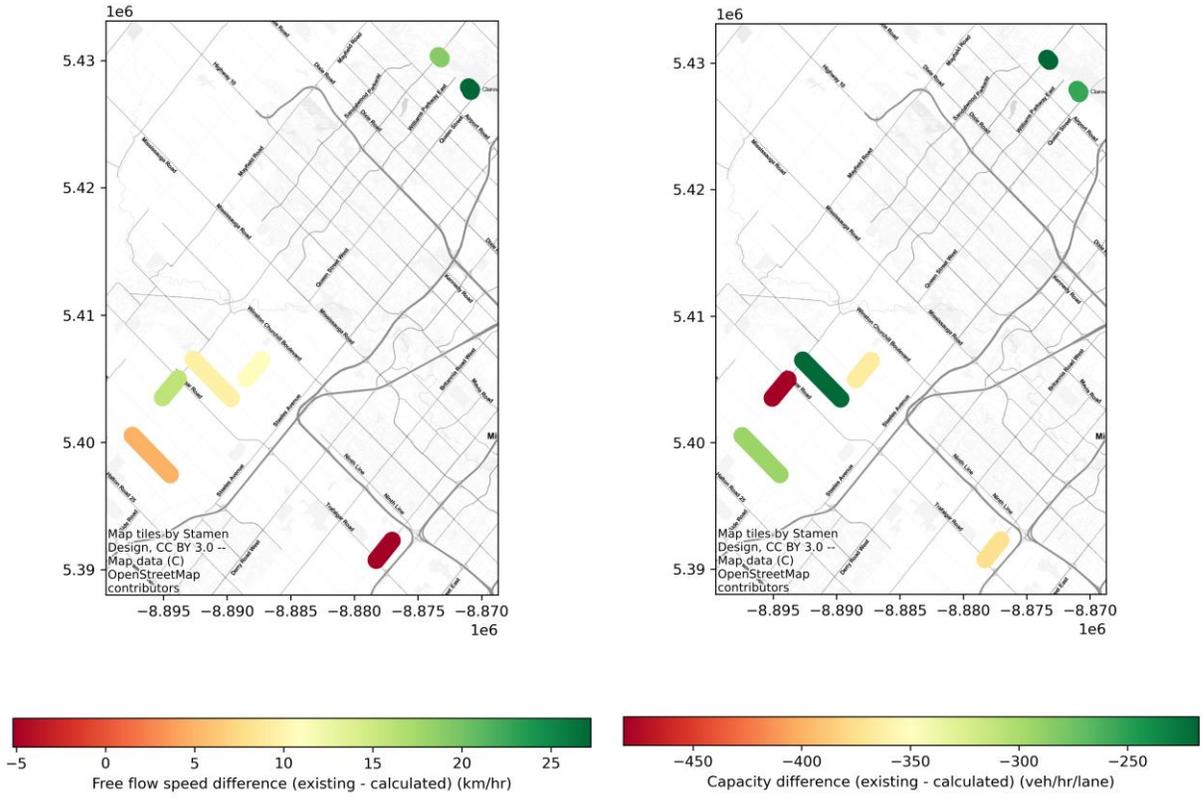


Figure 6.23-Existing and calculated values difference of free flow speed and capacity, VDF code: 22

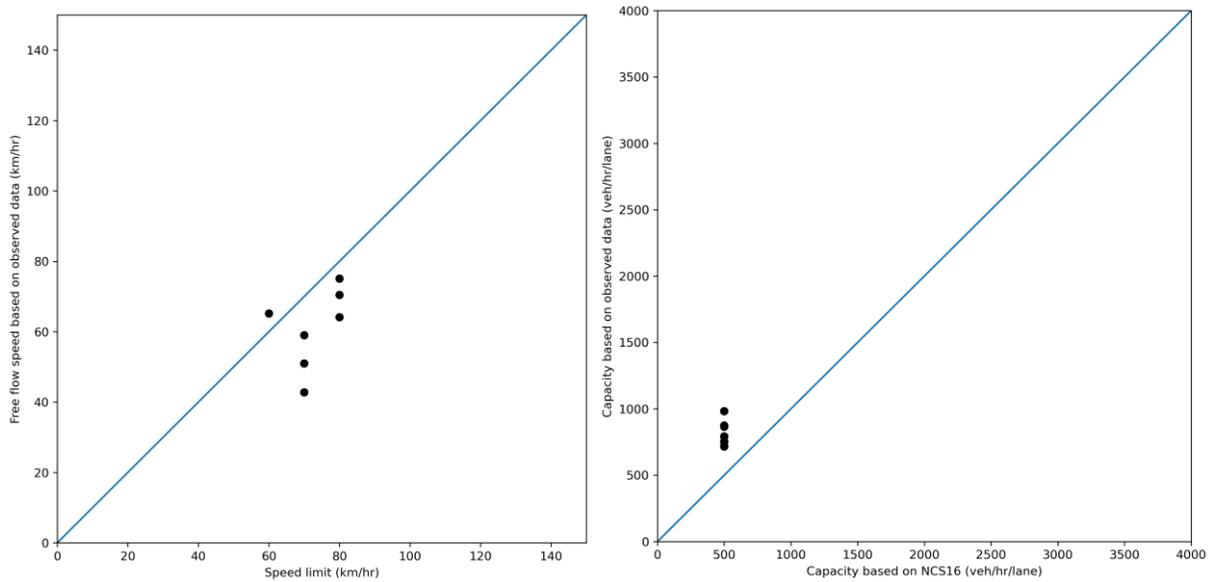


Figure 6.24-Comparison curve between existing and calibrated free flow speed (left) and capacity (right), VDF code: 22

Suburban Principal Urban Arterials (VDF code: 30)

