Implementing Model Predictive Control based Variable Speed Limit on Urban Freeways: Data Imputation, Model Modification and Field Test Analysis

by

Yuwei Bie

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science
in
Transportation Engineering

Department of Civil and Environmental Engineering
University of Alberta

©Yuwei Bie, 2016

ABSTRACT

Among different freeway traffic control strategies, Variable Speed Limit (VSL) shows its excellence in terms of control scale, technical feasibility and the capability of improving driving environment and traffic throughput. The Model Predictive Control (MPC) based VSL method provides a close form control loop enabling optimized variable speed limit value. The MPC-VSL control system relies heavily on a stable real time data source, an accurate traffic state prediction model and timely feedback from field implementation. The Vehicle Detection Stations (VDS) system is responsible for providing real time traffic flow related data. Most of the time VDS system works well, however, there are occasions when one set of loop lost data due to hardware failure, and this thesis provides imputation algorithm for missing data. The macroscopic traffic state prediction model in MPC-VSL control scheme is the modified METANET model. The feasibility of modifying one critical term in the original METANET model, namely "desire speed", is tested in this thesis with different weather conditions using real field weather and loop detector data. The last part of thesis will be evolutional analysis of VSL field test that was conducted on Whitemud Drive, Edmonton from August 13 to September 4 of 2015, borrowing the concept of time domain analysis scheme and system robustness analytical tool.

ACKNOWLEDGEMENT

This thesis could not have been completed without the help and support of many people, only a few of whom are listed below.

I would like to thank my committee members Dr. Tony Qiu, Dr. Karim El-Basyouny and Dr. John Doucette, and I am especially grateful to my supervisor, Dr. Qiu, who has always encouraged me to do my best during my past two years at the University of Alberta. I benefit a lot from his ideas, advices and experiences.

Thanks to my team members in the Centre for Smart Transportation which has a wonderful research and collaboration atmosphere. I would like to thank Dr. Amy Kim, Dr. Xu Wang, Dr. Zhen Huang, Dr. Gang Liu, Dr. Qianqian Du and Cheng Lan. Their instruction provides significant support not only on the research ideas, but also on the research techniques. I would also like to thank Ying Luo, Xu Han, Lu Mao, Yahui Ke, Ling Shi and Can Zhang who have always been generous offering their help during my graduate studies.

Also thanks to the City of Edmonton Transportation group who provides loop detector data to support all my researches.

In addition, I wish to express appreciation to my family and friends. Only with their support and understanding can I get together courage to complete master degree and continue to be a researcher.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION ····· 1

1.1 Background ····· 1

1.2 Problem Statement and Research Motivation ····· 2

1.3 Research Objectives ····· 3

1.4 Structure of Thesis ····· 3

CHAPTER 2. LITERATURE REVIEW ····· 5

2.1 Review of Loop Detector data and Missing Data Treatment ····· 5

2.2 The Traffic Prediction Model in the System: METANET ····· 8

2.3 Variable Speed Limit with Model Predictive Control ····· 11

CHAPTER 3. MISSING TRAFFIC DATA IMPUTATION METHOD ····· 14

3.1 Introduction of VDS System ····· 14

3.2 Data Missing Detector Diagnostics Algorithm ····· 16

3.3 Missing Data Imputation Algorithm ····· 19

3.4 Imputation Approach ····· 22

3.4.1 Methods for Comparison ····· 23

3.4.2 Aggregation Level ····· 24

3.4.3 Imputation Experiment Results ····· 24

3.5 Summary of Loop Detector Data and Imputation ····· 32

CHAPTER 4. A CASE STUDY OF MODIFYING METANET MODEL ····· 35

4.1 Introduction and Background ····· 35

4.2 The Modified METANET Model with Consideration of Weather Condition ····· 36

4.2.1 Notations and assumptions ····· 37

4.2.2 The framework of METANET prediction model ····· 38

4.2.3 The weather specific fundamental diagrams ····· 41

4.2.4 Fundamental diagram and METANET calibration methods ····· 43

4.3 Modelling Weather Impacts on Fundamental Diagram ·····45

4.4 Case Study in Edmonton ·····48

4.5 Summary of Modifying Traffic Prediction Model ·····59

CHAPTER 5. THE IMPLEMENTATION AND EVALUATION OF MPC-VSL FIELD TEST 60

5.1 Introduction and Background·····60

5.1.1 Description of Variable Speed Limit Testbed ····· 61

5.2 Model Prediction Control Based VSL Algorithm ·····63

5.2.1 METANET Prediction Model and Modification····· 64

5.2.2 Object Function and Constraints ····· 67

5.3 Time Domain Analysis of Speed Control ·····68

5.4 Analysis of Measure of Effeciveness·····73

5.5 Summary of MPC-VSL Field Test ·····79

CHAPTER 6. CONCLUSION AND DISCUSSIONS ·····80

6.1 General Conclusions·····80

6.2 Limitation of this Thesis ·····82

6.3 Recommendations for Future Researches ·····84

LIST OF TABLES

Table 1 Volume Values Correlation between Lanes (VDS 1035).....21

Table 2 Coefficients and Statistics of MLR Model for Volume Measurement.....25

Table 3 Coefficients and Statistics of MLR Model for Density Measurement 25

Table 4 RMSE Value of Volume Imputation Using MLR, PLR and ASD Methods.....27

Table 5 RMSE Value of Density Imputation Using MLR, PLR And ASD Methods 27

Table 6 The key weather factors impacting FD features. 47

Table 7 Fundamental Diagram Features in Three Conditions 51

Table 8 Estimated Weather Factor Parameters And Statistics.....52

Table 9 Calibrated METANET Global Parameters 54

Table 10 Quantitative Traffic Prediction Performances: Speed 57

Table 11 Quantitative Traffic Prediction Performances: Density 57

Table 12 Quantities of Step Response of MPC-VSL Control System.....72

Table 13 Statistics of Unstructured Uncertainty of MOE versus VSL Overshoot 77

LIST OF FIGURES

Figure 1 Vehicle detection stations map of Whitemud Drive ····· 16

Figure 2 Volume data correlation among 3 lanes of VDS 1035 on Whitemud Drive, Edmonton ····· 20

Figure 3 Absolute errors of three methods for volume and density imputation ····· 28

Figure 4 Performance of volume data imputation for verification group using MLR, PLR and ASD methods ····· 30

Figure 5 Performance of density data imputation for verification group using MLR, PLR and ASD methods ····· 32

Figure 6 Illustration of triangular FD and its variation ····· 42

Figure 7 The significant weather factors impacting FD ····· 47

Figure 8 Ground snow and speed of wind under three weather conditions. ····· 49

Figure 9 The location of target VDS on Whitemud Drive, Edmonton. ····· 49

Figure 10 Field traffic data and calibrated FD under three weather conditions ····· 51

Figure 11 Speed prediction and density prediction accuracy on one heavy snow condition day (Nov.18) ····· 54

Figure 12 Speed prediction and density prediction accuracy on few snow condition day (Nov.14) ····· 55

Figure 13 Speed prediction and density prediction accuracy on one good weather condition day (May 02) ····· 56

Figure 14 VDS and DMS location on westbound testbed of Whitemud Drive ····· 62

Figure 15 DynaTAM interface and DMS on westbound testbed of Whitemud Drive ·· 63

Figure 16 Block Diagram of MPC-VSL control ····· 64

Figure 17 MPC-VSL Optimizer Decision Tree ····· 68

Figure 18 Speed Profiles of Valid Field Test Days near DMS 1 ····· 70

Figure 19 Block Diagram of TTT and TTD ····· 73

Figure 20 Comparing field measured TTT, model predicted TTT and adjusted model predicted TTT ····· 75

Figure 21 Comparing field measured TTD, model predicted TTD and adjusted model predicted TTD ····· 75

Figure 22 Boxplot of Unstructured Uncertainty of TTT/TTD over VSL Overshoot Value. ····· 77

LIST OF ABBREVIATIONS

VSL	Variable Speed Limit
VDS	Vehicle Detection Station
MPC	Model Predictive Control
EB	eastbound
WB	westbound
FD	Fundamental Diagram
TTT	Total Travel Time
TTD	Total Travel Distance
MOE	Measure of Effectiveness
FFS	Free Flow Speed
MLR	multiple linear regression
PLR	pairwise linear regression
ASD	“average of surrounding detectors” data imputation method

CHAPTER 1. INTRODUCTION

This chapter presents the background of model predictive control based variable speed limit including its data source, algorithm and implementation. In this part, the author will also describe the research motivation, research objectives and the structure of this thesis.

1.1 Background

Variable Speed Limit (VSL) is one of the main control methods in Active Traffic Management (ATM). Other control methods include ramp metering and route guidance. The design and implementation of VSL has close relationship with the development of Intelligent Transportation Systems (ITS). The new technologies support VSL to be efficient and reliable traffic control method for the following reasons: 1) Various sources of data enhance each other, such as data fusion of loop detector data, probe vehicle data, connected vehicle GPS data, etc. Those data technologies set human free and make data stream real time. 2) The development of computer software makes it easier to simulate various conditions before implementation. 3) The network of road facilities, vehicles and traffic management center enables the implementation of large scale VSL control.

The philosophy of VSL is that it adjusts the speed limit of certain segments of freeways on the near upstream of congestion prone locations. The decision of VSL values can be made based on past experience, feedback of current traffic states, and model predicted future traffic states. The prediction based VSL strategy is thought to have the following advantages over the former

two simple strategies: 1) The MPC-VSL includes an independent set of traffic flow model that is able to make relatively accurate short term traffic state prediction that is theoretically possible to prevent worse traffic condition from happening. 2) Based on prediction result, there is an optimized VSL decision that considers objective functions. 3) There is possibility that some components of system can be replaced such as prediction model or optimizer without changing the scheme of MPC-VSL control.

1.2 Problem Statement and Research Motivation

Since the MPC-VSL been proposed and simulated in VISSIM software, the reliability of this system has never been tested in real world. There are several vulnerable components in this system that can be bottleneck of the successfulness of VSL field implementation.

The first vulnerable point is the reliability of data source. Currently the system rely solely on duel loop detector data from field VDS and the loop detectors sometimes lost one lane data for more than one day, and the major reason is hardware failure. If human engineering cannot fix hardware problem timely, an imputation method is required. The second vulnerable point is the performance of prediction model in the system. Its accuracy significantly impact the decision result.

Since the 4-week VSL pilot field test was conducted in Edmonton, the problems listed above seem to be more urgent to be solved. The traffic data during test days, either field collected data or predicted data within the system is precious

for analysing and evaluation of this MPC-VSL system. The VSL field test is a good trigger for me to look into every detail of this control system, from loop detector database, prediction algorithm and field performance.

1.3 Research Objectives

This thesis looks into the problematic points of MPC-VSL system with three angles. There are three specific goals of this thesis:

1) Developing online loop detector data missing diagnose and imputation method for potential use.

2) Testing the prediction performance of METANET model with modification of one critical term. The case study is introducing weather factors into this term to improve METANET prediction accuracy under different weather conditions. This term in next step will be modified for MPC-VSL control.

3) Analysis of the performance of MPC-VSL field test, to be specific, doing time domain analysis for speed response in VSL control case and calculate time domain specifications. Also analysis robustness of prediction model as well as evaluating measure of effectiveness.

1.4 Structure of Thesis

This thesis includes 5 chapters:

Chapter 1 introduces the background of Model Predictive Control based variable speed limit method using loop detector data, problem statement and research objectives.

Chapter 2 is the literature review chapter, which reviews loop detector data and missing data imputation, METANET model formulation and weather specific modification, the came into being of MPC-VSL control method formulation and VSL simulation and implementation history.

Chapter 3 describes loop detector data collection and missing data imputation method.

Chapter 4 describes how original METANET prediction model is modified and take weather specific METANET model as case study.

Chapter 5 introduces a VSL field test conducted in Edmonton and analyzes its performance within system control scheme.

CHAPTER 2. LITERATURE REVIEW

This chapter reviews loop detector missing data imputation methods, METANET model formulation and weather specific modification methods, the MPC-VSL control method formulation and VSL simulation and implementation technology.

2.1 Review of Loop Detector data and Missing Data Treatment

Duel loop detector systems are widely used in transportation control practice for their ability to detect vehicle presence, so that they provide all necessary traffic variables in high frequency and accuracy. The more complex physical feature over single loop detectors indicates higher possibility of hardware failure and data missing. The American Association of State Highway & Transportation Officials (AASHTO) Guidelines for Traffic Data Programs do not recommend substituting estimated values for missing data points or sections, for the reason that errors will be relatively random and cannot be quantified. However, quickly developing intelligent transportation system control projects rely heavily on real-time full traffic ground truth data, so the missing data issue has become a major hurdle in applying sensor data to most traffic control programs. In response to this challenge, various imputation methods have been developed over the past few decades.

Earlier ad hoc methods were utilized to impute missing transportation data. This category of method usually uses replacement, average and weighted average techniques. Some commonly used ad hoc methods include the following: 1) The

historical average is used to represent a value for a given time of day or day of the week for imputing missing values [1]; 2) The weighted average of surrounding upstream and downstream stations is used to impute the missing value for the station, and the value is then divided among the lanes using historical lane distribution; 3) The average of the surrounding time periods involves averaging the values from the 10-minute intervals before and after the missing value, and can only be used when upstream and downstream time period data is available. Ad hoc methods have been effective in cases where the amount of missing data is small and road conditions are both consistent and recurrent.

After 1990th, using statistically principled techniques became a trend. Chen et al. [2] proposed the pairwise linear regression method, in which each missing value is imputed using estimated values for all neighboring loop detectors. This technique performs well and is commonly accepted. Al-Deek and Chandra [3]used pairwise second-order models with speed, volume and occupancy interaction terms and observed good results. Smith et al. [4] proposed a two-tiered approach, in which a simple ad hoc technique, such as the historical average approach, is used in daytime real-time imputation, while a computationally intensive but more advanced technique is employed to refine the imputed values during nighttime.

Other imputation approaches that have been developed employ a spatial-temporal relationship to impute a small amount of missing data. Qu et al. [5] utilized a probabilistic principal component analysis method. The study showed that the fluctuations of traffic flow were Gaussian and that the principal

component analysis can reveal the characteristics of traffic flow. Li et al. [6] indicate that the hidden spatial-temporal dependence is nonlinear and could be better retrieved by the kernel probabilistic principle component analysis based method. Asif et al. [7] developed a method that overcame the issue of incomplete historical data and addressed large and diverse road networks. The imputation method was based on fixed point continuation and canonical polyadic decomposition. The expectation maximization and data augmentation methods proposed by Smith et al. [4] for missing transportation data also served well in panel transportation data imputation when the missing rate was not high. These data imputation approaches have been shown to be both accurate and computationally efficient when the missing type is random and the missing proportion is small, but there were drastic increasing errors when the missing rate was higher and missing data was from one lane. Therefore, these imputation methods were not suitable for real-time imputation.

The abovementioned imputation methods impute one estimate for each missing value, and those techniques can be called single imputation methods. The multiple imputation method assumes that imputation should consider the uncertainty of the value to impute [8]. The multiple imputation technique can work in combination with various imputation methods. Kristian Henrickson et al. first implemented a proven predictive mean matching multiple imputation method and applied it to loop detector volume data collected on Interstate 5 in Washington State, using iterative multiple imputation with chained equations [9]. This add-on approach improves the accuracy of a complex statistical imputation

approach; however, it is still not suitable for real-time imputation due to the fact that it is time-consuming. So that for larger scale of data missing instead of random multiple data point missing, a fast and robust imputation method is needed such as the method proposed in Chapter 3, which is based on multiple linear regression method and the assumption of homogeneous lanes.

2.2 The Traffic Prediction Model in the System: METANET

Over the past two decades, the focus of efforts in modelling and forecasting macroscopic traffic states has transitioned from univariate temporal correlation to multivariate temporal-spatial correlation and from linear to nonlinear forms. Those models may be loosely classified as statistical and non-statistical methods. Some examples are included in the class of time series models, such as the seasonal autoregressive integrated moving average model [10] and Kalman Filter state-space model [11], neural network [12], nonparametric regression [13], stochastic Newell's three-detector method [14] and other empirical models.