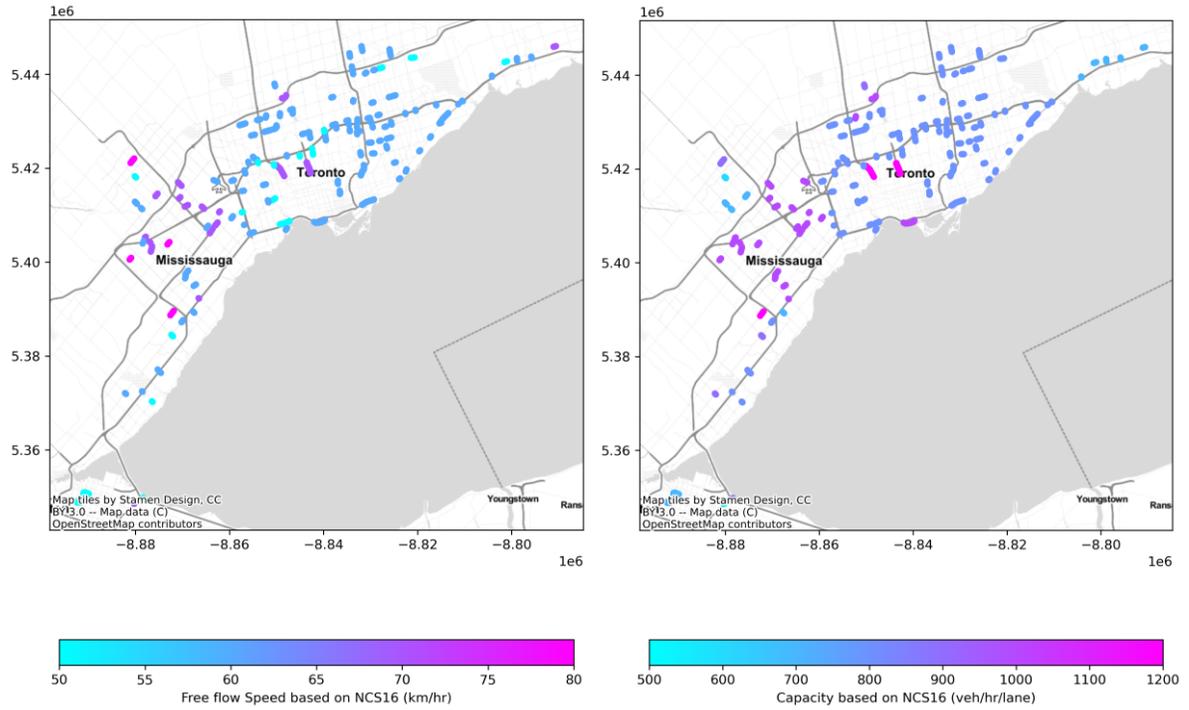


Figure 6.25-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 30

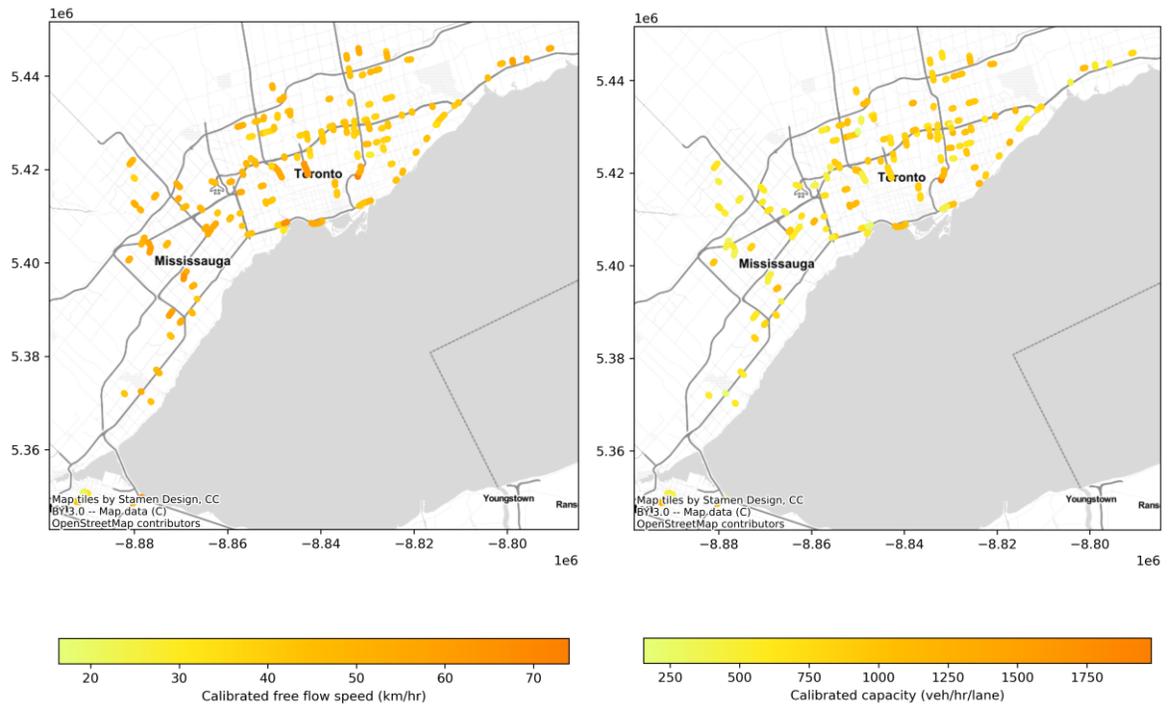


Figure 6.26-Calibrated capacity (right) and free flow speed (left), VDF code: 30

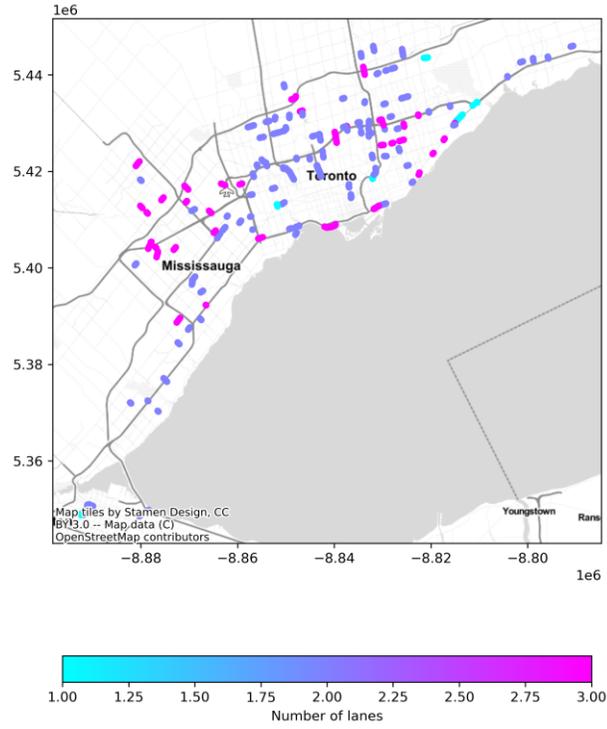


Figure 6.27-Number of lanes, VDF code: 30

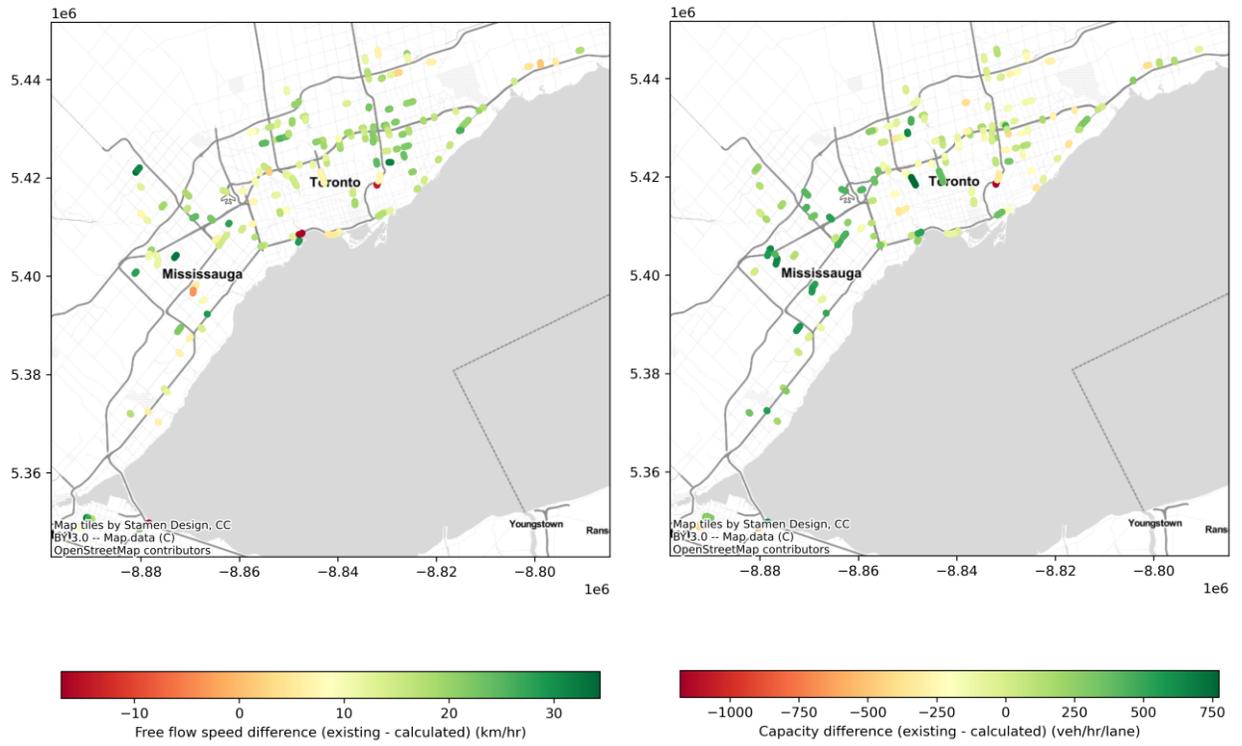


Figure 6.28-Existing and calculated values difference of free flow speed and capacity, VDF code: 30

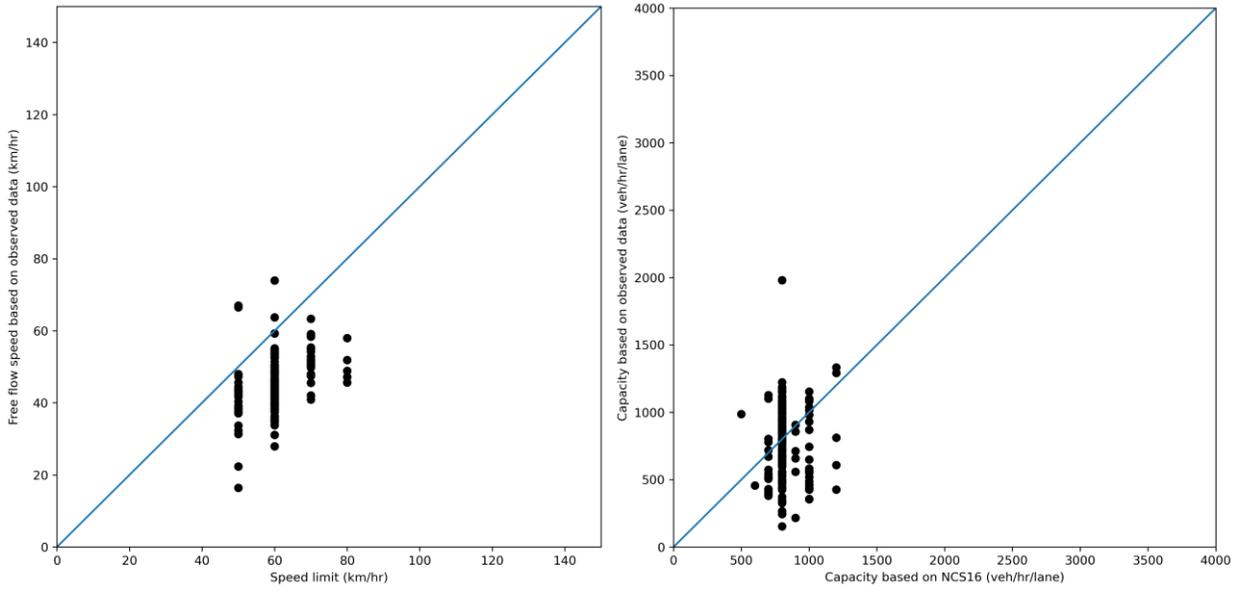


Figure 6.29-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 30

Suburban Collector Road (VDF code: 31)