Another class of works is based on the use of macroscopic traffic flow theory to estimate the internal traffic state for any intermediate point on a freeway or arterial segment from the boundary conditions. Macroscopic models consider traffic flow as fluid instead of individual vehicles. Three variables are capable of describing traffic stream characteristics: flow, density and mean speed. Macroscopic traffic flow models are classified as first, second or higher order, depending on the number of differential equations included [15]. Of all first-order models, the most used is the Lighthill–Whitham–Richards model [16][17], which uses one partial differential equation to describe the vehicle flow conservation

law. This model was also the first combination of a traffic flow model with a static fundamental diagram. Another representative first-order model is the Cell Transmission Model [18][19][20], which is a discretized and simplified version of the Lighthill–Whitham–Richards model. The Payne model [21] is the oldest second order traffic flow model. Besides the flow conservation law equation, the Payne model also includes one partial differential equation that describes mean speed dynamics. This model can replicate traffic phenomenon with higher accuracy. There are other types of second order traffic flow models such as variation kinematic waves [22], second-order traffic flow model with Kalman filter [23], CTM-based second-order traffic flow model with particle filtering [24], the Lighthill–Whitham–Richards partial differential equation with the Lagrangian measurements [25], Newell's simplified kinematic wave model [26][27]. The selected second-order traffic flow model in this paper is METANET [28][29], which is a discretized and enhanced version of the Lighthill–Whitham–Richards model combined with the Payne model. METANET model fulfill the simplicity and convenience requirement since it has a space-time discrete, explicit analytical state-space form and allow for convenient discretization intervals [15]. In addition, previous studies have demonstrated that the METANET model is highly accurate and relatively easy to calibrate, which makes it one of the most frequently utilized macroscopic traffic flow models in a variety of traffic engineering tasks and research. This model can also be used for optimal real-time traffic control of freeway traffic. In this paper, the macroscopic traffic state prediction will be based on the METANET model.

The modification of METANET can be of various purposes. In this thesis, the ultimate modification of METANET is to make it suitable for short time traffic prediction in variable speed limit control environment. A simpler trial of modifying the same term in equation is conducted in Chapter 4 which is a case study modifying METANET to make if capable of functioning well in varying weather conditions. Weather affects many fundamental aspects of road conditions. Maze provided evidence that traffic demand, safety, operations and flow can be reduced by rain, snow, fog, cold, and wind at different levels[30]. Some quantity analysis has been done by researchers in the past years. Methods based on aggregated flow and speed measurements from local sensors are used to estimate capacity and free flow speed at adverse weather[31]. HCM (2010) also recommended how much bad weather had impact on road conditions[32]. However, Kwon showed that in many cases the manual always underestimate or overestimate the real effects [33]. Hashim et al. showed the empirical analysis of the extent of highway capacity loss due to rainfall [34]. After that Hou et al. proposed that in mesoscopic network simulation weather factors can be introduced in calibration of traffic flow model for adverse weather [35]. In 2013, William et al. modelled the effects of rainfall intensity on traffic speed, flow and density relationships, and calibrate using hourly rainfall data from Hong Kong [36]. The papers indicated that adverse weather has different impact towards different locations and times and it is very necessary to identify the true impact for estimation. So that weather factors is introduced to METANET.

2.3 Variable Speed Limit with Model Predictive Control

Among all active traffic management strategies such as ramp metering, variable speed limit and route guidance, VSL control method changes posted speed limit based on real-time road, traffic, and weather conditions, and it can offer considerable promise in restoring the credibility of speed limits and improving safety and mobility by restricting speeds during adverse conditions. Thus, over time, two general views have evolved on the use of variable speed limits. The first emphasizes the homogenization effect [37], whereas the second is more focused on avoiding or mitigating traffic flow breakdown by reducing the input flow at bottlenecks by means of speed limits [38].

Among all active traffic demand management strategies such as ramp metering, variable speed limit and route guidance, VSL control method performs well in control scale, control method flexibility and feasibility. VSL method changes posted speed limit based on real-time road, traffic, and weather conditions [39], and it keeps the credibility of speed limits that under adverse conditions the speed limits are able to maintain traffic safety and the highest possible traffic throughput. There exist two basic thoughts about the effect of VSL. One emphasizes that the main role VSL plays is smooth traffic, the aim of control is not to reduce average speed, but to reduce speed difference [40], whereas the second is more focused on avoiding or mitigating traffic flow breakdown by reducing the input flow at bottlenecks by reducing speed limits [38].

Considering both two fundamental contributions of VSL, traffic engineers may achieve a more homogeneous traffic density over the freeway links and between lanes, at the same time preventing the high traffic density that leads to traffic breakdown. The more uniform speed distribution and density distribution reduces crash potential. The most commonly used VSL strategies works in a reactive manner with simple policies. To be specific, the VSL decision is triggered by abnormal traffic states that detected real time, such as a high traffic flow or low average speed. The VSL strategy generally includes two or three fixed values and chose by switching parameters. The safety benefits of implementing the VSL control have been well-acknowledged. To achieve more improvement in traffic throughput a more complex control algorithm that includes optimized VSL decision is needed. Proactive VSL control is one direction that goes further than previous reactive manner.

To predict the effects of a control measure several techniques can be used [41], such as case-based reasoning [42], knowledge-based systems in which policy is made upon previous practices [43], and model-based prediction [44]. In this paper we use for the predictions the macroscopic traffic flow model METANET first described in [28] and [29]. To find the optimal combination of control strategy that utilize traffic state prediction model we apply a model predictive control (MPC) framework[45][46][47] in which a close loop is formed that the future system state is predicted, optimal decision is made based on prediction result and then the optimal decision go back to implementation and influence the system state in the future step. MPC is an optimal control scheme

applied in a rolling horizon framework. Optimal control is successfully applied by Kotsialos et al. [48][49] to coordinate or integrate traffic control measures. Both optimal control and MPC have the advantage that the controller generates optimal control signal or decision according to a user-defined objective function. The objective function we use is a weighted summation of total travel time (TTT) and total travel distance (TTD), in which total travel time is minimized and total travel distance maximized simultaneously.

The simulation and field implementation of VSL had been done by many institutions, however, very few researchers conducted computer simulation of MPC based VSL and no field test is done using MPC-VSL control strategy. Hegyi et al. implemented a MPC-based VSL control, and the simulation in PARAMICS software resulted in a 32% reduction in TTT. [50] Hadiuzzaman et al. conducted MPC-based VSL simulation in VISSIM software that result in 38.8% reduction in TTT and 8.1% increase in TTD. [39] The field experiment of MPC-VSL in this paper first provides unique real data that reveals more problem than ideal computer simulation, and that the real world sensitivity and robustness of MPC-VSL control system is first checked in this paper.

CHAPTER 3. MISSING TRAFFIC DATA IMPUTATION METHOD

This chapter introduces the basic condition of VDS system and the need to introduce missing data imputation algorithm when some hardware are broken. The one month field data of one VDS is taken as case study. One lane of data is removed and imputation is tried. In the end the performance of imputation is evaluated by comparison of real data and imputed data.

3.1 Introduction of VDS System

The Vehicle Detection Stations (VDS) on a 10-kilometer corridor of Whitemud Drive in Edmonton, Canada, collects and stores traffic data from dual loop detectors. This section of road plays an important role in people and freight transportation in the city. It has two curves and several bottleneck locations where the number of lanes drops. The VDS system currently has 28 VDS in total, and each station has three or four dual loop stations. The data reporting frequency is 20 seconds. Each dual loop reports the volume q—the number of vehicles crossing the loop detector during a 20-second time interval—and space mean speed measurement v, as well as the occupancy measurement, which cannot be used directly in traffic control and cannot be transformed accurately into density. The accurate density measurement ρ is calculated by $\rho = q/v$.

The dual loop detector is currently the most prevalent and reliable data source for traffic operation related projects. These detectors provide engineers and researchers with mass traffic data every day, with a high update frequency and

accuracy that enables loop detector data to be input into a real-time traffic control module, such as variable speed limit. However, due to the fact that loop detectors are embedded into road pavement, hardware failure and errors constantly result in missing data. Nowadays, with the increase in data size and accuracy requirements for traffic control models, the treatment of missing data has become key to improve loop detector data quality.

Most of time the VDS system functions well, however, some of our loop stations lost data of one lane for some days. In this case, it is easy to tell when data is missing because, once a lane is detected absent of data at the beginning of a day, the imputation should carry on at least for the whole day.

The missing data pattern can be Missing At Random (MAR), which is caused by random events without consistent underlying hardware issues; the missing data holes or physically unreasonable values can be safely removed or imputed using prior distribution-related imputation methods. In this case, the loop detectors on Whitemud Drive do not experience MAR much; instead, the main problem is Missing Not At Random (MNAR). The loop detector stations on our field test road usually contain three or four dual loops, with each one representing one lane. If one detector fails to record data, the consequence is that data for the specific lane will be missed for a long period of time, e.g. a day or a week. Therefore, this field study examines missing lane data as a type. It is also regarded as a major missing type in other jurisdictions. The Figure 3 shows the VDS system map on Whitemud Drive, Edmonton.

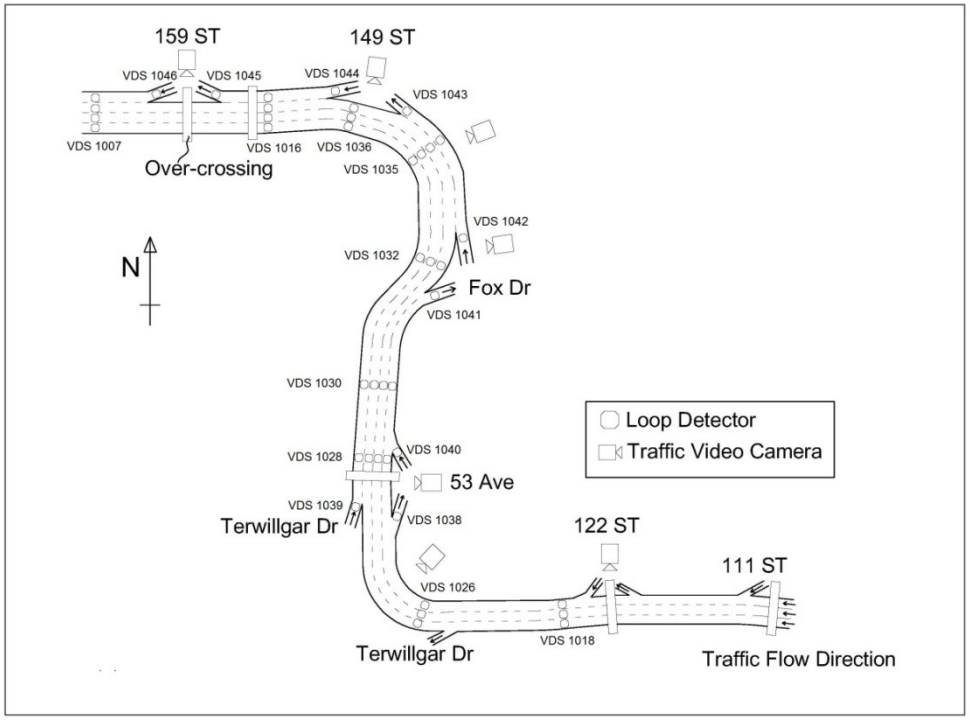


Figure 1 Vehicle detection stations map of Whitemud Drive

3.2 Data Missing Detector Diagnostics Algorithm

Existing data reliability tests include the threshold method [21] and "acceptable region" method [51] . The diagnostic algorithm in this paper is built upon those methods. The threshold method places thresholds on minimum and maximum flow, speed and density values, and is reported invalid if the detected value falls beyond the feasible region. The "acceptable region" method is similar. A region in the k-q plane is defined and data samples are declared acceptable if they fall into the defined region; this method was founded by researchers at the University of Washington and is referred to as the Washington Algorithm. [51] These above methods define the boundaries of the feasible region based on historical data. In

one specific case, the boundaries were fixed and had to be calibrated before online diagnosis.

The diagnosis algorithm in this research was based on corresponding lanes at the upstream stations. Previous methods tended to report $(q = 0, \rho = 0)$ points in the fundamental diagram as bad data points and eliminate them or impute with a non-zero value, which brought about significant positive bias to loop detector data. In the field, good detectors frequently report $(q = 0, \rho = 0)$ due to the high updating frequency; for example, during a 20-second time interval, there is a big possibility that no vehicle has passed. If $(q = 0, \rho = 0)$ are accepted as good data points, then the missing lane phenomenon would fail to be diagnosed. The speed measurement was not used as one of the diagnosis criteria, as it has the same result as using flow and density measurements.

The diagnosis method we proposed is based on corresponding upstream. For a specific lane, if the upstream lane had non-zero flow while itself had zero flow, then imputation was activated for that day. At the beginning of each day, the first five minutes were used for diagnosis, and imputation started after that. This diagnosis is feasible when the VDS positions are known. This algorithm is called the Daily Diagnosis Algorithm (DDA). The input to the algorithm is the 20-second frequency data recorded by the VDS of neighboring locations:

$$\Delta_j^m(d) = \begin{cases} 1, \text{ if } \sum_{i=1}^{15} q_j^m(i,d) = 0 \\ \qquad \text{and} \\ \qquad \sum_{i=1}^{15} q_j^{m-1}(i,d) \neq 0 \\ 0, \text{ otherwise} \end{cases} \tag{3-1}$$

where, m is the index of our target VDS, and we assume that VDS No. m-1 is the station upstream of station No. m. The 20-second flow of lane j is $q_j^m(i,d)$; i is the index of the 20-second sample number and i =1,2,3,…,4320; d is the index of the day of one month and d =1,2,3,…,30. The output of the algorithm is the diagnosing index $\Delta_j^m(d)$ for the day d. $\Delta_j^m(d) = 1$ if the loop is bad, and $\Delta_j^m(d) = 0$ if the loop is good. The DDA provides only one diagnosis result for one lane each day. The result of the diagnosis from (1) decides whether the specific lane should be imputed or not:

$$\begin{cases} \begin{cases} q_{true,j}^m(i,d) = q_{meas,j}^m(i,d) \\ \rho_{true,j}^m(i,d) = \rho_{meas,j}^m(i,d) \end{cases}, \text{if } \Delta_j^m(d) == 0 \\ \begin{cases} q_{true,j}^m(i,d) = q_{esti,j}^m(i,d) + \varepsilon_{q,j}^m(i,d) \\ \rho_{true,j}^m(i,d) = \rho_{esti,j}^m(i,d) + \varepsilon_{\rho,j}^m(i,d) \end{cases}, \text{if } \Delta_j^m(d) == 1 \end{cases} \tag{3-2}$$

where, $q_{meas,j}^m(i,d)$ and $\rho_{meas,j}^m(i,d)$ are measured values of flow and density. When measurable values from the field exist and $\Delta_j^m(d) = 0$, the imputation program is not triggered and measured values are treated as true values. $q_{esti,j}^m(i,d)$ and $\rho_{esti,j}^m(i,d)$ are estimated values yielded when $\Delta_j^m(d) = 1$ and the imputation program has started. In the process of imputation, $\varepsilon_{q,j}^m(i,d)$ and $\varepsilon_{\rho,j}^m(i,d)$ are error values that should be independent of estimated values, and be minimized. If imputation is triggered, only the flow measurement q and density measurement ρ are imputed; speed would be calculated via $v = q/\rho$. The reason why q and ρ are chosen as imputation variables, while speed is imputed indirectly, is that flow and density are found to be highly correlated among nearby lanes,

while the speed measurement shows no similar trend. The following section explains the imputation algorithm design in detail.

3.3 Missing Data Imputation Algorithm

The multiple linear regression (MLR) model is proposed to describe the behavior of loop detectors in the same station using historical data. It is well accepted that volume and density values are highly correlated between adjacent loops at the same station (1, 3, 4, 6), while the speed value has not shown such a correlation. Since speed values have random characteristics and are difficult to impute directly, we have chosen to impute flow and density first, then calculate speed via equation $= q/\rho$. In previous studies using the pairwise linear regression method (3), it was found that neighbor detector pairs have a higher correlation than non-neighbor loops at the same station, which was not true in our data. We found that all loops at the same station are highly correlated and share transportation characteristics. Taking all other loops into consideration makes the most use of historical information. The reason why a higher-order model was not considered is that different lanes in the same road segment are intuitively homogenous. Another advantage of MLR over pairwise linear regression is that there is no need to choose which loop to define as "neighbor" if one loop has two adjacent loops, and this saves computing time and reduces complexity. Finally, MLR is more stable in performance than pairwise regression because sometimes one loop detector has a high correlation with its left hand neighbor, while at other times it has a higher correlation with the right hand one. By taking all other loops in the same location

into consideration, the model should be unique and performance should be more robust.