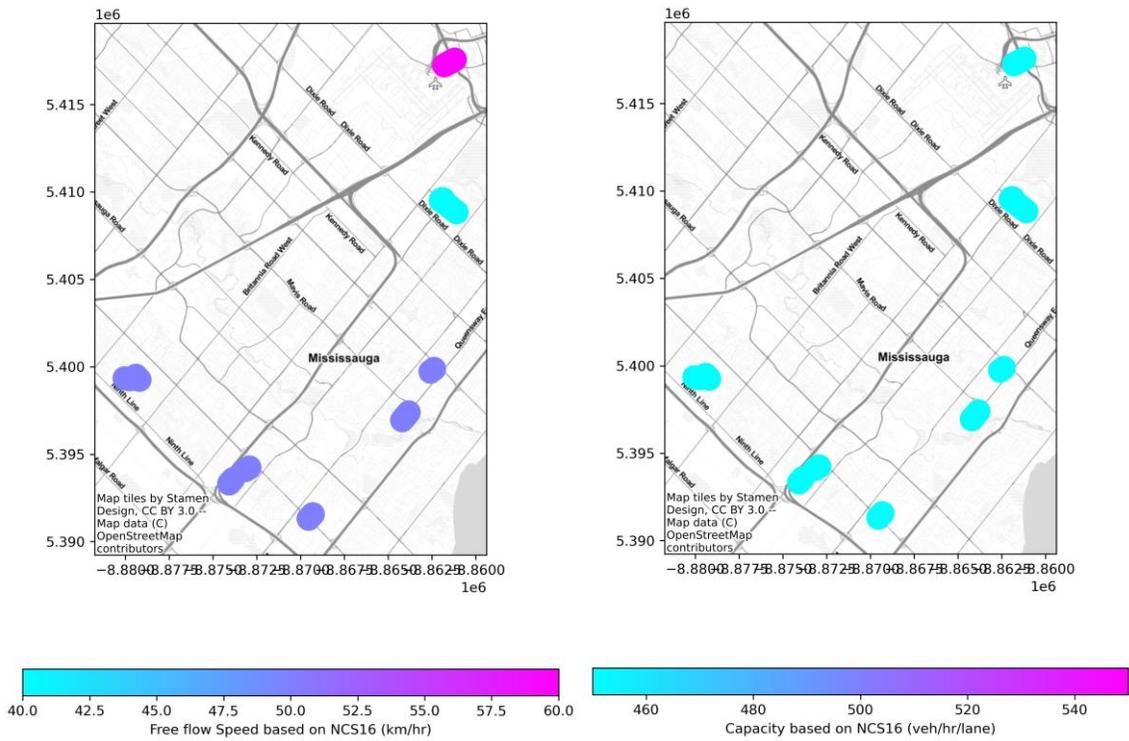


Figure 6.30-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 31

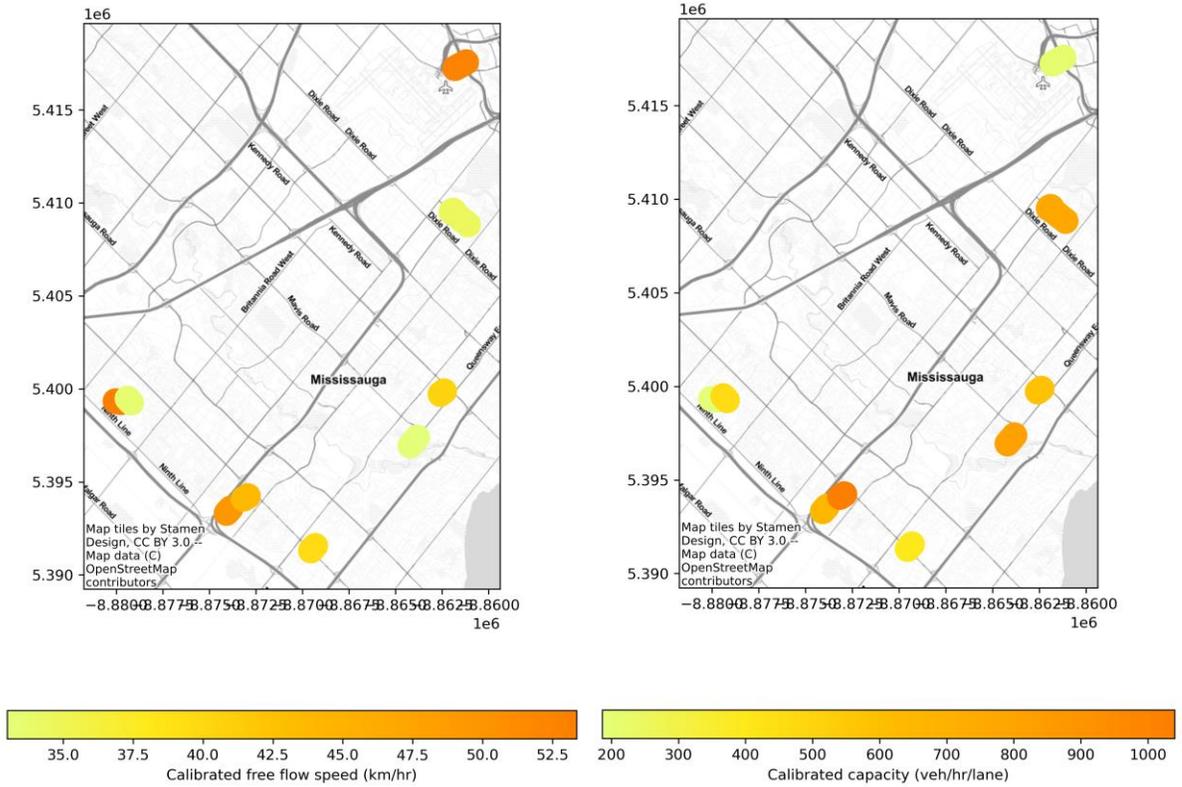


Figure 6.31-Calibrated capacity (right) and free flow speed (left), VDF code: 31

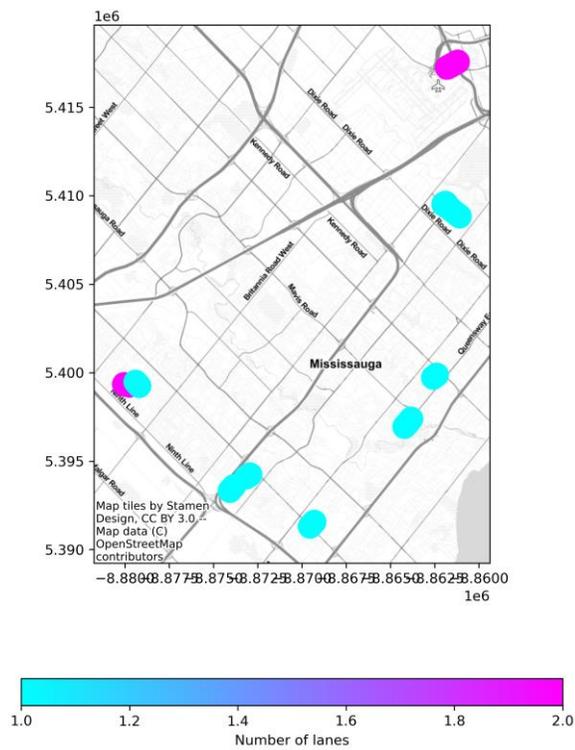


Figure 6.32-Number of lanes, VDF code: 31

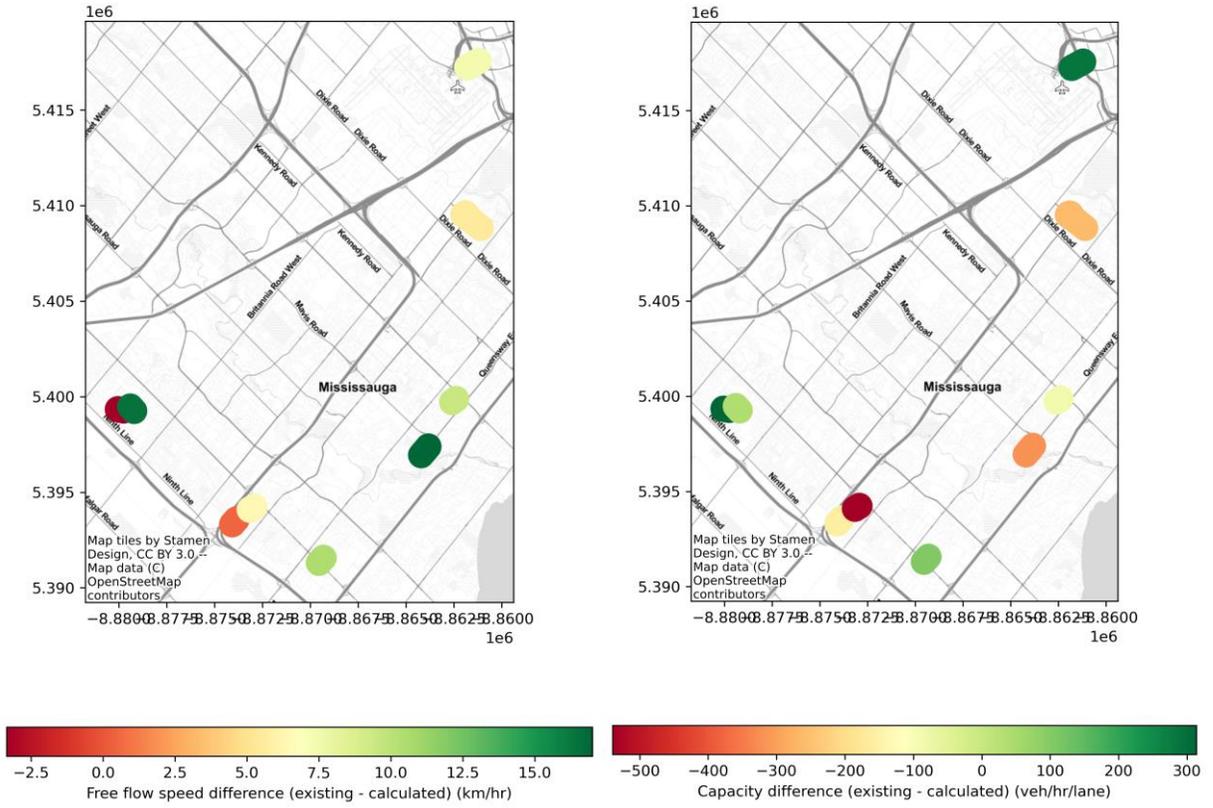


Figure 6.33-Existing and calculated values difference of free flow speed and capacity, VDF code: 31

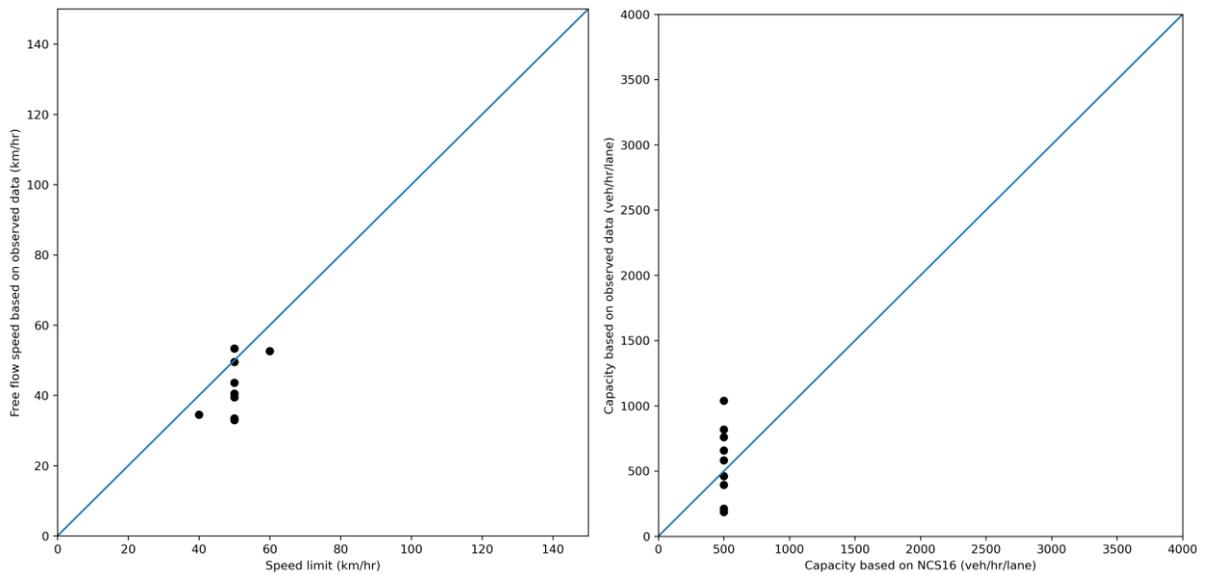


Figure 6.34-Comparison curve between existing and calibrated free flow speed (left) and capacity (right), VDF code: 31

Major Urban Arterials (VDF code: 40)