To exhibit the linear relationship among loop detectors at the same station, take VDS 1035 in Whitemud Drive as an example. Figure 2 shows three pairings of all three loop detectors in VDS 1035 as well as a 3D scatter figure of all three lanes. Table 1 shows that all lanes in the same station have high correlation.

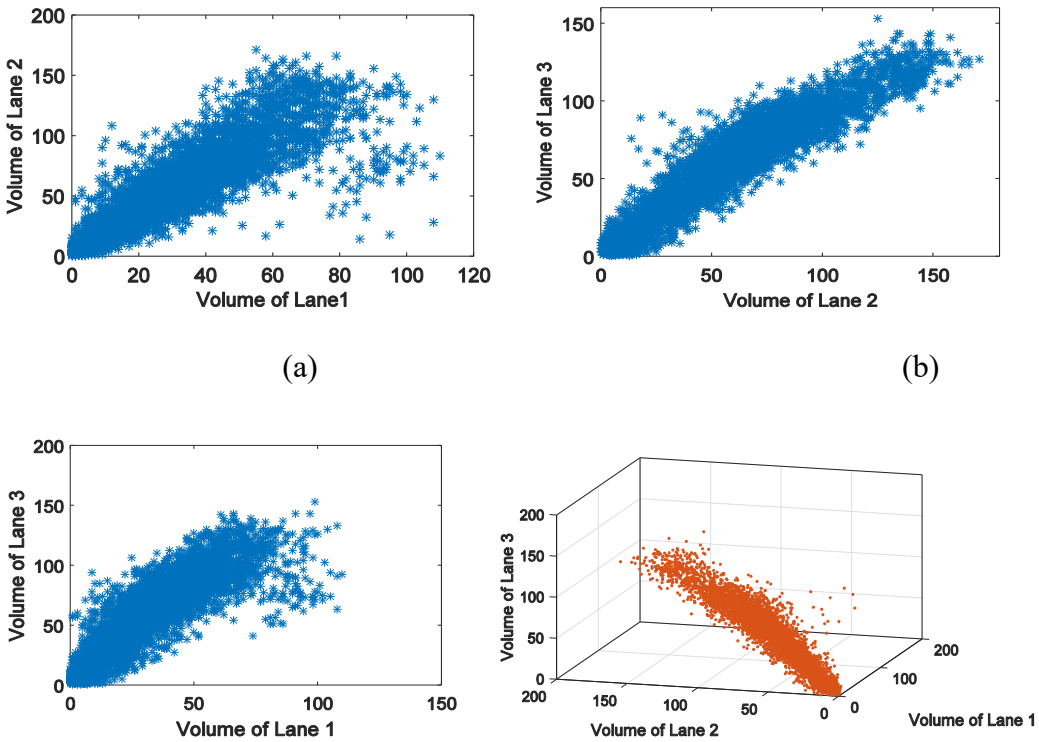


Figure 2 **Volume data correlation among 3 lanes (VDS 1035 of Whitemud Drive, Edmonton) (a): lane 1 vs. lane 2. (b) lane 2 vs. lane 3. (c) lane 1 vs. lane 3. (d) lane 1 vs. lane 2 vs. lane 3.**

Table 1 Volume Values Correlation between Lanes (VDS 1035)

Variables	*Adjusted R-squared*
q1 vs. q2	0.78
q2 vs. q3	0.90
q1 vs. q3	0.77

We use the following multiple linear regression models to relate the volume and density measurements from one lane to all the other lanes:

$$q^{m}_{esti,j}(i,d) = \alpha^{j}_{0} + \alpha^{j}_{1} q^{m}_{meas,j'}(i,d) + \alpha^{j}_{2} q^{m}_{meas,j''}(i,d) + \varepsilon^{m}_{q,j}(i,d)$$

$$\rho^{m}_{esti,j}(i,d) = \beta^{j}_{0} + \beta^{j}_{1} \rho^{m}_{meas,j'}(i,d) + \beta^{j}_{2} \rho^{m}_{meas,j''}(i,d) + \varepsilon^{m}_{\rho,j}(i,d) \tag{3-3}$$

In the above equation, if j=1, then j'=2, j''=3; if j=2, then j'=1, j''=3; if j=3, then j'=1, j''=2. For each target lane j, the parameters α^{j} and β^{j} were estimated using days of historical data. In this case, we used 21 days of historical data. The regression method was least square regression, which means when calculating imputation values error term $\varepsilon^{m}_{q,j}(i,d)$ and $\varepsilon^{m}_{\rho,j}(i,d)$can be treated as zero.

$$\alpha^{j}_{0}, \alpha^{j}_{1}, \alpha^{j}_{2} = \arg\min \left(\frac{1}{n}\sum_{t=1}^{n}\left[q^{m}_{esti,j}(i,d) - \alpha^{j}_{0} - \alpha^{j}_{1} q^{m}_{meas,j'}(i,d) - \alpha^{j}_{2} q^{m}_{meas,j''}(i,d)\right]^{2}\right)$$

$$\beta_0^j, \beta_1^j, \beta_2^j = \arg\min \left(\frac{1}{n}\sum_{t=1}^{n}\left[\rho_{esti,j}^{m}(i,d) - \beta_0^j - \beta_1^j \rho_{meas,j'}^{m}(i,d) - \beta_2^j \rho_{meas,j''}^{m}(i,d)\right]^2\right) \tag{3-4}$$

3.4 Imputation Approach

In this part, a multiple linear regression model was configured in MatLab software. Of the VDSs on the Whitemud Drive road segment, this paper chose VDS 1035 as a test station. The dataset used was VDS 1035 data from November 1, 2013 to November 30, 2013, approximately 129600 records in 20-second intervals and 8640 records in five-minute aggregated intervals. This VDS functions well, without losing data. All data model learning and results comparison was based on the above data population, which meant that for some days, specific lane data would be removed artificially to test the performance of the imputation algorithm. The VDS 1035 is a congestion prone location on one main curve of Whitemud Drive. It has three regular lanes and one off-ramp lane as shown in figure 1 before. In this study we concentrate only on regular lanes.

The whole diagnosis and imputation process is as follows: First, divide the whole month of November 2013 into two groups: the learning group and verification group. The learning group is November 1-21 and the verification group is November 22-30. Second, aggregate the learning group's data as a whole, and conduct multiple linear regressions on volume and density values. Each lane takes turns to be imputed. This step generated six sets of parameters, i.e. 18

parameters in total. Finally, the data of the verification group is used to evaluate the imputation algorithm. The data of each of the three lanes in the verification group was artificially removed and then imputed.

3.4.1 Methods for Comparison

The pairwise linear regression (PLR) and average of surrounding detectors (ASD) methods have also been implemented in this research for comparison with MLR. Pairwise linear regression, as mentioned before, shows high correlation among neighbor loop measurements. Two loops are defined as "neighbors" if they are in the same station in different lanes, or if they are in adjacent ?. In this case, only loops in the same location (station) are regarded as neighbors. The following pairwise linear model was used:

$$PLR: q^{m}_{esti,j}(i,d) = \alpha^{j}_{0} + \alpha^{j}_{1} q^{m}_{meas,j'}(i,d) + noise$$

$$\rho^{m}_{esti,j}(i,d) = \beta^{j}_{0} + \beta^{j}_{1} \rho^{m}_{meas,j'}(i,d) + noise$$

$$ASD: q^{m}_{esti,j}(i,d) = \frac{1}{2}(q^{m}_{meas,j'}(i,d) + q^{m}_{meas,j''}(i,d))$$

(3-5)

For each pair of neighbor lanes (j, j'), the least squares estimation method was used to determine the parameters $(\alpha^{j}_{0}, \alpha^{j}_{1})$ and $(\beta^{j}_{0}, \beta^{j}_{1})$. Chen et al. used five days of historical data to estimate the parameters, while in this study 21 days of historical data, commensurate with the MLR method, are used.

3.4.2 Aggregation Level

The aggregation interval in this research was five minutes. The raw 20-second data had characteristics of too many ($q = 0, \rho = 0$) points and too small of a volume value, usually between 0~7. Raw data fluctuates abruptly between zero and small numbers. This phenomenon conceals the nature of traffic data by making the trend of data obscure. Five-minute aggregation level was chosen because it was the shortest time interval compromising statistics data mining and the most possible usage of high frequency traffic data. Commonly used traffic prediction models also receive data input at five-minute aggregated intervals.

3.4.3 Imputation Experiment Results

The MLR of volume and density imputation, as well as reference methods PLR and ASD, were conducted. In this research, the root mean square error (RMSE) was used to evaluate the performance. Compared to mean absolute error (MAE) and mean absolute percentage error (MAPE) that are typically used in other similar studies, RMSE does not show bias for a big missing percentage as MAE and MAPE do, so it is suitable for this research. As shown below, in this research M represents the number of samples in the verification. The RMSE of q and ρ have the same formation:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{M}\left(x_{esti,j}^{m}(i,d) - x_{true,j}^{m}(i,d)\right)^2}{M}}, x = p \text{ or } x = \rho \tag{3-6}$$

Table 2 and 3 below show coefficient values, tests and model goodness of fit from MLR. It is clear that the three sets of regression showed a high adjusted R-squared value, which generally means any of the three lanes were highly correlated with the other two. The high F statistics and corresponding low to zero P value indicate that the parameters were all significant at a high confidence level. Please note here that those parameters were specific to the aggregation level. They were generated from five-minute aggregation intervals and should be adjusted when applied to other aggregation intervals.

Table 2 Coefficients and Statistics of MLR Model for Volume Measurement

Lane Index j	*1*	*2*	*3*
α_0^j	0.062	-1.689	6.001
α_1^j	0.285	0.344	0.290
α_2^j	0.273	0.834	0.730
R-squared	0.805	0.917	0.915
F statistic*	11801.772	31653.115	30951.074
P value	0	0	0

Table 3 Coefficients and Statistics of MLR Model for Estimated Density

Lane Index j	*1*	*2*	*3*
β_0^j	0.308	0.017	0.662
β_1^j	0.146	0.117	0.244

β_2^j	0.377	0.956	0.773
R-squared	0.681	0.908	0.915
F statistic*	6001.410	27933.557	30464.596
P value	0	0	0

$*$ Significance level$\alpha = 0.05$.

Table 4 and 5 show the RMSE performance of the MLR method as well as the other two comparison methods. For volume imputation, the MLR method consistently outperformed PLR and ASD. Among regression-based imputation methods, PLR has been regarded as the most accurate and computational effective method and deemed the only one suitable for real-time imputation. MLR did not show as large of an advantage over PLR as over ASD, but it saved more time in the regression process; that is to say, for one specific lane, the imputation model was unique. In PLR, the definition of "neighbor" was obscure and the information of other highly correlated lanes was unutilized (3). In density imputation, MLR still shows a smaller RMSE value than the other methods. Considering the time-saving advantage and convenience of MLR, it is still assumed to be a better method.

Figure 3 shows boxplots of the absolute imputation error of the three methods. Here, absolute value is non-aggregated so that each method has one absolute error vector. The upper and lower bottoms of boxes show 25% and 75% error values respectively, and red horizontal line shows mean absolute error. It is clear that in volume imputation, MLR showed the lowest mean absolute error and

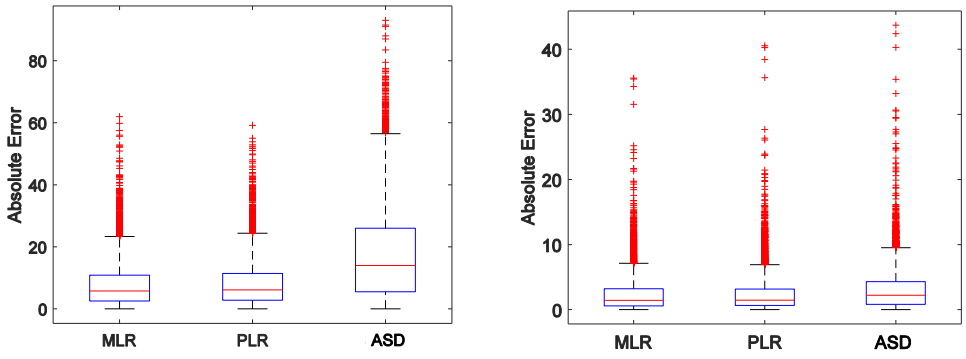
also the lowest variance. In density imputation, MLR and PLR both had the lowest mean absolute error and the least variance between the 75th and 25th percentile values, compared to ASD. Furthermore, MLR had the lowest variance in terms of the whole error vector dataset, and its maximum absolute error is the smallest among the three methods.

Table 4 RMSE Value of Volume Imputation Using MLR, PLR and ASD Methods

Lane Index	*1*	*2*	*3*
MLR	9.954	10.761	10.816
PLR	10.095	11.727	11.000
ASD	31.243	18.474	17.346

Table 5 RMSE Value of Density Imputation Using MLR, PLR and ASD Methods

Lane Index	*1*	*2*	*3*
MLR	2.280	2.318	5.352
PLR	2.103	2.470	5.525
ASD	5.865	3.673	3.077

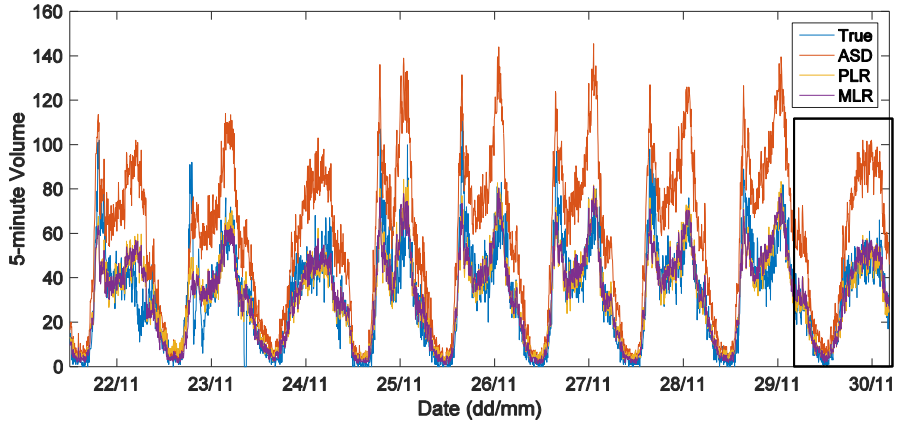


(a) (b)

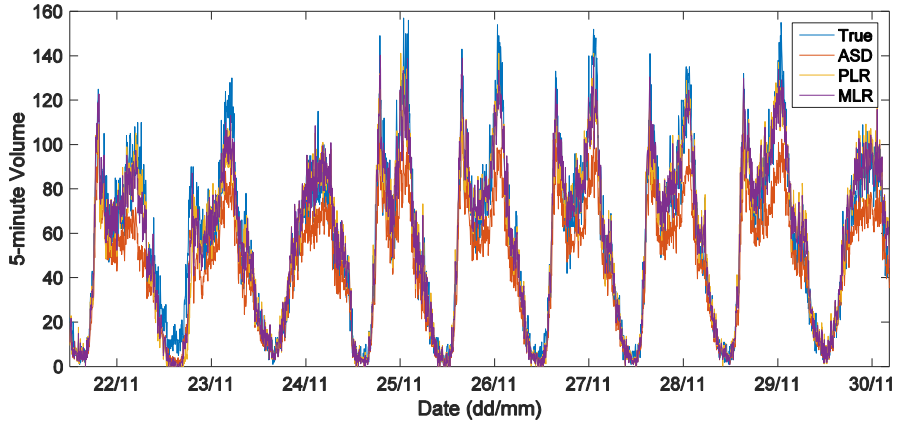
Figure 3 Absolute errors of three methods for (a) volume, and (b) density imputation

The quantitative measurement of imputation error cannot show the overall performance of imputation for the following two reasons. First, the large amount of data loop detectors collected contains some random characteristics. So that the exact calculation of RMSE was not repeatable; i.e., if we had the chance to collect data again for the same lane at the same time, the value of RMSE could change. Second, the main goal of lane imputation was to capture the sensitive trend of data fluctuation. A successful volume and density imputation should show the recurrent as well as non-recurrent increase and decrease as with true data. This trend fit can be clearly examined from graphics.

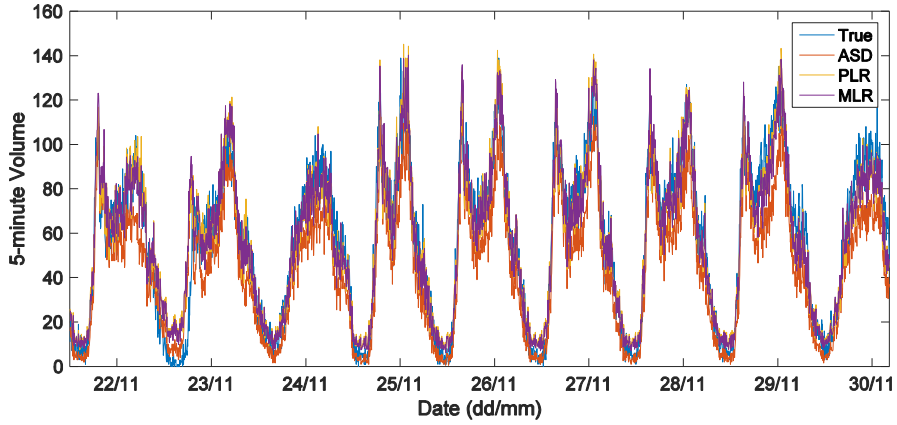
Figure 4 graphically shows the performance of volume data imputation for the verification group on VDS 1035 if lane 1, 2 or 3 constantly miss data, respectively. Part (d) of Figure 4 is a zoomed-in picture showing the imputation for the specific day of November 30 for missing lane 1.


160
140
120
100
80
60
40
20
0
5-minute Volume
22/11
23/11
24/11
25/11
26/11
27/11
28/11
29/11
30/11
Date (dd/mm)
True
ASD
PLR
MLR

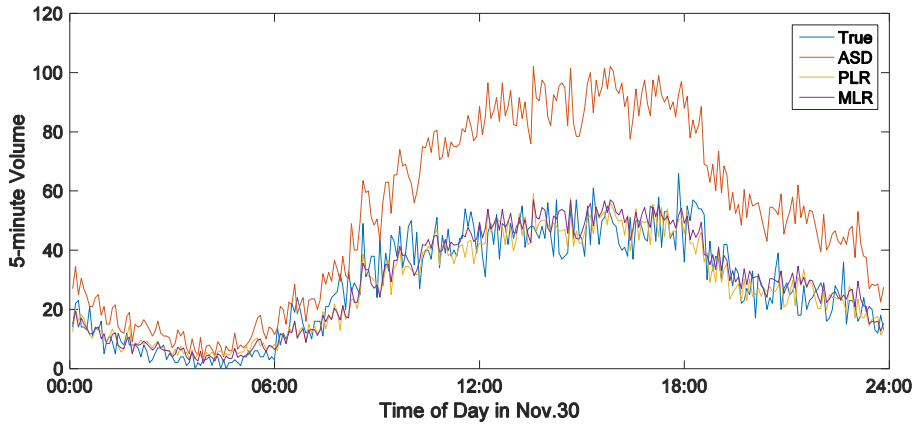
(a)


160
140
120
100
80
60
40
20
0
5-minute Volume
22/11
23/11
24/11
25/11
26/11
27/11
28/11
29/11
30/11
Date (dd/mm)
True
ASD
PLR
MLR

(b)


160
140
120
100
80
60
40
20
0
5-minute Volume
22/11
23/11
24/11
25/11
26/11
27/11
28/11
29/11
30/11
Date (dd/mm)
True
ASD
PLR
MLR

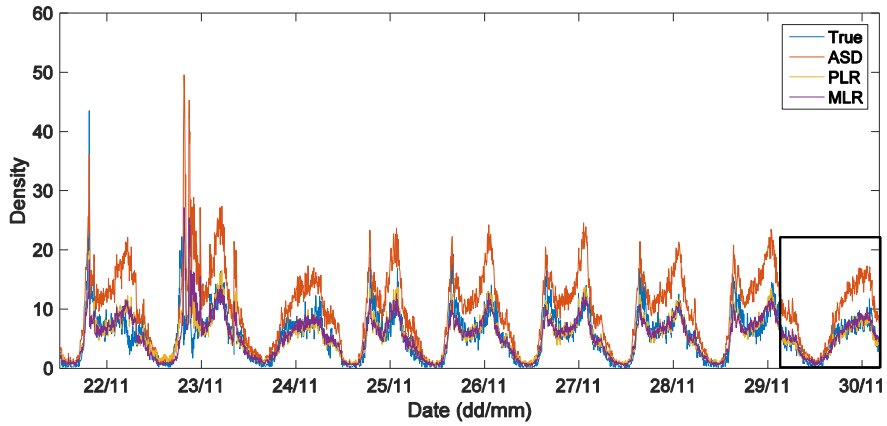
(c)



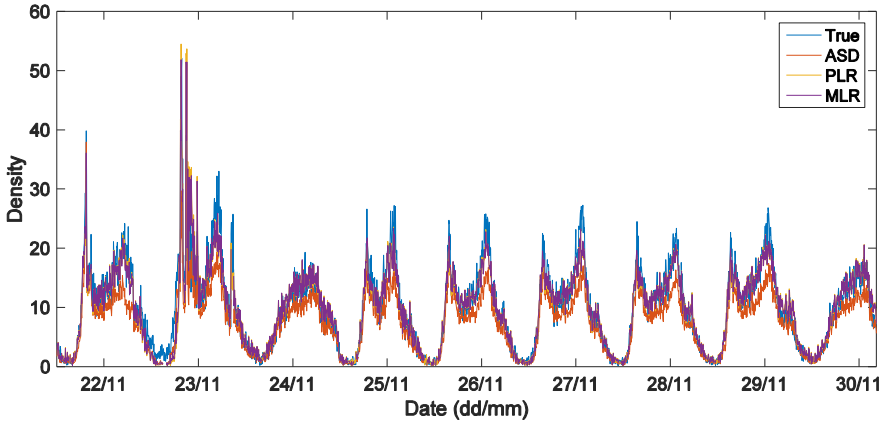
(d)

Figure 4 **Performance of volume data imputation for verification group using MLR, PLR and ASD methods versus true data: (a) missing lane 1, (b) missing lane 2, (c) missing lane 3, and (d) missing lane 1 on Nov. 30, 2013.**

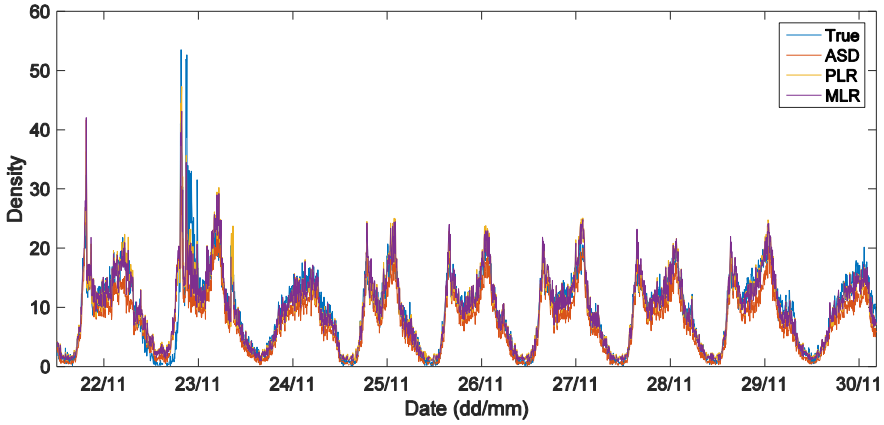
Figure 5 graphically shows the performance of the density imputation for the verification group of November 2013 for VDS 1035 if lane 1, 2 or 3 constantly miss data, respectively. Part (d) of Figure 4 is a zoomed-in picture showing the imputation for the day of November 30 for missing lane 1.



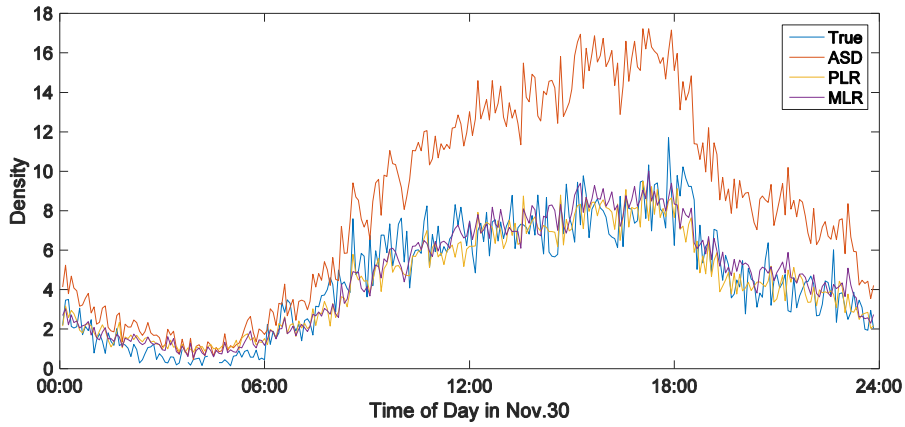
(a)


60
50
40
30
20
10
0
Density
22/11
23/11
24/11
25/11
26/11
27/11
28/11
29/11
30/11
Date (dd/mm)
True
ASD
PLR
MLR

(b)


60
50
40
30
20
10
0
Density
22/11
23/11
24/11
25/11
26/11
27/11
28/11
29/11
30/11
Date (dd/mm)
True
ASD
PLR
MLR

(c)



(d)

Figure 5 **Performance of density data imputation for verification group using MLR, PLR and ASD methods versus true data: (a) missing lane 1, (b) missing lane 2, (c) missing lane 3, and (d) missing lane 1 on Nov. 30, 2013.**

3.5 Summary of Loop Detector Data and Imputation

In this part, loop detector data is shown to be reliable especially when the imputation algorithm is proved to be good. DDA is presented to detect invalid loop detectors and then impute missing volume and density data based on real-time data from lanes at the same station with a complete dataset. Existing methods of imputation mainly focus on offline panel data imputation. The missing type is usually random. Previous methods had high accuracy, relatively high computation complexity and were perfectly suitable for data stored offline for future research use. Those characteristics made many previous methods not suitable for online imputation, for that the computing and display time have to be within 20 seconds. In the field, hardware damage or instability leads to constantly missing data for

one or more specific lanes. If hardware failure cannot be fixed timely, the imputation algorithm for the missing lane data should be triggered.

As specifically designed, the diagnosis and imputation method for online loop detector systems should generate qualified data for real-time traffic control. The DDA diagnosis method detects missing data every day with the help of the first five minutes of data at the beginning of that day. The diagnosis is not solely based on one station but instead on the station and its upstream station together. Traditionally, the most used imputation methods for missing lane data include historical data filling, average of previous time periods, ASD and so on. They are quick interpolation techniques but lack accuracy, and they cannot capture the trend of ongoing traffic conditions. The MLR imputation method in this paper fully utilizes historical information and is more than simple interpolation. PLR has been used as a benchmark to justify the quality of imputation. MLR consistently outperformed PLR, especially on volume imputation. Taking into account its operational ease, stability and shorter computing time, the MLR method has a greater advantage over PLR.

However, there is still great potential to improve the algorithms in this study. The diagnosis algorithm here relies on data from the upstream loop detector station. When the upstream corresponding lane is also absent of data, the algorithm fails. The case suggested above is rare but is still a possibility. The imputation algorithm is currently suitable for missing data from one lane. If two or more lanes at the same station all miss data, the imputation can still be carried out, but the accuracy is untested. Imputation relies on good historical data from all

lanes, and if the historical data has bias, then the imputation will be biased. Due to the formation of the imputation regression model, the imputed values sometimes are a little smaller than zero, although this kind of error does not have an observable influence on the traffic control algorithm.

CHAPTER 4. A CASE STUDY OF MODIFYING METANET MODEL

This chapter presents the feasibility of modifying one critical term "desire speed" in METANET traffic state prediction model to suite different needs including VSL control environment and varying weather environment. To be specific, this chapter tries making the desire speed term weather specific to improve prediction performance using real field data, and that the winter of Edmonton is taken as example.

4.1 Introduction and Background

This part will proposes a weather factor modelling method for macroscopic traffic prediction model, considering adverse weather brings about significant impact on road conditions and traffic dynamics. We suggest that adverse weather as a set of exogenous factors lower the free flow speed, shift critical density, decrease flow capacity and make the freeway more prone to congestion. The non-negligible different road nature arouse the question of building weather-specific fundamental diagrams so that the traffic state prediction model can be more accurate under varying weather conditions, and then to be more reliable to be applied to dynamic traffic control. The entire process of traffic state prediction is computed by METANET model. Prediction error of weather-specific prediction and conventional overall prediction is compared. Real data collected by loop detectors on freeway Whitemud Drive, Edmonton, Canada is used for parameter calibration

and prediction error evaluation. The results show that proposed weather models reasonably enhanced the accuracy of macro traffic state prediction model than conventional one.

In this part, the weather factor modelling is developed and then be inserted to traffic state prediction model using field data. In comparison with the previous work, this study has three contributions. Firstly, more than one weather factors are considered and filtered under a wider range of weather conditions, varying from the worst to the best weather, on different macroscopic traffic variables. Whereas previous works only consider one category of extreme weather and one traffic variable. Secondly, weather factors are successfully introduced into fundamental diagram as well as traffic prediction model. Thirdly, this paper shows clear quantitative result that weather significantly impact the traffic dynamics on freeways using high resolution field data.

4.2 The Modified METANET Model with Consideration of Weather Condition

The model description is separated into four parts. Section 4.2.1 will be notations and assumptions. Section 4.2.2 and 4.2.3 will present the description of the traffic state prediction model and the description of the fundamental diagram used in this study. Section 4.2.4 will describe model calibration skills.

4.2.1 Notations and assumptions

For consistency, the freeway are divided into N sections with lengths Δ_i, $i = 1,...,N$, each having at most one on-ramp and off-ramp. All the variables used throughout this paper are defined as following.

T	Data collection interval=20s
Δ_i	Length of section *i*
λ_i	Number of lanes at section *i*
Ω	Set of space and time (x,t)
$\rho(x,t)$	Traffic density at time *t*, space *x*.
$v(x,t)$	Vehicle space mean speed at time *t*, space *x*.
$r(t,x)$	On-ramp flow at time *t*, space *x*
$s(t,x)$	Off-ramp flow at time *t*, space *x*
$q_i(k)$	Number of vehicles in the freeway section i at time $k \cdot T$ divided by the length Δ_i
$\rho_i(k)$	Traffic density at time step *k*, section *i*.
$r_i(k)$	On-ramp flow at time step *k*, section *i*
$s_i(k)$	Off-ramp flow at time step *k*, section *i*
$V(\rho_i(k))$	Desire speed in speed dynamics in METANET model
Θ	Set of unknown parameters of METANT model
$\rho_{cr,i}$	Critical density of FD at section *i*
$\rho_{jam,i}$	Jam density of FD at section *i*

$v_{f,i}$	Free flow speed of FD at section *i*
θ	Capacity drop fraction of FD at section *i*
$w_i^{v_{f,i}}$	Weather adjust factor for free flow speed of FD at section *i*
$w_i^{\rho_{cr,i}}$	Weather adjust factor for critical density of FD at section *i*
$w_i^{\rho_{jam,i}}$	Weather adjust factor for jam density of FD at section *i*

4.2.2 The framework of METANET prediction model

The development of mathematical model that describes the dynamic evolution of three traffic variables enables the short time prediction of macroscopic traffic statues. METANET model has three dynamics that describes flow, density and speed. Among three dynamics, the flow dynamics and density dynamics are exact analytical models without parameter calibration, and are derivate from matter conservation law [52][53] written as follows:

$$\frac{\partial \rho(t,x)}{\partial t} + \frac{\partial q(t,x)}{\partial x} = r(t,x) - s(t,x) \tag{4-1}$$

This conservation law equation indicated the fact that the vehicle entering one section will eventually exit, either to next section of main road or the off-ramp. The flow and density dynamics together are the first order part of METANET model. However with solely first order dynamics one cannot describe the dynamics of speed change, although there is certain relationship between speed and density, speed do not change instantaneously with density in the real world.