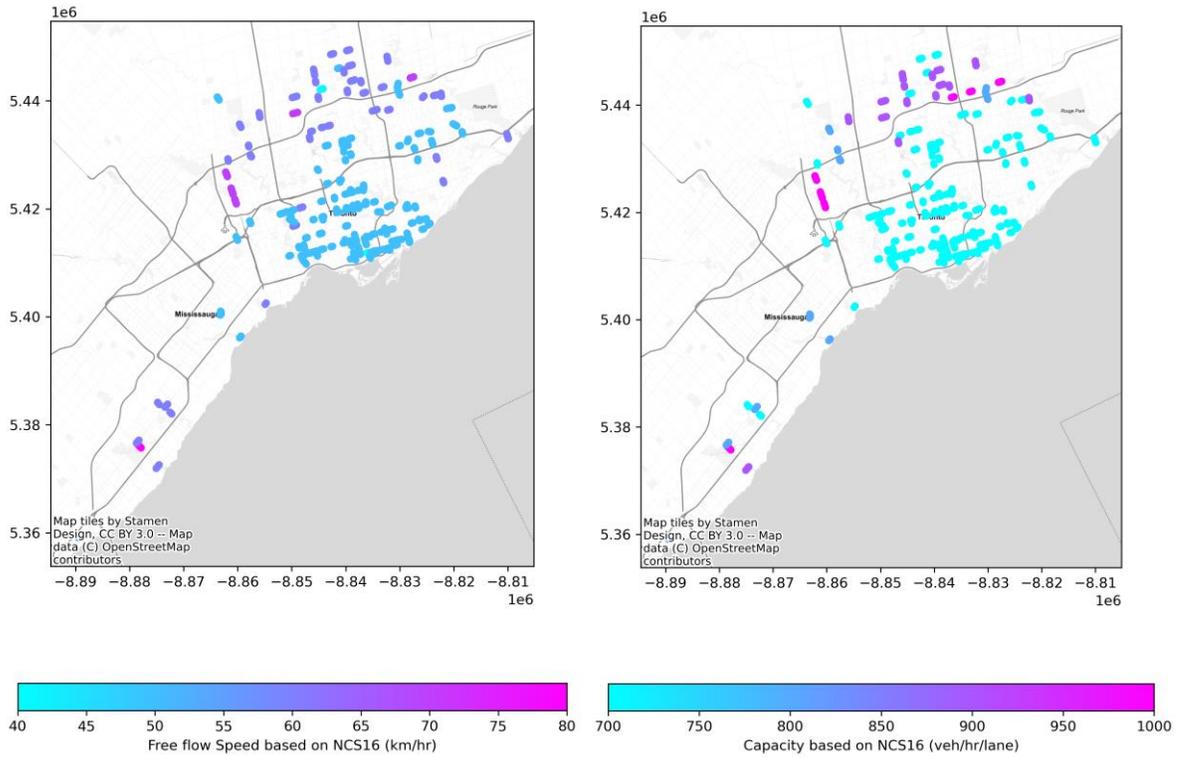


Figure 6.35-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 40

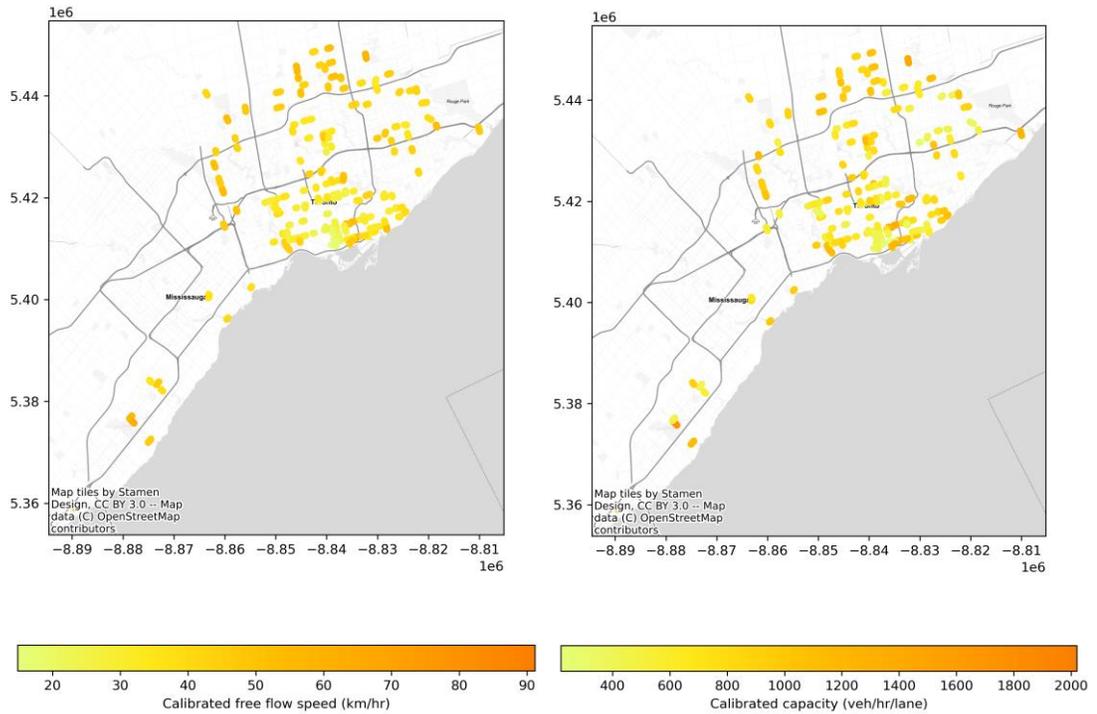


Figure 6.36-Calibrated capacity (right) and free flow speed (left), VDF code: 40

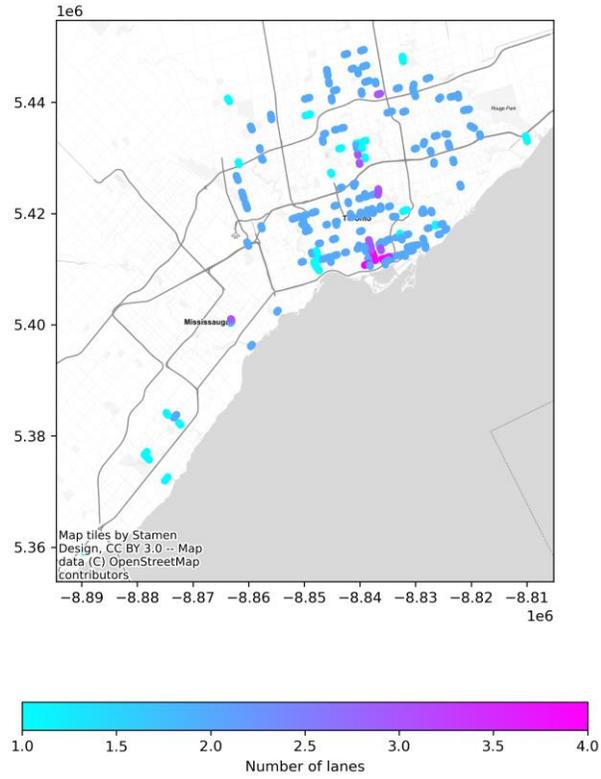


Figure 6.37-Number of lanes, VDF code: 40

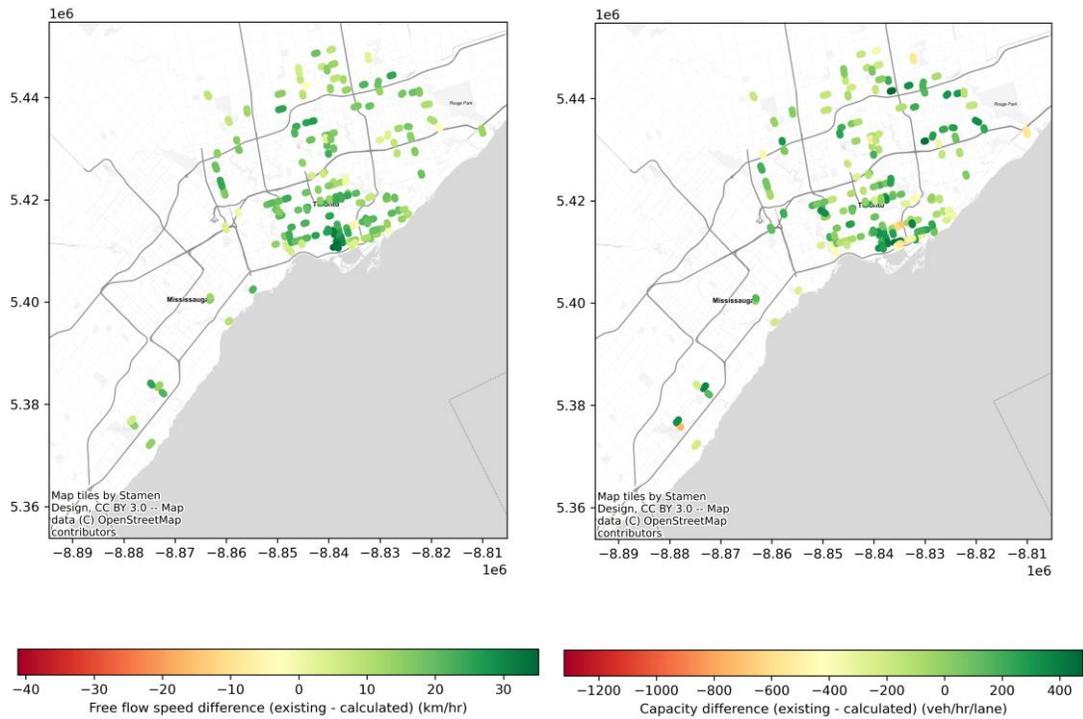


Figure 6.38-Existing and calculated values difference of free flow speed and capacity, VDF code: 40

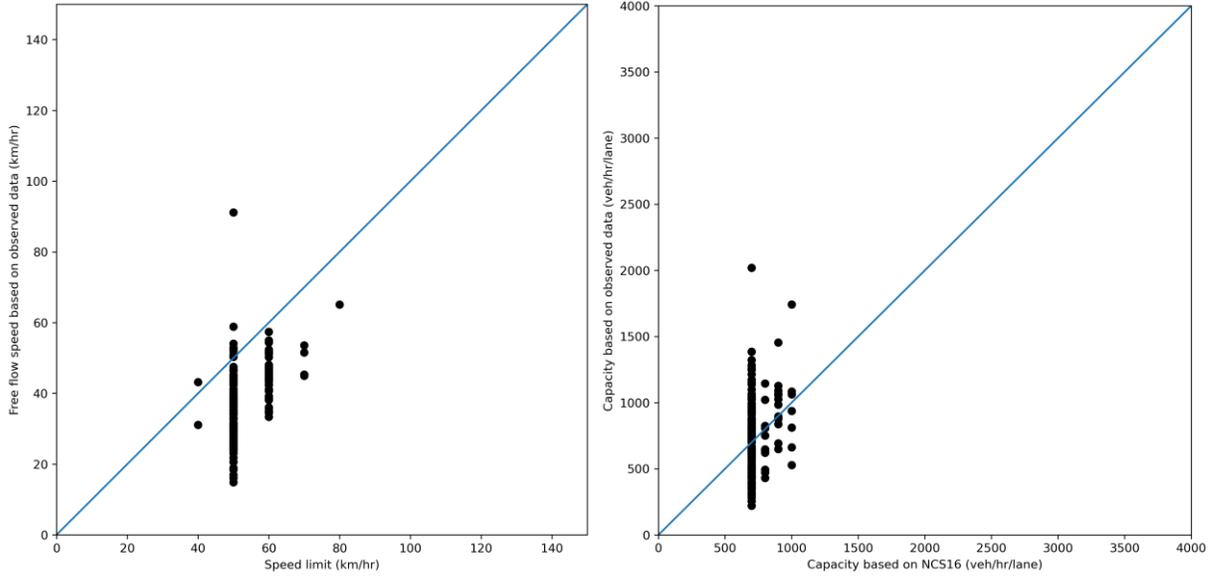


Figure 6.39-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 40

Minor Urban Arterials (VDF code: 42)