Payne proposed that a small time delay should be applied to the speed-density relationship in fundamental diagram as follows:

$$v(x,t+\tau)=V\left[\rho(x+\Delta x,t)\right] \tag{4-2}$$

Papageorgiou et al. expanded the left side of the above equation in a Taylor series respect to τ and right hand side with respect to Δx, after rearranging terms we will get the following equation:

$$\tau\cdot\frac{\partial v}{\partial t}=V(\rho)-v-\frac{v}{\rho}\cdot\frac{\partial\rho}{\partial x} \tag{4-3}$$

Where the arguments x and t are depressed for convenience, $\mathrm{d}v/\mathrm{d}t$ is the acceleration of an observer moving with the traffic stream [28], $\mathrm{d}v/\mathrm{d}t$ is written as follows:

$$\frac{\mathrm{d}v}{\mathrm{d}t}=\frac{\partial v}{\partial t}+v\cdot\frac{\partial v}{\partial x} \tag{4-4}$$

Substituting (4-3) into (4-4) then we can get the continuous form of speed dynamics model

$$\frac{\mathrm{d}v}{\mathrm{d}t}=v\cdot\frac{\partial v}{\partial x}+\left[V(\rho)-v-\frac{v}{\rho}\cdot\frac{\partial\rho}{\partial x}\right]/\tau \tag{4-5}$$

Equation (4-1), (4-5) and the identical equation $q=\rho\cdot v$ together forms the three dynamics of METANET model. With the introduction of speed dynamics model, METANET is expected to be one of the most accurate macroscopic traffic dynamics model. The discretized model formulation is written as follows:

Density Dynamics

$$\rho_i(k+1)=\rho_i(k)+\frac{T}{\Delta_i\lambda_i}\left[\lambda_{i-1}q_{i-1}(k)-\lambda_i q_i(k)+r_i(k)-s_i(k)\right] \tag{4-6}$$

Speed Dynamics

$$v_i(k+1) = v_i(k) + \frac{T}{\tau}\left[V\left(\rho_i(k)\right) - v_i(k)\right] + \frac{T}{\Delta_i} v_i(k)\left[v_{i-1}(k) - v_i(k)\right] - \frac{\eta \cdot T\left[\rho_{i-1}(k) - \rho_i(k)\right]}{\tau \cdot \Delta_i \cdot \left[\rho_i(k) + \kappa\right]}$$

(4-7)

Flow Dynamics

$$q_i\left(k+1\right) = \rho_i\left(k+1\right) \cdot v_i\left(k+1\right) \tag{4-8}$$

In the model τ, κ, α and η (km^2/h) are global model parameters to be calibrated using the historical data. The desired speed $V\left(\rho_i\left(k\right)\right)$ (km/h) in speed dynamics is represented by

$$V\left(\rho_i\left(k\right)\right) = v_{f,i} \exp\left[-\frac{1}{\alpha}\left(\frac{\rho_i\left(k\right)}{\rho_{c,i}}\right)^{\alpha}\right] \tag{4-9}$$

The desired speed variable leaves room for the introduction of weather factors because it involves two important parameters calibrated from the fundamental diagram: free flow speed and critical density, and the fundamental diagram is impacted by weather. Density dynamics do not involve parameters and other parameters in speed dynamics are calibrated globally for goodness of fit. Note that of the four terms making up the speed dynamics, each term has a physical meaning. In equation (4-7), the second term is referred to as the relaxation term, describing that with a lag time item τ, the mean speed v of the link gets relaxed to the desired speed which largely depends on parameters of FD. The selection of the desired speed is critical to reflect the driver behavior and from previous practice we chose the format of equation (4-9). The third is the convection term meaning that vehicles travelling from upstream link $i-1$ to current link i gradually adapt their speed rather than instantaneously. The fourth is anticipation

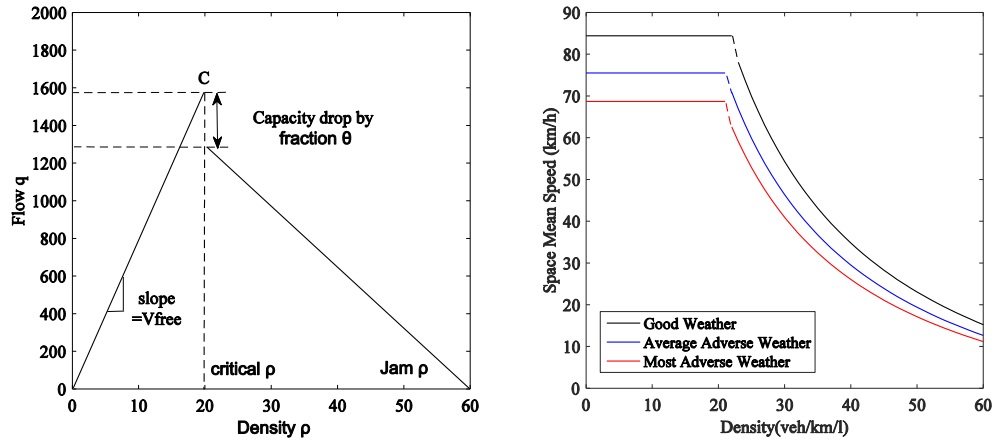
term meaning that drivers are always keeping an eye on the traffic condition ahead. If a driver observes high traffic density in the downstream link $i+1$, he then reacts as slowing down, and vice versa. The constant $\kappa > 0$ is added to keep the anticipation term limited when density is low. It comes to the conclusion that the desire speed item is the right place to insert weather factor considering weather specific fundamental diagrams.

The Courant–Friedichs–Lewy (CFL) condition is followed, which means to grant that vehicles cannot travel beyond one link within computing time interval *T*, so that *T* satisfies the following condition:

$$T \le \frac{\Delta_i}{\max\left(v_{f,i}\right)}, i = 1, 2, ..., N \tag{4-10}$$

4.2.3 The weather specific fundamental diagrams

It is assumed that the fundamental diagram (FD) changes according to different weather conditions, since weather significantly impacts driver behavior and the driving environment. The weather-specific FD defines different $V\left(\rho_i\left(k\right)\right)$ for different weather conditions in the METANET model. To be specific, the parameters of FD such as free flow speed, critical density, jam density and capacity drop vary with weather conditions instead of being constants as regarded before. And those key parameters in the FD are essential for the accuracy of the traffic prediction model. The triangular FD is still used here shown as Figure 6(a). Note that, to better exhibit the different free flow speeds under different weather conditions, we also demonstrate the variation form of the triangular FD, for which the vertical coordinate is space mean speed.



(a) (b)

Figure 6 **Illustration of (a) triangular FD and (b) its variation with speed as vertical coordinate.**

From the point of the space mean speed, the triangular FD describes that before density reaching critical density, space mean speed keeps in free flow speed. After reaching critical density, traffic becomes congested and the road segment capacity drops by fraction θ due to the unsatisfying driving environment. After reaching the congestion point, traffic flow decreases linearly to zero, which is when density also reaches maximum, and that density is referred to as jam density. In the variation form of the triangular FD, space mean speed behaves similarly, remaining constant up until the congestion point, and after that, speed will drop together with capacity. During congestion, the space mean speed will drop as inverse proportional function of density. Figure 6 (b) shows that we anticipate that under different weather conditions FD will shift. The expression of the variation form of the FD is as follows:

$$v_i(k+1)=\begin{cases} v_{f,i}, & \text{if } \rho_i(k+1)\in(0,\rho_{cr,i}) \\ \dfrac{v_{f,i}\cdot\rho_{cr,i}}{\rho_i(k+1)}\left(1-\theta-\dfrac{\rho_i(k+1)-\rho_{cr,i}}{\rho_{jam,i}-\rho_{cr,i}}\right), & \text{otherwise} \end{cases} \tag{4-11}$$

In the equation (4-11) weather factors will be introduced to adjust the free flow speed, critical density and jam density of FD under varying weather conditions. In part 3 of this paper, the modeling of three weather factors will be presented. Equation (4-12) describes the new form of flow dynamics considering capacity drop and weather factors and this equation will replace the original flow dynamics in METANET model.

$$q_i(k+1)=\begin{cases} \rho_i(k+1)\cdot v_i(k+1), \text{ if } \rho_i(k+1)\in(0,\rho_{cr,i}) \\ \min(\rho_i(k+1)\cdot v_i(k+1),(1-\theta)\cdot v_{f,i}\cdot\rho_{cr,i}, \text{ otherwise} \end{cases} \tag{4-12}$$

4.2.4 Fundamental diagram and METANET calibration methods

The parameter estimation of the FD is based on data collected by conventional loop detectors. For the parameter calibration of FDs, free flow speed $v_{f,i}$, critical density $\rho_{cr,i}$, capacity *C*, jam density $\rho_{jam,i}$ and capacity drop fraction θ must be estimated. For calibration, the data format we use is (ρ_i, q_i) data points, where the horizontal coordinate is density and the vertical coordinate is flow. The procedure of calibrating a triangular FD is as follows:

- Step1: The identification of $\rho_{cr,i}$ and capacity C. In the definition of triangular FD, the summit point of the triangle roughly indicates critical density and capacity. To be specific, we plot all (ρ_i, q_i) points and find the 3rd largest q_i, take it as capacity C and the corresponding ρ_i as $\rho_{cr,i}$. The

reason not to choose the largest flow is that rarely the largest flow is extremely high as an outlier, this might due to the detection error.

- Step 2: The identification of $v_{f,i}$. After defining $\rho_{cr,i}$, the whole dataset can be divided into two parts: the left-side triangle represents uncongested traffic conditions and the right-side triangle represents congested traffic conditions. Then we calculate the slope of each data point distributed in the left side and take an average as $v_{f,i}$, which is described in equation (4-13). In this equation *n* represents the number of data points within the left-side triangle.

$$v_{f,i} = \frac{1}{n}\left(\sum_{ii=1}^{n} \frac{q_{ii}}{\rho_{ii}}\right), \rho_{ii} \in \left(0, \rho_{cr,i}\right) \tag{4-13}$$

- Step 3: The identification of $\rho_{jam,i}$ and capacity drop fraction θ. The jam density $\rho_{jam,i}$ represents a theoretical value of when the road section is totally congested and all vehicles have stopped moving; however, in the real dataset, this point is seldom observed. When determining the right-side triangle, we fix the $\rho_{jam,i}$ with an empirical value. In the case study section, the empirical value will be decided. The slope of the right side is determined by least square fit. After determine the foot and slope of right side, the intersecting point of the right side and the vertical auxiliary line passing through previous capacity *C* is the new capacity after dropping, and the capacity drop fraction θ is calculated as follows:

$$\theta = 1 - \frac{\left(\rho_{cr,i} - \rho_{jam}\right)\sum_{ii=1}^{n}\left(\rho_{ii} - \bar{\rho}\right)\left(q_{ii} - \bar{q}\right)}{C \cdot \sum_{ii=1}^{n}\left(\rho_{ii} - \bar{\rho}\right)^2}, \rho_{ii} \in \left(\rho_{cr,i}, \rho_{jam_{,i}}\right) \tag{4-14}$$

For the parameter estimation of METANET model, τ, κ, α and η (km^2/h) are global model parameters to be calibrated using recent historical data. In this study, the global parameters of METANET model are fixed, which are calibrated before this study, using loop detector data of the same road segment via the following expression.

$$\Theta^* = \arg\min_{\Theta}\left[\sum_{ii}^{N}\left\{v^{ii}{}_{obs} - f(v^{ii}{}_{obs} \mid \Theta)\right\}^2 + \sum_{ii}^{N}\left\{\rho^{ii}{}_{obs} - f(\rho^{ii}{}_{obs} \mid \Theta)\right\}^2\right] \tag{4-15}$$

4.3 Modelling Weather Impacts on Fundamental Diagram

The historical weather data provided by Canadian government website, historical climate data webpage, Edmonton area includes three critical categories of weather data: temperature, amount of snow on the ground, and maximum wind speed. We assume that Free Flow Speed (FFS), capacity, and critical density of FD of each day are impacted by weather. Those three categories of weather index are capable of express most kinds of weather in Edmonton, Canada. Since people in Edmonton seldom experience rainfall that is heavy enough to influence visibility, fog is also rare to see, so that the most common adverse weather is snowy weather. HCM has pointed out that little snow would reduce capacity by 5-10% and heavy snow reduces capacity by 25-30%. From observing and previous data,

we assume that ongoing snow may have significant impact on FFS. Due to the variable "snow on the ground" may reflect a mean value, "derivative of snow on the ground" is added besides the origin variable "snow on the ground" which is formulated as below.

$$\frac{\mathrm{d}SG_j}{\mathrm{d}t_j} = \frac{SG_j - SG_{j-1}}{t_j - t_{j-1}} + \varepsilon,\ \varepsilon \sim \tag{4-16}$$

SG_j is the depth of the snow on the ground at day j. t_j is the date and ε is the error term. In the dataset, SG_j is the depth of snow measured with centimeter in the area of the city where the target VDS located.

$$DSG_j = \frac{SG_j - SG_{j-1}}{t_j - t_{j-1}} \tag{4-17}$$

DSG_j is a derivative of snow on the ground which is used in our model. We assume that the change of SG_j indicates that some weather events are going on. The positive DSG means that snow is going on, and the negative DSG means that the weather is becoming better and snow is melting. We also use $Temp$ to represent temperature and MWS to represent max wind speed. To filter the key factors impacting FD features, we did T-test prior to our case study utilizing one loop detector station data of November of 2013 in Edmonton, Canada. From T-test, we found that DSG has significant impact on both FFS and capacity at 95% confidence level. Surprisingly, MWS and $Temp$ are found not significant so that we suggest that drivers of this area are confident to drive normally under different speed of wind and temperature.

Table 6 The key weather factors impacting FD features.

*Significance**	*SG*	*DSG*	*MWS*	*Temp*
FFS	YES	YES	NO	NO
Capacity	NO	YES	NO	NO
Critical Density	NO	NO	NO	NO

* Significance at 95% confidence level

The flow chart in Figure 7 shows how DSG and SG affect FFS, capacity and critical density in the model of this paper.

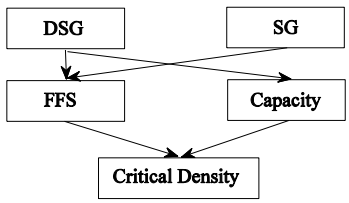


Figure 7 The significant weather factors impacting FD.

The fundamental diagram under varying weather is formulated as below. Three weather factors $w_i^{v_{f,i}}$, $w_i^{\rho_{cr,i}}$ and $w_j^{C_{cr},\mathrm{i}}$ are introduced to adjust the free flow speed, critical density and capacity of FD under varying weather conditions. The capacity adjust factor $w_j^{C_{cr},\mathrm{i}}$ will not appear in the following FD expression but the critical density adjust factor $w_i^{\rho_{cr,i}}$ is computed from $w_j^{C_{cr},\mathrm{i}}$ as written as follows

$$v_i(k+1)=\begin{cases} w_j^{v_{f,i}}\cdot v_{f,i}, & \text{if } \rho_i(k+1)\in(0,\rho_{cr,i}) \\ \dfrac{w_j^{v_{f,i}}\cdot v_{f,i}\cdot w_j^{\rho_{cr,i}}\cdot\rho_{cr,i}}{\rho_i(k+1)}\left(1-\theta-\dfrac{\rho_i(k+1)-w_i^{\rho_{cr,i}}\cdot\rho_{cr,i}}{\rho_{jam,i}-w_j^{\rho_{cr,i}}\cdot\rho_{cr,i}}\right), & \text{otherwise} \end{cases} \tag{4-18}$$

So that the flow dynamics will be updated as follows

$$q_i(k+1)=\begin{cases} \rho_i(k+1)\cdot v_i(k+1), & \text{if } \rho_i(k+1)\in(0,\rho_{cr,i}) \\ \min(\rho_i(k+1)\cdot v_i(k+1),(1-\theta)\cdot w_j^{v_{f,i}}\cdot v_{f,i}\cdot w_j^{\rho_{cr,i}}\cdot\rho_{cr,i}, & \text{otherwise} \end{cases} \tag{4-19}$$

In which

$$w_j^{v_{f,i}} = \beta_0 + \beta_1 DSG + \beta_2 SG \tag{4-20}$$

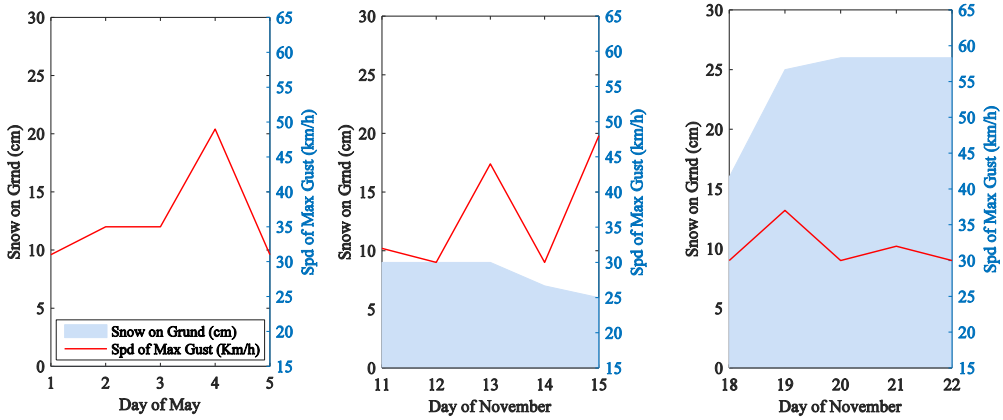
$$w_j^{C_{cr},\mathrm{i}} = \alpha_0 + \alpha_1 DSG \tag{4-21}$$

$$w_j^{\rho_{cr,i}} = \frac{w_j^{C_{cr},\mathrm{i}}}{w_j^{v_{f,i}}} + \delta = \frac{\alpha_0 + \alpha_1 DSG}{\beta_0 + \beta_1 DSG + \beta_2 SG} + \delta \tag{4-22}$$

Where δ is adjusted factor. This model describes that FFS is mainly impacted by SG_j which indicates the current weather situation and DSG which indicates the ongoing weather events such as snow storm going on. Capacity of FD is mainly impacted by DSG solely. And critical density factor is calculated from the above two factors with a constant adjust factor added.

4.4 Case Study in Edmonton

The calibration and validation of the models are based on the data of target VDSs from May and November of 2013. In this study, three weather conditions are involved in testing the stability of the models. Accordingly, three different fundamental diagrams and weather-related parameters are generated. In Edmonton, Canada, the main type of adverse weather is snowy weather, so the three weather conditions are categorized as “good weather condition,” “light snow condition” and “heavy snow condition.” During the day of May 01-05, the snow melt and temperature indicated "good weather conditions." From November 11th to 15th, "light snow conditions" were observed, with the snow on the ground measuring around 6-9 cm. During the days of November 18th to 22nd, "heavy snow conditions" were present, as the amount of snow on the ground ranged 16-26 cm. Figure 3 shows the indexes of the three weather conditions.



(a) (b) (c)

Figure 8 **Ground snow and speed of wind under three weather conditions.**

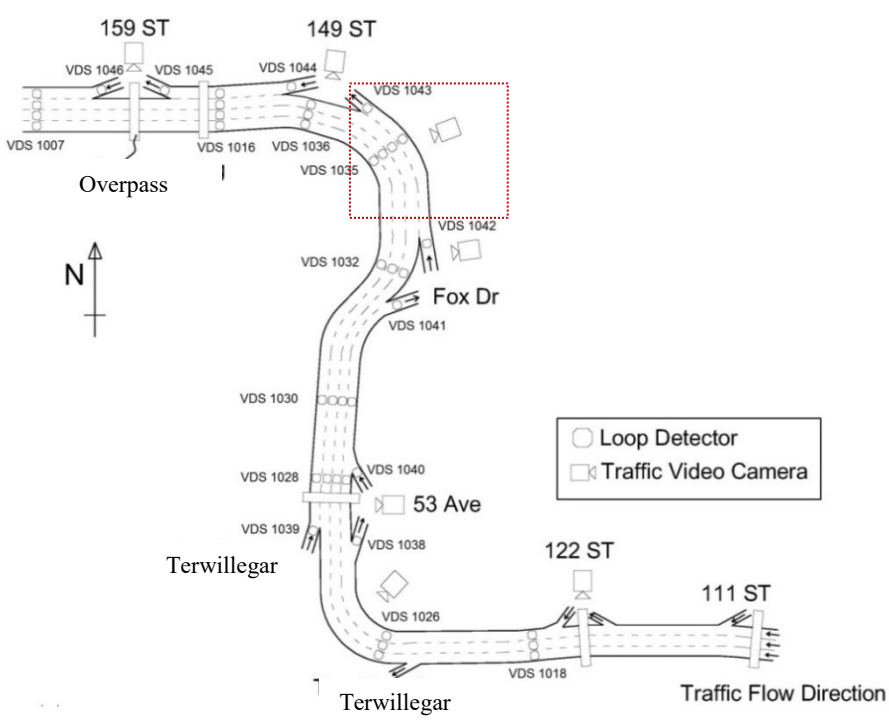


Figure 9 **The location of target VDS on Whitemud Drive, Edmonton**.

Table 7 shows the calibrated FD features and weather factors, while Figure 5 combines data points and the calibrated FD together to help visualize the difference in macroscopic traffic conditions under different types of weather. From the field data we observed that the triangular FD holds. Note that in the calibration process, some obvious outlier points have been eliminated. The significant differences between the types of weather are visibly evident. For free flow speed, good weather conditions experienced the highest free flow speed, which is higher than the posted speed limit (80 km/h); under light snow conditions, the free flow speed is slightly lower than that under good weather and almost equal to the posted speed limit; when it snows heavily, the free flow speed drops drastically to 66.7 km/h. In terms of capacity, which cannot be directly observed from the FD, under good weather conditions the capacity is 1647 veh/h/l, while under light snow conditions the capacity is slightly lower with a number of 1572 veh/h/l. However, under heavy snow conditions the capacity drops as low as 1323 veh/h/l. In terms of critical density, the difference among the weather conditions is not as significant as the previous two parameters. Good weather conditions still show the largest critical density of 24.84 veh/km/l, and under light snow conditions and heavy snow conditions, the critical densities are 22.30 veh/km/l and 21.46 veh/km/l respectively. Jam density represents the ability of a road segment to accommodate vehicles, so in this paper we assume that jam density does not change since the length and number of lanes does not change with the weather.

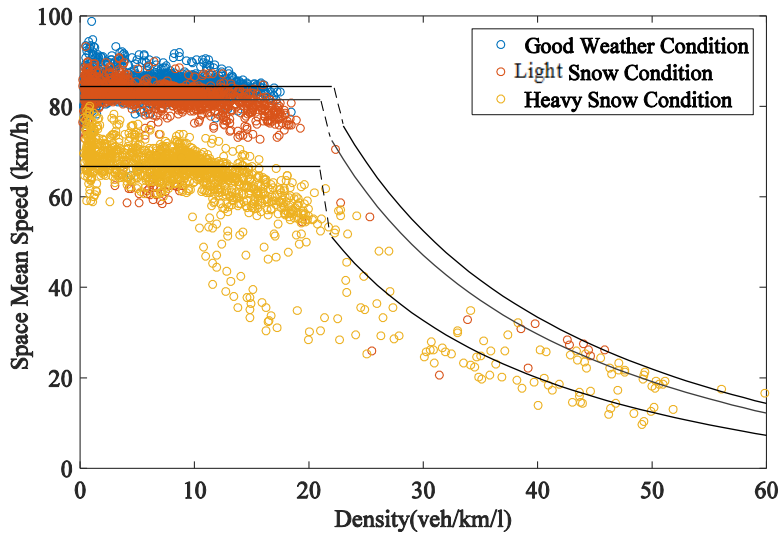


Figure 10 **Field traffic data and calibrated FD under three weather conditions**

Table 7 Fundamental Diagram Features in Three Conditions

Features	*Good weather condition*	*Light snow condition*	*Heavy snow condition*
$v_{f,i}$ (km/h)	84.38	81.52	66.66
Capacity (veh/h/l)	1647	1572	1323
$\rho_{cr,i}$ (veh/km/l)	24.84	22.30	21.46
$\rho_{jam,i}$ (veh/km/l)	100	100	100
θ (Capacity Drop)	0.05	0.09	0.17

Table 8 Estimated Weather Factor Parameters and Statistics

	α_0	α_1	β_0	β_1
Coefficients with	0.873	-0.01796	0.9648	-0.01737
	(0.83, 0.90)	(-0.029, -0.006)	(0.925, 1.004)	(-0.0236, -0.0111)
	β_2	δ		
(95% confidence bounds)	-0.00105	0.1344		
	(-0.0034, 0.00136)			
Adjusted R-square	$w_j^{C_{cr},i}$ model:0.7924	$w_j^{v_{f,i}}$ model:0.833		

Table 8 shows the regression results of the weather factor parameters and statistics. The weather data used in this regression include the DSG and SG from May 01-05, November 11-15 and November 18-22 of 2013. Loop detector data for the same period is also used to calibrate the FDs, and the data frequency is 20 seconds. The negative α_1 indicates that capacity decreases with a positive DSG, and a positive DSG indicates that snow is accumulating and road conditions are worsening. The negative β_1 and β_2 indicate that the free flow speed will decrease with a positive value of DSG and SG, which implies that the snowfall is relatively heavy and snow is accumulating. Note that all the parameters fall within 95% confidence bounds and both regressions have a good fit, see adjusted R-square.

To validate the necessity of considering weather factors in METANET prediction model, we conducted prediction simulation in which the seed of each

round of calculation is field data. And we compared the prediction accuracy of not adding weather factor and adding weather factor conditions. The prediction scenario is set to be 10 min with a calculating frequency of 20s with rolling horizon. The calculation interval is 20 seconds and that to predict traffic status 10 minutes in advance. And we do 30 iterations at each round of calculation. Since the prediction simulation is not real time, we evaluate the prediction accuracy by comparing the traffic state of 10 min later predicted via METANET and the real traffic state of 10 min later in dataset.

In each pair of the following three comparisons, the "not adding weather factor" condition always use the fixed parameters $v_{f,i}$ = 80.06; $\rho_{cr,i}$ = 23.83. Note that those fix parameters together with other global parameters are used in our previous practices, they represents a general situation that under uncongested situation free flow speed can be slightly higher than speed limit 80 km/h, however under unsatisfying weather condition those setting might problematic. Based on weather specific FD assumption mentioned above, we change the value of $v_{f,i}$ and $\rho_{cr,i}$ given calibrated weather factors, while other global parameters of METANET remains unchanged. From each kind of weather conditions we picked up one day to conduct prediction experiment. For heavy snow weather the chosen day is Nov.18, and for light snow weather the day is Nov.14, and for good weather that day is May.02. During each day the traffic state prediction is from 6 AM to 9 PM. Here as follows we only exhibit speed and density prediction results for that flow prediction is calculated through flow identical equation from speed and density. The global parameter values are shown in Table 9.

Table 9 Calibrated METANET Global Parameters

τ	η	κ	α
120.00	37.98	10.00	2.29

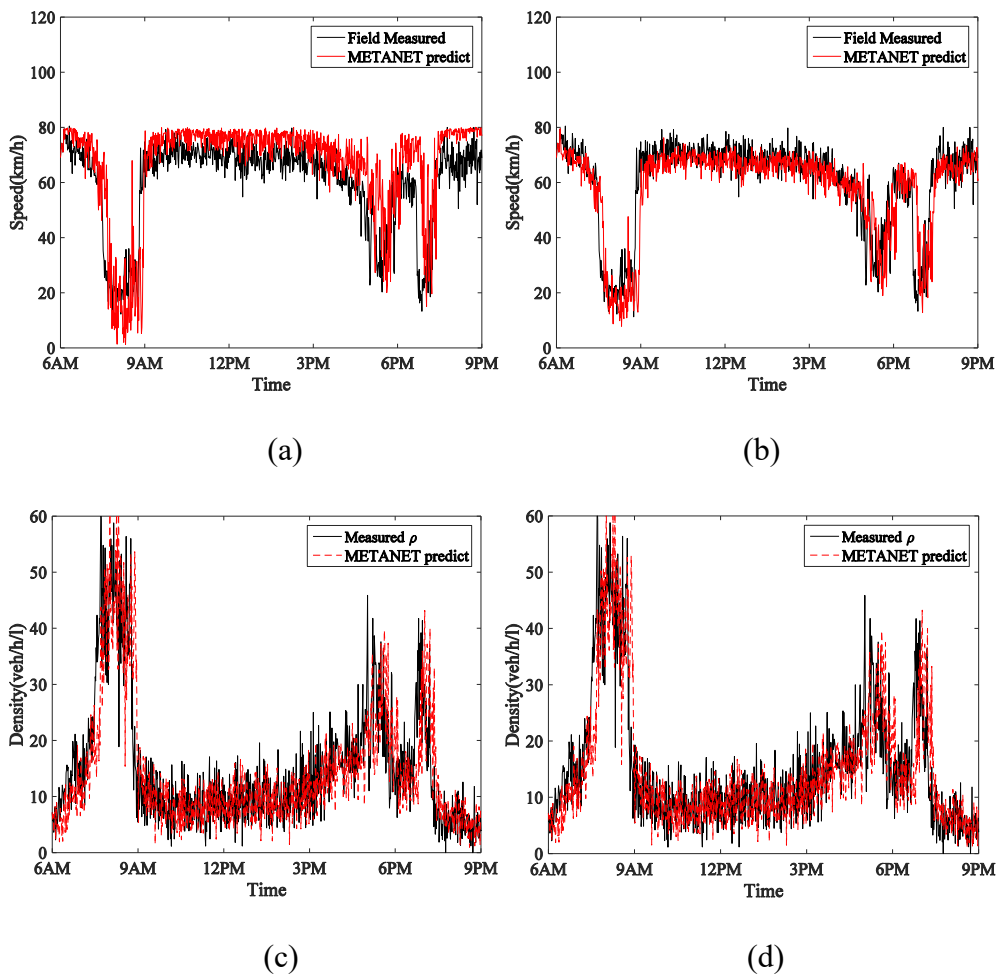


Figure 11 **Speed prediction (a, b) and density prediction (c, d) accuracy on one heavy snow condition day (Nov.18) (a) (c)** $v_{f,i}$ **= 80.06;** $\rho_{cr,i}$ **= 23.83. (b) (d)** $v_{f,i}$ **= 66.50;** $\rho_{cr,i}$ **= 21.46.**