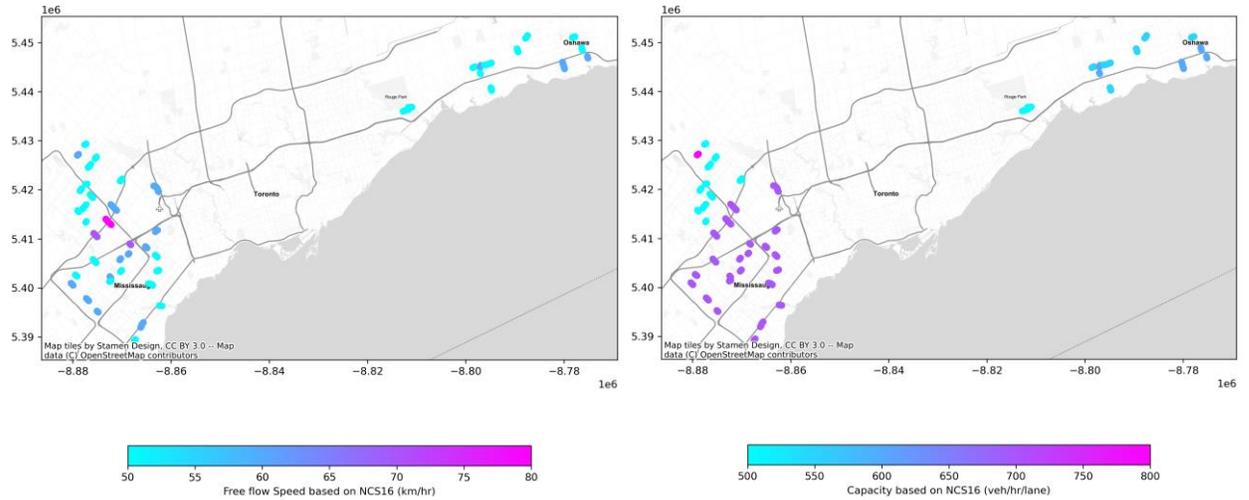


Figure 6.40-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 42

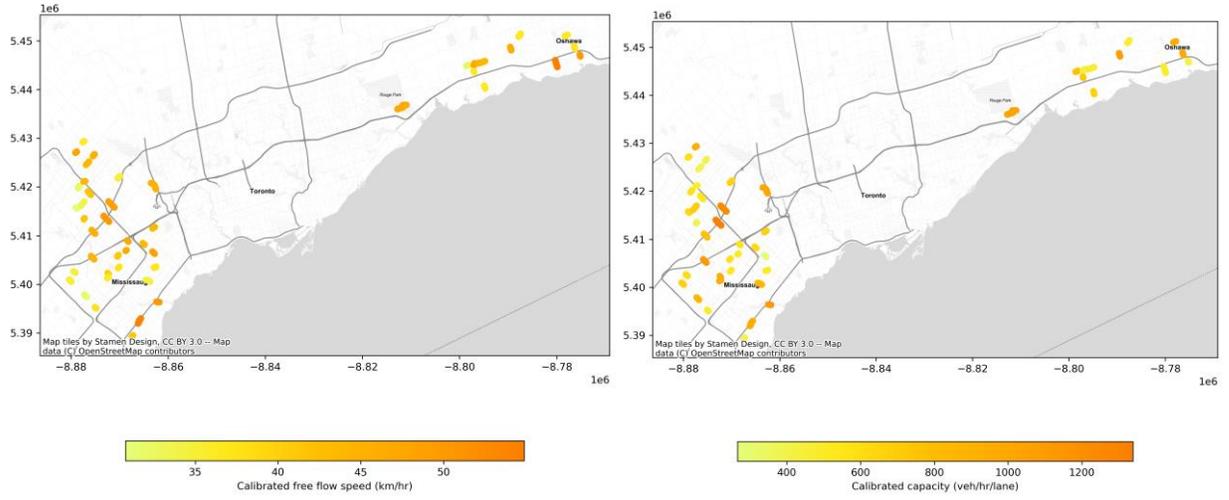


Figure 6.41-Calibrated capacity (right) and free flow speed (left), VDF code: 42

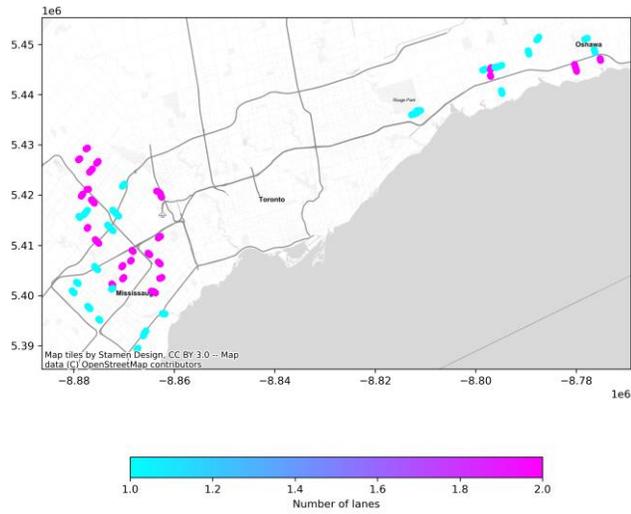


Figure 6.42-Number of lanes, VDF code: 42

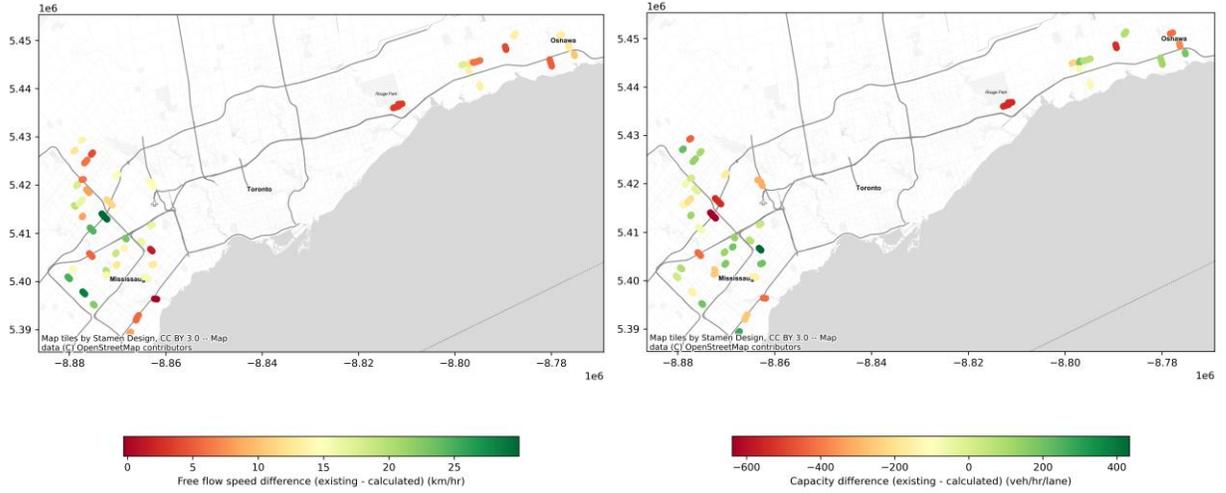


Figure 6.43-Existing and calculated values difference of free flow speed and capacity, VDF code: 42

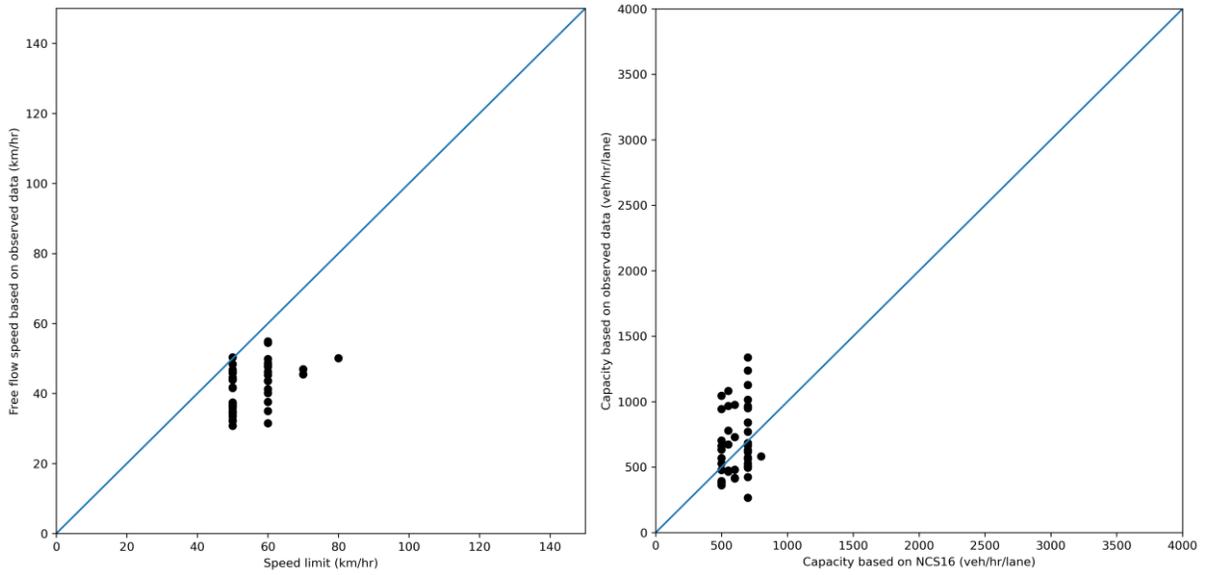


Figure 6.44-Comparison curve between existing and calibrated free flow speed (left) and capacity (right) , VDF code: 42

Downtown/City Center Roads (VDF code: 50)