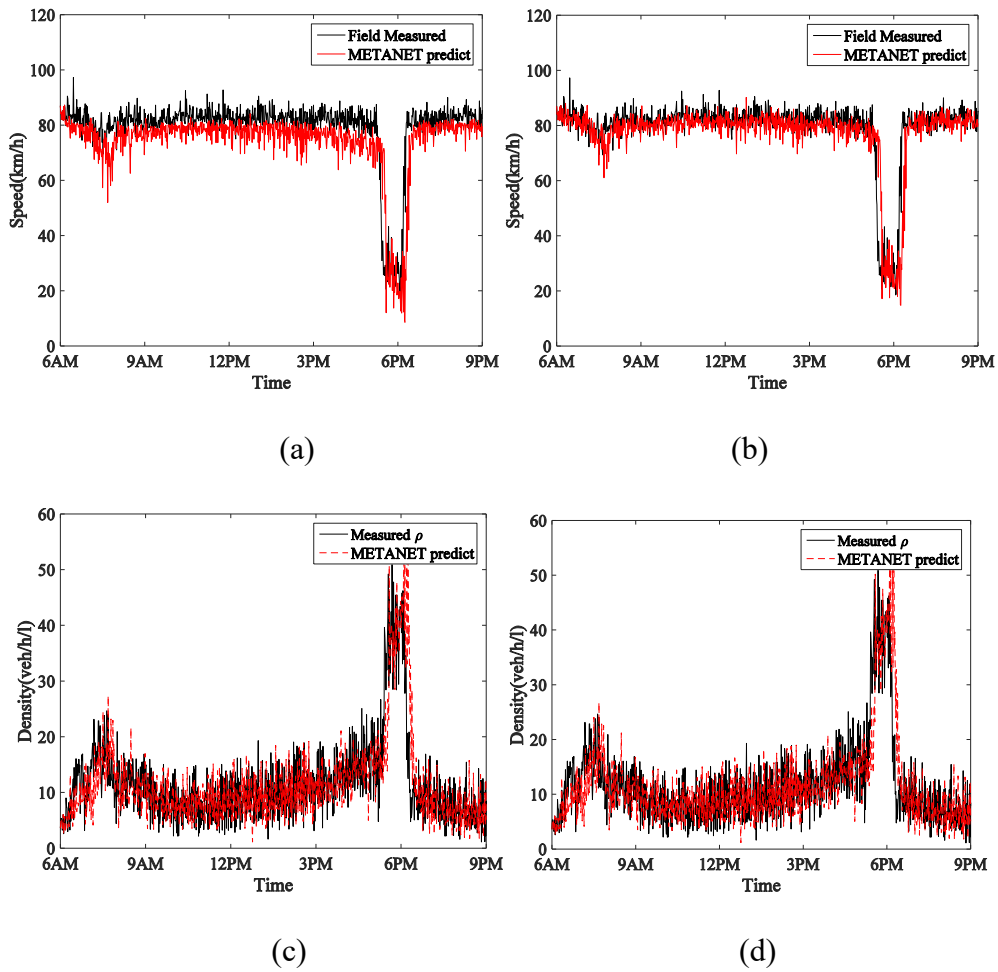


Figure 12 **Speed prediction (a, b) and density prediction (c, d) accuracy on few snow condition day (Nov.14) (a) (c)** $v_{f,i}$ **= 80.06;** $\rho_{cr,i}$ **= 23.83. (b) (d)** $v_{f,i}$ **= 81.08;** $\rho_{cr,i}$ **= 22.70.**

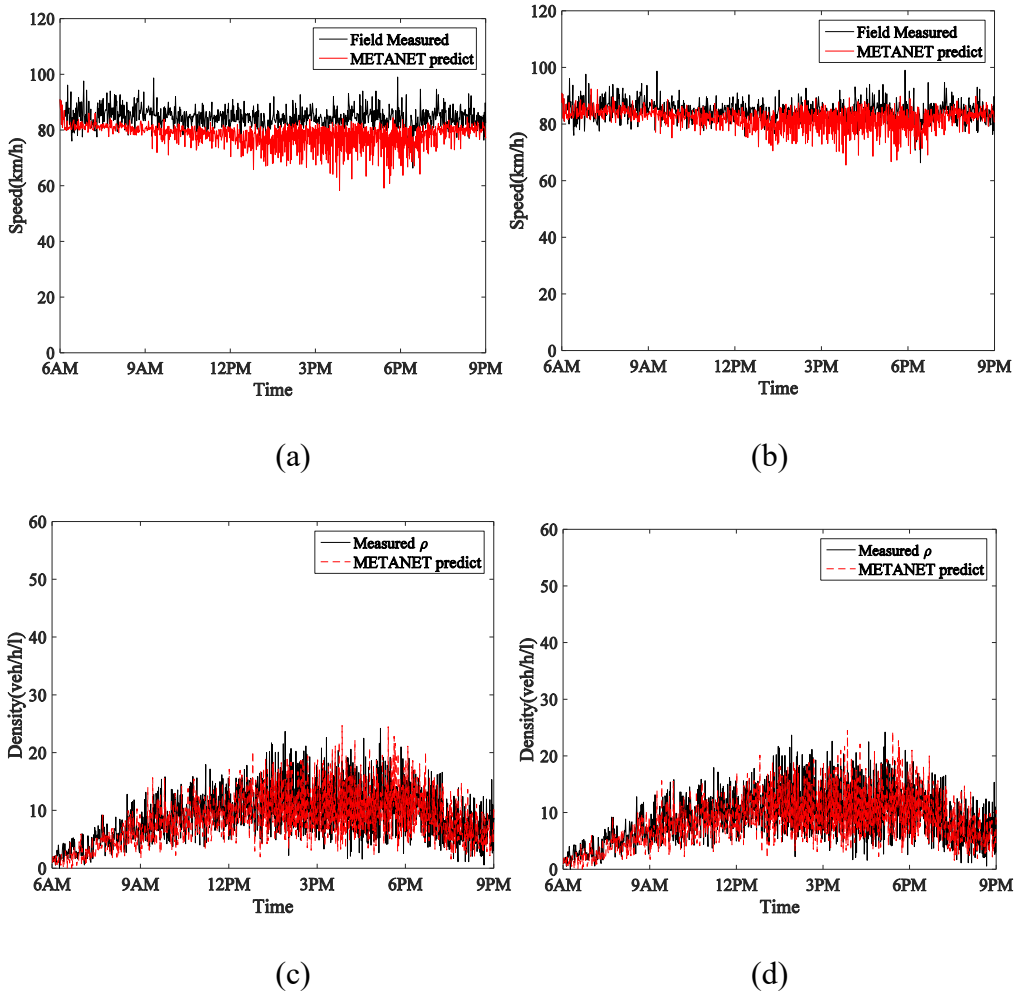


Figure 13 **Speed prediction (a, b) and density prediction (c, d) accuracy on one good weather condition day (May 02) (a) (c)** $v_{f,i}$ **= 80.06;** $\rho_{cr,i}$ **= 23.83. (b) (d)** $v_{f,i}$ **= 84.06;** $\rho_{cr,i}$ **= 23.83.**

Table 10 Quantitative Traffic Prediction Performances: Speed

Feature: RMSE of speed prediction		*Good weather condition*	*Light snow condition*	*Heavy snow condition*
Daytime Period *6AM-9PM*	Conventional	6.88	9.12	11.69
	Weather specific	**5.32**	**8.21**	**9.52**
AM Peak Hours *7AM-9PM*	Conventional	6.47	8.35	16.21
	Weather specific	**4.54**	**6.81**	**14.96**
PM Peak Hours *4PM-7PM*	Conventional	9.39	15.89	17.43
	Weather specific	**6.45**	**15.29**	**13.51**

Table 11 Quantitative Traffic Prediction Performances: Density (Daytime)

Feature: RMSE of density prediction	*Good weather condition*	*Light snow condition*	*Heavy snow condition*
Conventional prediction	4.46	6.02	7.17
Weather specific prediction	**4.45**	**5.98**	**7.14**

From all figures, Table 10 and Table 11 is it found that in speed prediction simulation observable difference can be seen while the difference in density

prediction is not visible. The reason is that the speed dynamics model heavily relies on the parameters while density dynamics is simply derived from mass conservation. Under heavy snow condition, the whole daytime speed RMSE of conventional prediction is 11.69 while in weather specific prediction the RMSE is 9.52, especially in PM peak hours the RMSE dropped from 17.43 to 13.51 which is more drastic than daytime average, and so does AM peak hours. Under light snow weather and good weather condition the whole daytime speed RMSE of conventional prediction is 9.12 while in weather specific prediction the RMSE is 8.21, especially in AM peak hours the RMSE dropped from 8.35 to 6.81 which is more drastic than daytime average and PM peak hours. Under good weather condition the whole daytime speed RMSE of conventional prediction is 6.88 while in weather specific prediction the RMSE is 5.32, especially in PM peak hours the RMSE dropped from 9.39 to 6.45 which is more drastic than daytime average and AM peak hours, however under good weather condition the original speed RMSE is small enough. In terms of density prediction, due to the nature of density dynamics described above, the convention prediction and weather specific prediction did not show significant difference. The RMSE difference ranges from 0.01~0.03 may be contributed to random computation error. And thus the prediction error of flow is proportional to the prediction error of speed. It can be concluded that the weather specific METANET prediction will be more helpful under the most adverse weather condition and during peak hours when speed drop violently.

4.5 Summary of Modifying Traffic Prediction Model

This part proved that weather conditions indeed impact driving environment and driver behavior that lead to the significant change of fundamental diagram, so that it is necessary to consider about weather specific fundamental diagram. If go further into traffic state prediction, we may conclude that this modification of METANET was successful in improving prediction accuracy.

The same weather variable filtering method can be applied to other cases if researchers have accessibility to higher resolution weather data such as hourly weather data or there are other weather events such as rain and fog. Using the same method, researchers will find out different significant variables that fit their situations. Through prediction simulation it is found that macroscopic traffic prediction accuracy is enhanced after introducing weather specific fundamental diagram parameters especially in speed prediction. Given global parameters previously calibrated and fixed, in adverse weather conditions weather factors perform better than favorable weather conditions. And that the same "desire speed" term modification technology can be applied to other application such as variable speed limit control.

CHAPTER 5. THE IMPLEMENTATION AND EVALUATION OF MPC-VSL FIELD TEST

This chapter gives an analysis of VSL field test performance. In this part, the time response of VSL speed curve is drawn and time domain analysis is done. Moreover, TTT and TTD prediction performance is analyzed.

5.1 Introduction and Background

The effectiveness of variable speed limits (VSL) in improving traffic throughput and saving travel time for urban freeway users is crucial for future traffic control operation. This paper models and examines the effectiveness of VSL control based on field test data collected from an urban freeway testbed in Edmonton, Canada. The VSL control algorithm is model predictive control (MPC)–based, in which the traffic state prediction model is a modified METANET model. The desire speed term in the METANET model is modified when VSL control is triggered. The value of VSL in each round of calculation is decided by a discrete choice model–based optimizer, and its objective is to minimize system total travel time (TTT) while maximizing total travel distance (TTD). The TTT and TTD compose the measure of effectiveness (MOE) of the MPC-VSL field test. This paper conducts a step response analysis of how VSL reacts to field speed reduction and evaluates the MOE of the MPC-VSL field test. In the evaluation of the MOE, the discrepancy between the predicted MOE and actual MOE, namely

the system uncertainty is attributed to system parameters and different VSL scenarios. This study demonstrates the performance of MPC-VSL field test, and the proposed analyzing frame serves future traffic control practices.

Posted speed limits are an essential measure for speed regulation, helping to reduce traffic accidents and improve traffic throughput. Since static speed limits are designed for ideal road conditions, they may not be suitable during adverse weather conditions or congestion. Variable speed limits (VSL) have been proposed as a means to recommend safe driving speeds during less ideal conditions. Implementation of VSL has been successful in European countries and the United States. [54] In Edmonton, Canada, a four-week VSL field test was implemented on a 10-kilometer stretch of freeway. This paper analyzes the performance of the VSL test, and conducts analyses of system stability robustness under the scheme of model predictive control. The results show the effectiveness of our VSL algorithm. This study provides evaluating scheme for future field implementation.

5.1.1 Description of Variable Speed Limit Testbed

The Vehicle Detection Stations (VDS) on a 10-kilometer corridor of Whitemud Drive in Edmonton, Canada, collect and store traffic data from dual loop detectors. This section of road plays an important role in people and freight transportation in the city. The VDS system currently has 28 VDS in total, and each station has three or four dual loops. The data recording frequency is 20 seconds. Each dual loop reports the volume q—the number of vehicles crossing the loop detector during a 20-second time interval—and mean speed measurement v, as well as the

occupancy measurement, which cannot be used directly in traffic control and cannot be transformed accurately into density. The accurate density measurement ρ is calculated by $\rho = q / v$. Figure 14 shows the testbed geometry, VDS locations and the location of DMS (Dynamic Message Signs) that will be used to display variable speed limits.

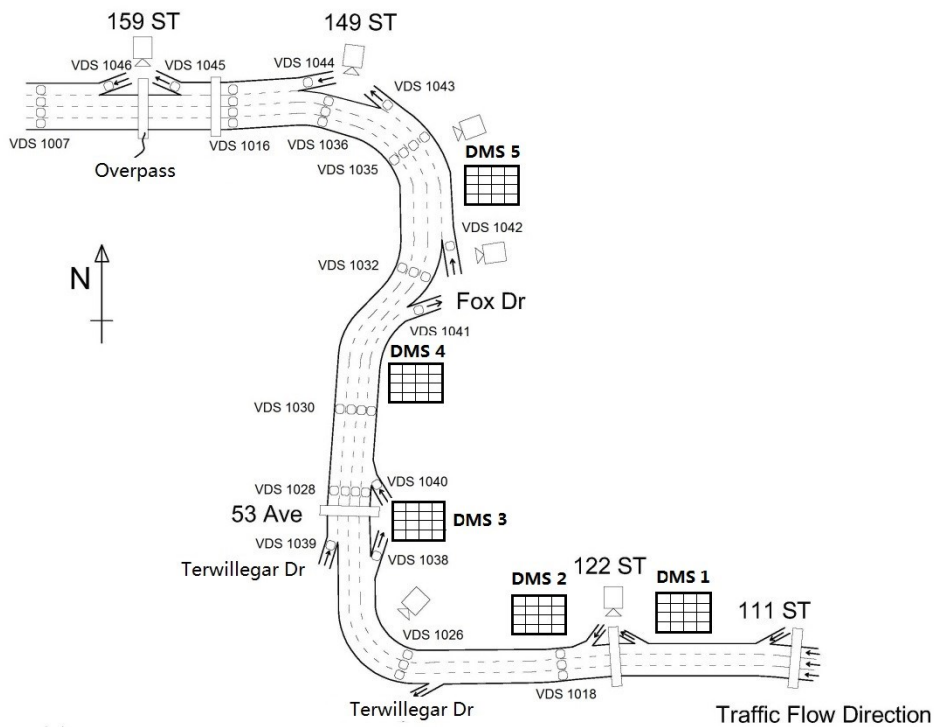


Figure 14 **VDS and DMS location on westbound testbed of Whitemud Drive**

The VSL control decisions are made and sent out by software called DynaTAM (Dynamic Analysis Tool for Active Traffic Demand Management). During real-time VSL control, data from loop detector stations is sent to computers in the traffic management center in the City of Edmonton. The DynaTAM software installed in the computers is responsible for providing speed limit suggestions. The decisions are then shown on screen with beeping

notifications for human operators to check. After the operators check the video cameras at corresponding downstream locations and confirm that the speed limit needs to be changed, the suggested speed limit values are then put into DMSs along the road. From August 13 to September 4 of 2015 on weekdays, the City of Edmonton conducted a four-week VSL pilot test on the 10-kilometer testbed of westbound Whitemud Drive. There were 5 DMSs in total located upstream of predefined congestion-prone points. Figure 2 demonstrates the interface of DynaTAM and the actual VSL display used in the field test. Due to legal reasons, the variable speed limits were not mandatory to drivers, and they were labeled as "Advisory Speed."

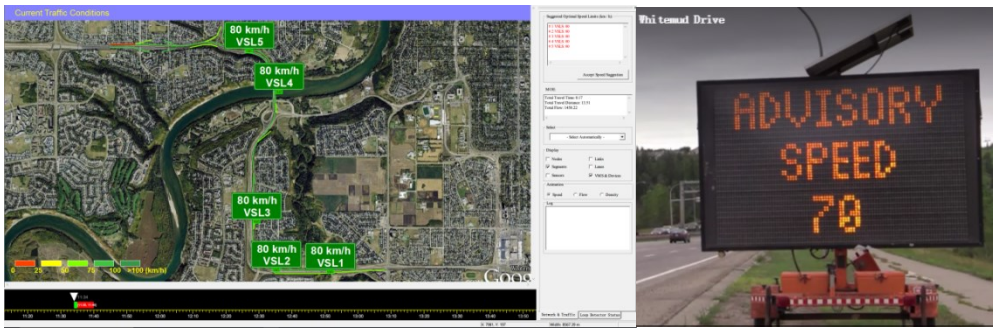


Figure 15 **DynaTAM interface and DMS on westbound testbed of Whitemud Drive**

5.2 Model Prediction Control Based VSL Algorithm

The model predictive control (MPC)–based VSL algorithm is composed of five parts: an original traffic prediction model, control case traffic prediction model, optimizer, objective and constraints. As shown in Figure 3, for VSL control road

segments, the system uses a modified METANET model for traffic state prediction, and the discrete choice optimizer makes a choice based on predicted traffic states given objective functions and constraints. After optimization, the optimal VSL choice is shown to operators, and at the same time, the VSL value is inputted to the VSL-METANET model again for the final traffic state prediction results to be recorded. Figure 3 shows a block diagram of the MPC-VSL control system.