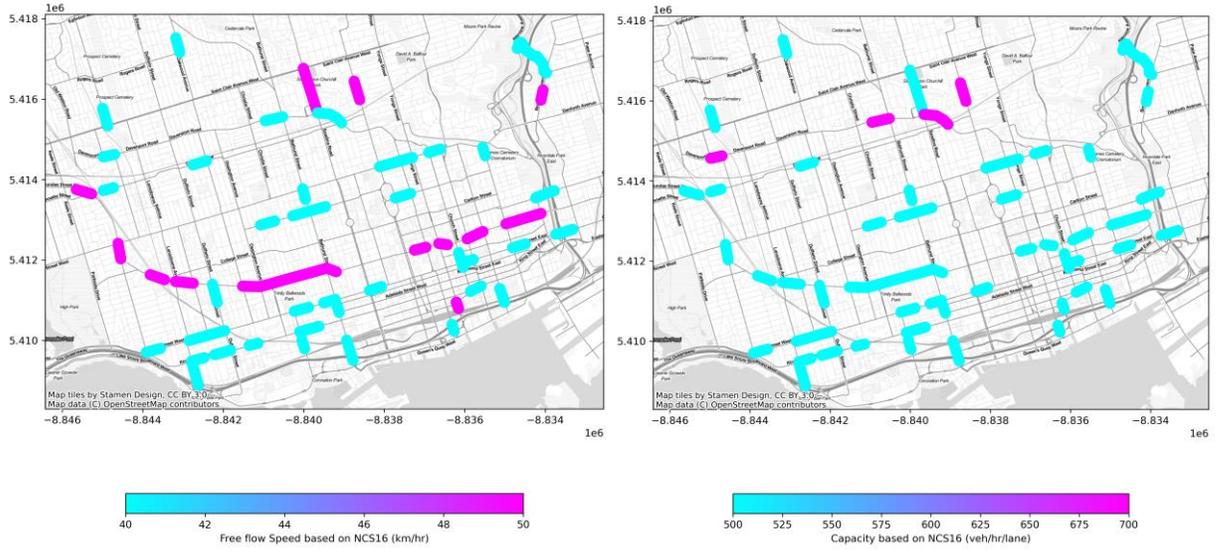


Figure 6.45-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 50

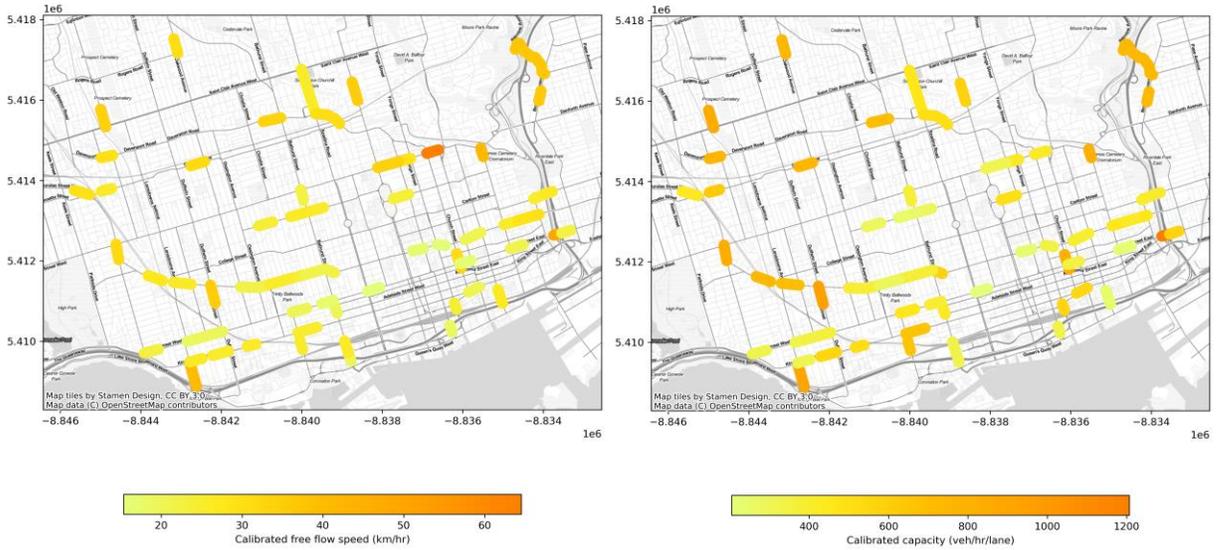


Figure 6.46-Calibrated capacity (right) and free flow speed (left), VDF code: 50

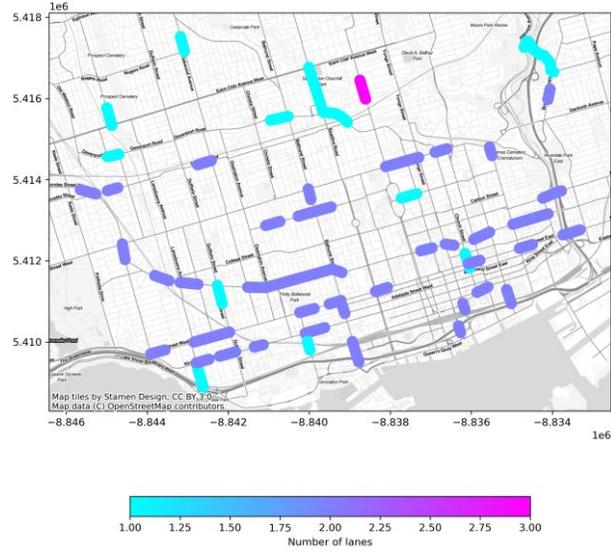


Figure 6.47-Number of lanes, VDF code: 50

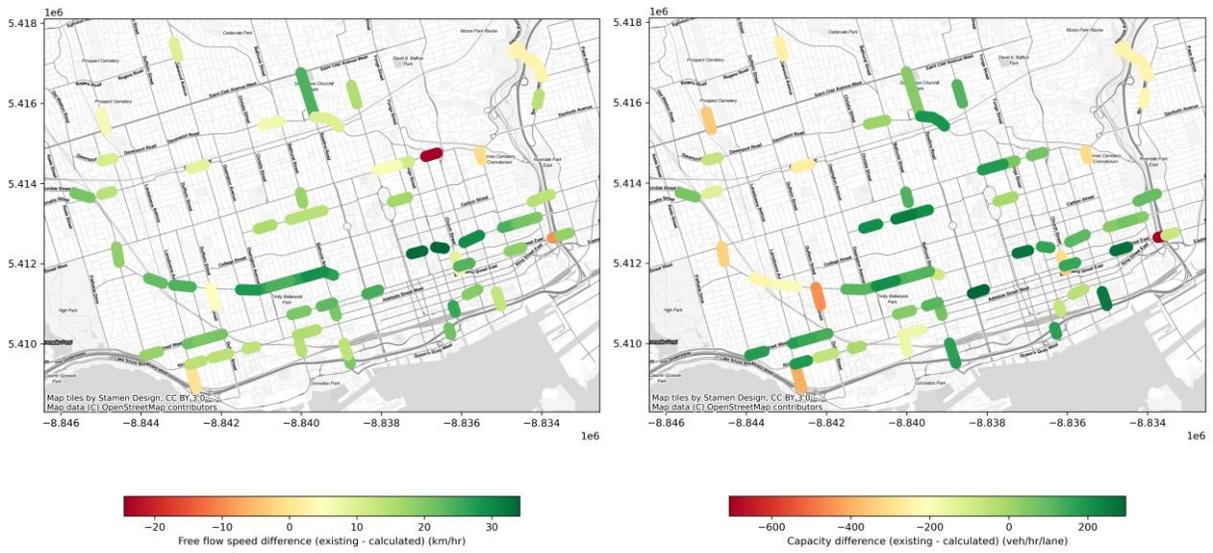


Figure 6.48-Existing and calculated values difference of free flow speed and capacity, VDF code: 50

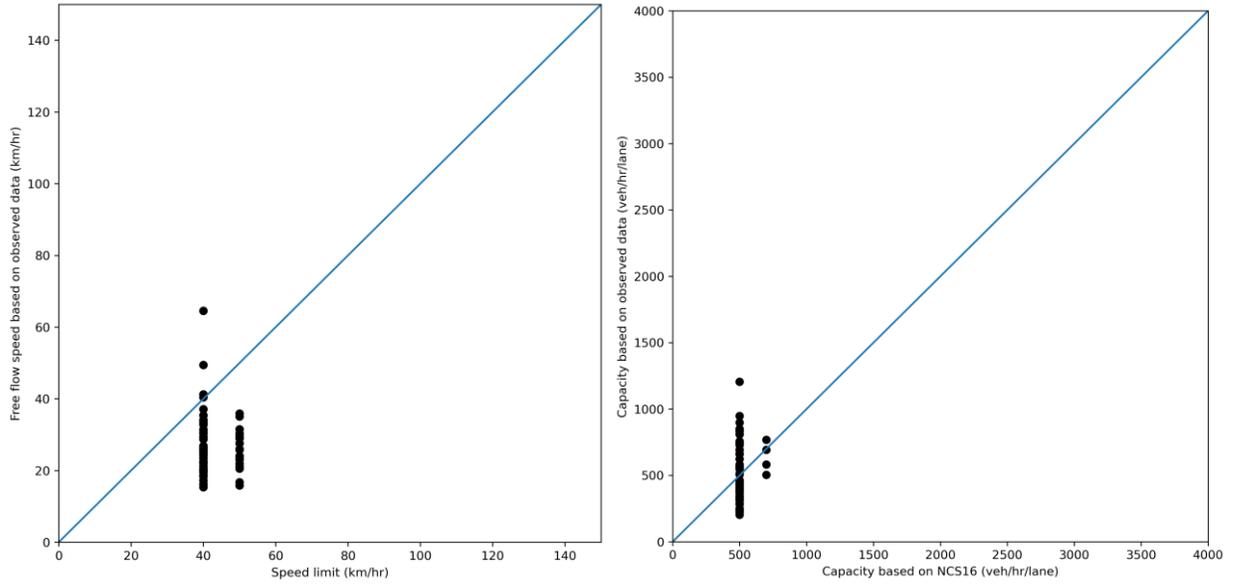


Figure 6.49-Comparison curve between existing and calibrated free flow speed (left) and capacity (right), VDF code: 50

Urban Collector Roads (VDF code: 51)

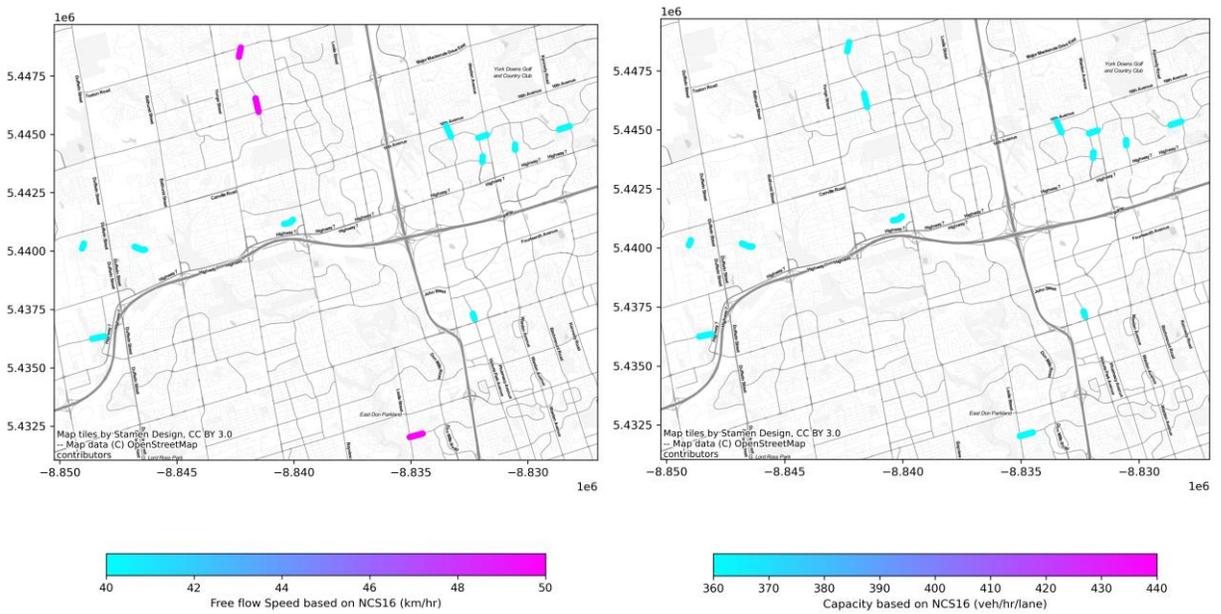


Figure 6.50-Capacity (right) and free flow speed (left) based on NCS16, VDF code: 51

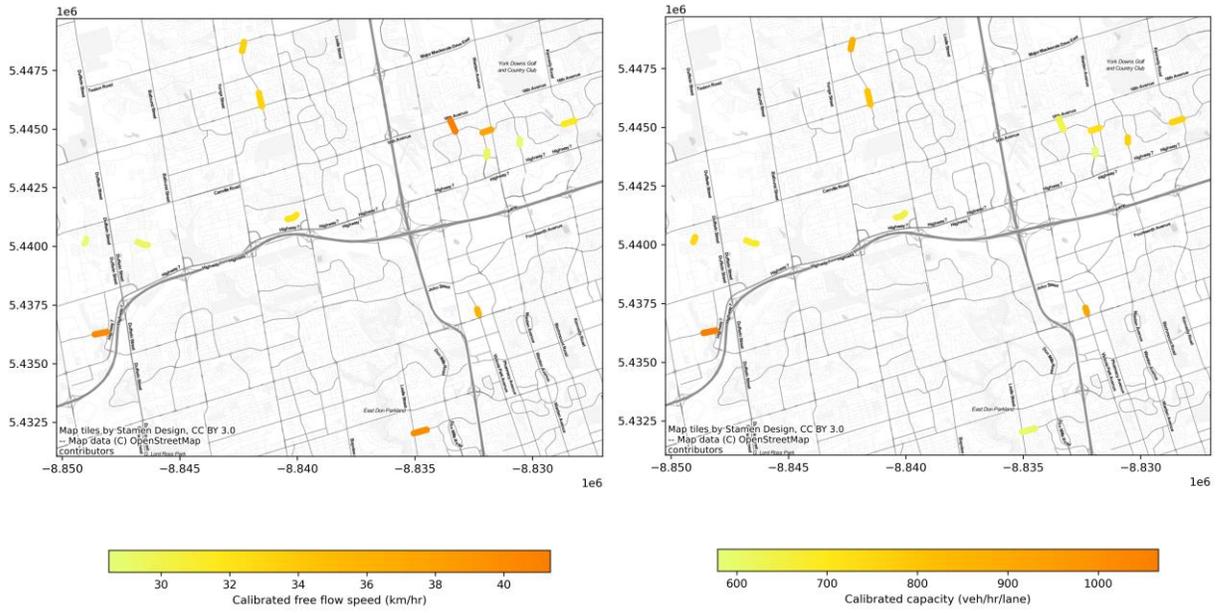


Figure 6.51-Calibrated capacity (right) and free flow speed (left), VDF code: 51

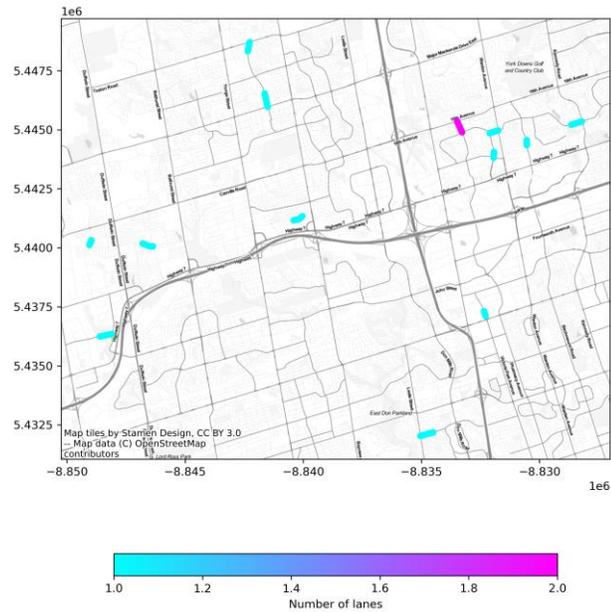


Figure 6.52-Number of lanes, VDF code: 51

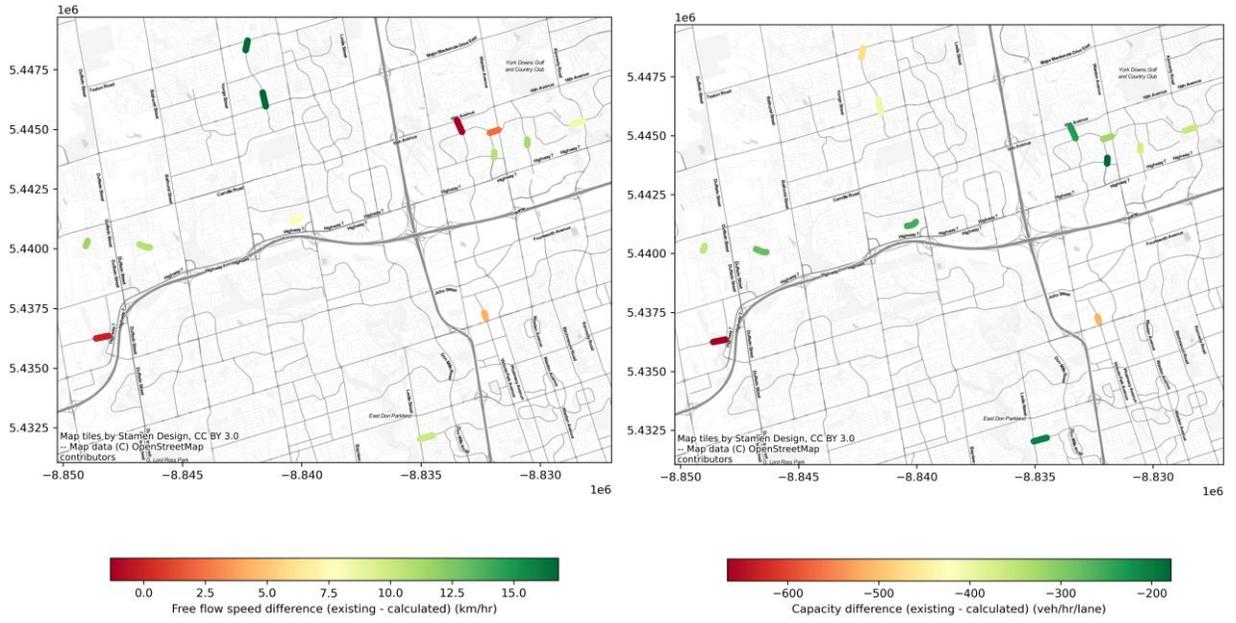


Figure 6.53-Existing and calculated values difference of free flow speed and capacity, VDF code: 51

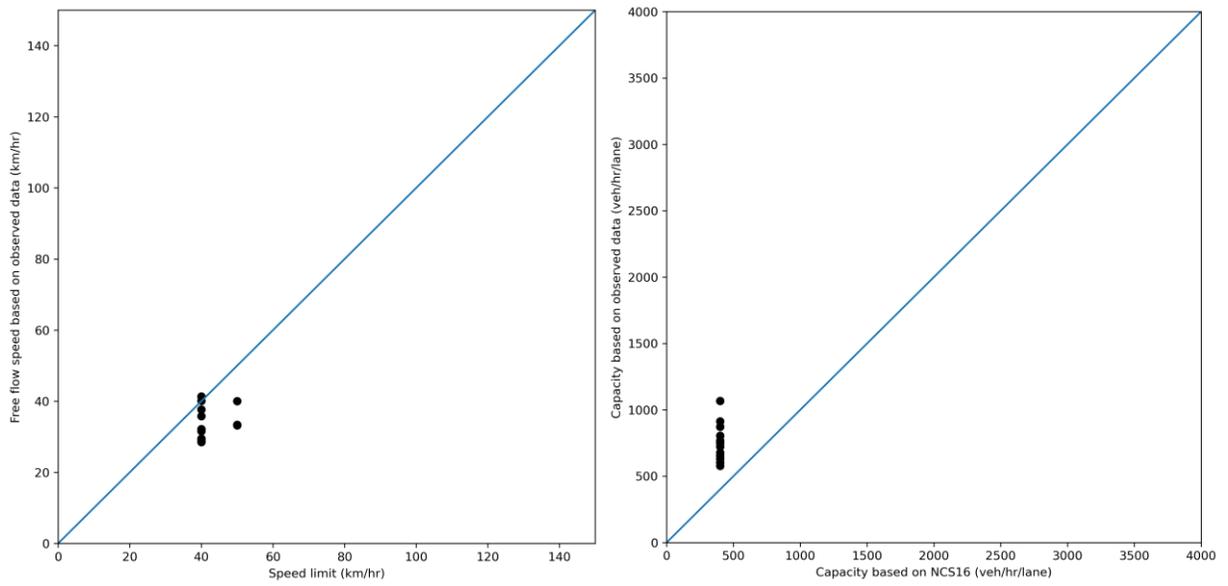


Figure 6.54-Comparison curve between existing and calibrated free flow speed (left) and capacity (right), VDF code: 51