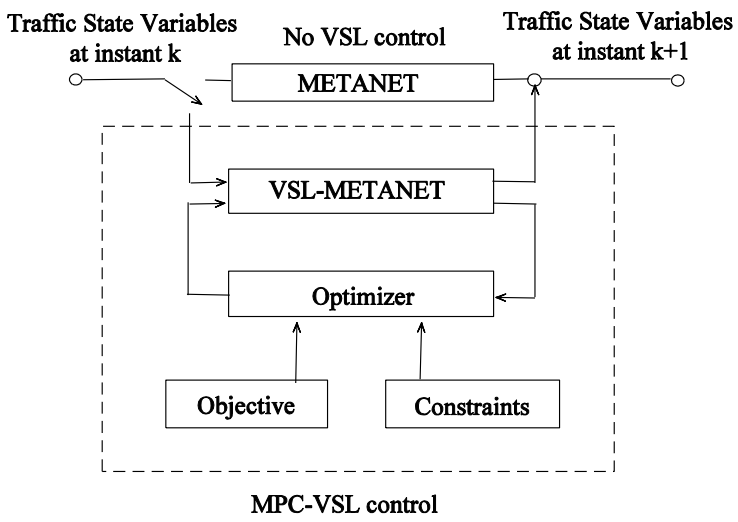


Figure 16 **Block Diagram of MPC-VSL control**

5.2.1 METANET Prediction Model and Modification

The METANET model as a second-order traffic flow model introduces speed dynamic functions instead of only including density dynamics to describe flow conservation law. This feature of the METANET model enables it to predict density, speed and flow variables respectively and accurately, as well as with a

small time interval. The discretized model formulation of the original METANET is written as follows:

Density Dynamics

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{\Delta_i \lambda_i}\left[\lambda_{i-1} q_{i-1}(k) - \lambda_i q_i(k) + r_i(k) - s_i(k)\right] \tag{5-1}$$

Speed Dynamics

$$v_i(k+1) = v_i(k) + \frac{T}{\tau}\left[V\left(\rho_i(k)\right) - v_i(k)\right] + \frac{T}{\Delta_i} v_i(k)\left[v_{i-1}(k) - v_i(k)\right] - \frac{\eta \cdot T\left[\rho_{i-1}(k) - \rho_i(k)\right]}{\tau \cdot \Delta_i \cdot \left[\rho_i(k) + \kappa\right]} \tag{5-2}$$

Flow Dynamics

$$q_i\left(k+1\right) = \rho_i\left(k+1\right) \cdot v_i\left(k+1\right) \tag{5-3}$$

Where, i is the index of links, and $i = 1, 2, ...M$, with M representing the number of testbed sections. K is the index of time instants, and T is the calculating time interval, where $T=20s$. Δ_i represents the segment length of link i, and λ is the lane number of link i. In the model, τ, κ, α and η (km^2/h) are global model parameters calibrated using the historical data. The desired speed $V\left(\rho_i\left(k\right)\right)$ (km/h) in speed dynamics is represented by the following expression:

$$V\left(\rho_i\left(k\right)\right) = v_{f,i} \exp\left[-\frac{1}{\alpha}\left(\frac{\rho_i\left(k\right)}{\rho_{c,i}}\right)^{\alpha}\right] \tag{5-4}$$

The triangular fundamental diagram (FD) is assumed in the control algorithm. In equation (5-4) $v_{f,i}$ means free flow speed and that $\rho_{c,i}$ represents for critical density

which is the density associated with capacity. In this model density dynamics does not involve parameters and other parameters in speed dynamics are calibrated globally for goodness of fit. Note that of the four terms making up the speed dynamics, and each of terms has physical meaning. In equation (5-2), the second term is referred to as the relaxation term, describing that with a lag time item τ, the mean speed v of the link gets relaxed to the desired speed which largely depends on parameters of FD. The selection of the desired speed is critical to reflect the driver behavior and from previous practice we chose the format of equation (5-4). The third is the convection term meaning that vehicles travelling from upstream link $i-1$ to current link i gradually adapt their speed rather than instantaneously. The fourth is anticipation term meaning that drivers are always keeping an eye on the traffic condition ahead. If a driver observes high traffic density in the downstream link $i+1$, he then reacts as slowing down, and vice versa. The constant $\kappa > 0$ is added to keep the anticipation term limited when density is low.

The METANET algorithm for the VSL control environment is modified by replacing desired speed with a modified one. In the control segments, the desired speed is assumed to be the posted advisory speed rather than the speed determined by the fundamental diagram as in the non-control case. In the VSL-METANET model, the desired speed term then becomes the optimized variable speed limit value, as decided by the optimizer in the MPC system.

Modified Desire Speed

$$V\left(\rho_i\left(k\right)\right) = u_i(k) \tag{5-5}$$

The Courant–Friedichs–Lewy (CFL) condition is followed, which means to grant that vehicles cannot travel beyond one link within computing time interval T, so that T satisfies the following condition:

$$T \le \frac{\Delta_i}{\max\left(v_{f,i}\right)}, i = 1, 2, ..., N \tag{5-6}$$

5.2.2 Object Function and Constraints

The objective function minimizes a weighted sum of TTT and TTD in a discrete choice fashion. In terms of definition, total travel time indicates a weighted sum of density of all time steps and links. Only minimizing TTT produces system bias wherein the optimizer tends to choose a lower speed limit to decrease the density, so that in this system TTD is maximized simultaneously. In the objective function, weights α_{TTT} and α_{TTD} are applied to TTT and TTD.

$$TTT = T\sum_{j=1}^{pr-1}\sum_{i=1}^{M}\lambda_i L_i \rho_i(k+j) \tag{5-7}$$

$$TTD = T\sum_{j=1}^{pr-1}\sum_{i=1}^{M}\lambda_i L_i \rho_i(k+j) v_i(k+j) \tag{5-8}$$

$$J = T\sum_{j=1}^{pr-1}\sum_{i=1}^{M}\lambda_i L_i \left[\alpha_{TTT}\rho_i(k+j) - \alpha_{TTD}\rho_i(k+j)v_i(k+j)\right] \tag{5-9}$$

The desired speed within the VSL environment is then decided by the optimizer. The optimization is subjected to several constraints of traffic safety, driver acceptance and traffic flow characteristics. [39] The first constraint is the upper limit of $u_i(k)$; since the regular speed limit is 80km/h, we consider that

under congested conditions $u_i(k) \le 80$. The second constraint is to maintain continuous traffic flow even under congestion. From local regulations, the lower bound of the speed limit is 30km/h, which means $u_i(k) \ge 30$. In Canada, the speed limit is in multiples of 10km/h, and in our system, to ensure safety the change in speed limit between two time instants is 0 or 10km/h, that is to say $u_i(k) - u_i(k+1) \in \{-10, 0, 10\}$. Given the above constraints, the optimizer becomes a discrete choice model described as follows.

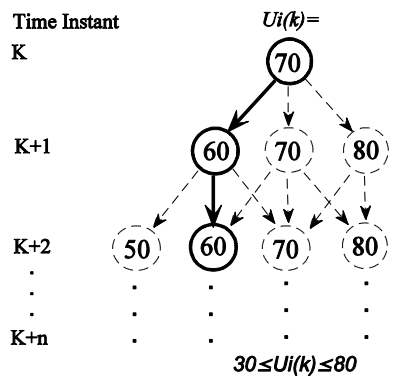


Figure 17 **MPC-VSL Optimizer Decision Tree**

5.3 Time Domain Analysis of Speed Control

The concept of time domain analysis is borrowed from automatic control. It is usually used to describe the time response of a system, and more specifically, how the input signal oscillates until stable given all theoretical control algorithms. Time domain analysis is usually conducted in pure theoretical calculation and simulation. In this study, we use time domain analysis concepts to describe the time response characteristics of the VSL as it responds to a drop in speed downstream at a congestion location. In Figure 5, the VSL line is observed to

capture the drop trend of the average speed for the corresponding location with little delay, as well as detect when the field average speed begins to recover.

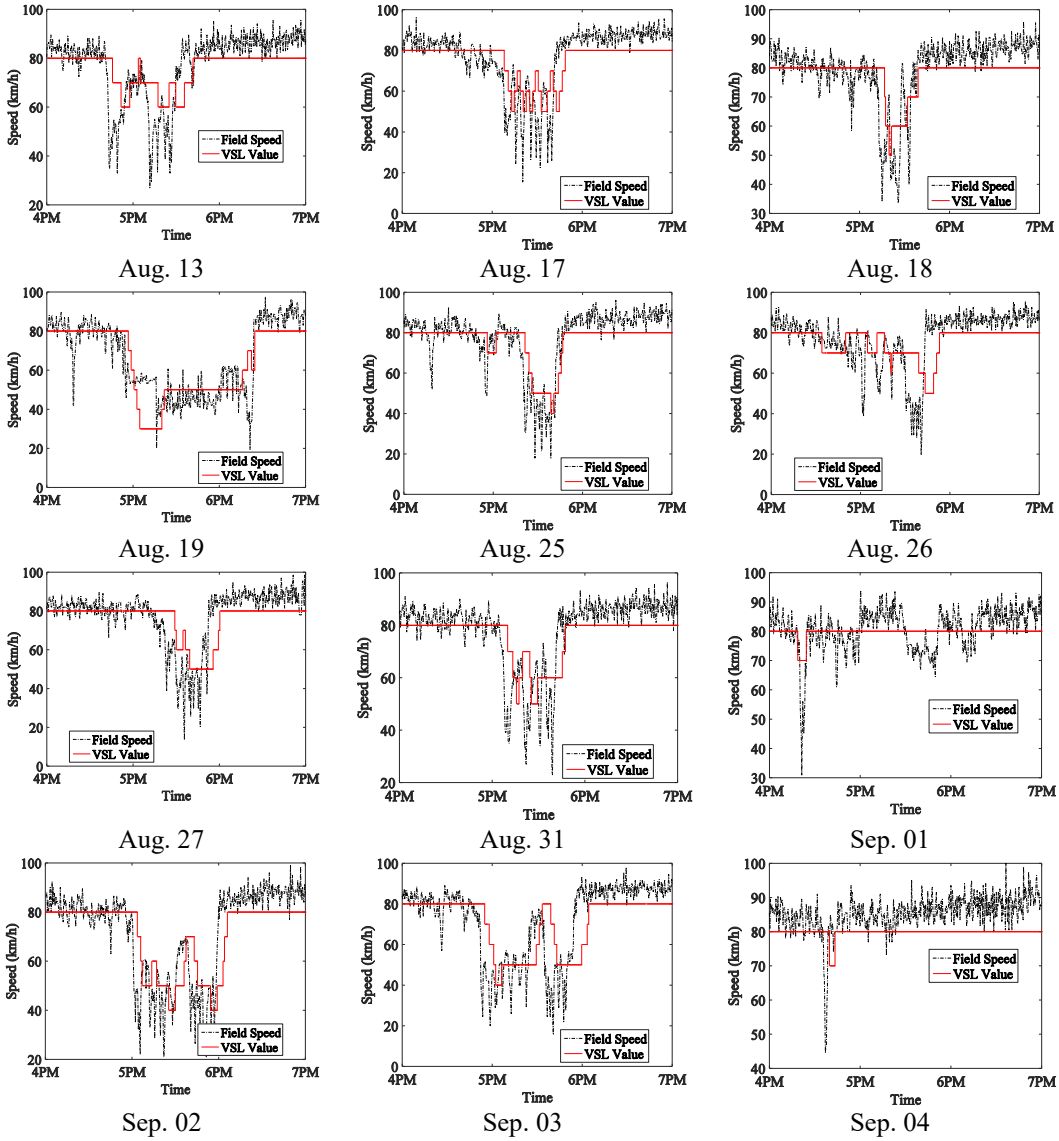


Figure 18 **Speed Profiles of Valid Field Test Days near DMS 1**

In the control system design language, there are certain quantities involved that associate with the step response of the system. The MPC-VSL control is more complex than a basic control system that besides standard quantities rise time, settling time, overshoot and peak time, latency time is added to better analyze the sensitivity of system response.

t_l =the latency time, which is the time lag between a drastic speed drop and VSL triggering, measured in minutes.

t_r = the rise time, which is the time it takes for the VSL control system to reach its extremum, measured in minutes.

t_s = the settling time, which is the time it takes for VSL control system transients to decay, measured in minutes.

$M_p = \frac{Maximun \quad Overshoot}{Final \quad Value}$. Overshoot in this paper refers to how much the VSL value deviates from 80km/h, measured as a positive number.

t_p =the peak time, which is the time it takes for the VSL system to reach the maximum overshoot point, measured in time instants.

In reality the VSL response curve includes more than one oscillation; the following table shows results for the first wave of VSL drop and recovery. The first oscillation indicates the capability of the system to be sensitive to traffic conditions and react correctly.

Table 12 Quantities of Step Response of MPC-VSL Control System

TEST DAY	*Speed Drop Began*	*VSL Control Began*	t_l	t_r	t_s	M_p	t_p
Aug. 13	4:41:20 PM	4:45:40 PM	4.33	6.00	12.00	25%	4:51:40 PM
Aug. 17	5:03:00 PM	5:08:00 PM	5.00	5.00	3.67	38%	5:13:00 PM
Aug. 18	5:12:00 PM	5:16:40 PM	4.67	3.00	20.00	38%	5:19:40 PM
Aug. 19	4:54:40 PM	4:56:40 PM	2.00	8.00	17.00	63%	5:04:40 PM
Aug. 25	5:17:20 PM	5:21:20 PM	4.00	17.33	8.00	50%	5:38:40 PM
Aug. 26	5:35:00 PM	5:39:00 PM	4.00	2.00	5.00	38%	5:41:00 PM
Aug. 27	5:22:20 PM	5:29:00 PM	6.67	1.00	4.67	38%	5:30:00 PM
Aug. 31	5:05:00 PM	5:10:00 PM	5.00	5.67	4.00	38%	5:15:40 PM
Sep. 01	4:18:20 PM	4:18:40 PM	0.33	2.67	3.00	13%	4:21:20 PM
Sep. 02	5:00:00 PM	5:03:40 PM	3.67	3.33	6.67	50%	5:07:00 PM
Sep. 03	4:50:20 PM	4:55:00 PM	4.67	7.33	31.00	50%	5:02:20 PM
Sep. o4	4:35:40 PM	4:39:40 PM	4.00	1.67	2.00	13%	4:41:20 PM

Typically those system step response quantities in Table 1 are equations that can be solved to meet certain requirements. In this case in contrast those quantities are constants to be evaluated empirically. Since no analyses on VSL field test system response exist for comparison, engineering experience is used to analyze the performance of the VSL step response. In this test, latency time ranges from 0.33 to 6.67 minutes, which we suggest is acceptable. The rise time is largely decided by the field condition, which was hard to judge, except for August

25, where the rise time falls between 1.00 and 8.00 minutes. The settling time reflects the length of time between the lowest VSL point and its first point that is nearest to 80km/h. If we consider the overshoot together with the settling time, the finding is that the bigger percentage of overshoot, the relatively longer settling time to take.

5.4 Analysis of Measure of Effectiveness

The measures of effectiveness (MOE) of MPC-VSL in this study are TTT and TTD, which are also components of the control objective function. In this section, a block diagram and an error transfer function are formulated to analyze the source of uncertainties for TTT and TTD.

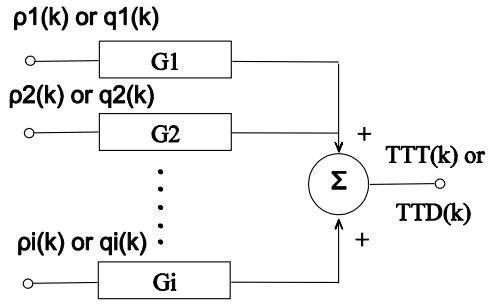


Figure 19 **Block Diagram of TTT and TTD**

In the above block diagram, at each time step, the TTT and TTD are calculated from a weighted sum of the density and flow of all links, in which the weight is the segment length of each link. The discrepancies of TTT and TTD, between the predicted and field-collected values of each, are attributed to the prediction error of density and flow, the segment length measurement error and the inaccuracy of the prediction model. To further explain, we borrow the concept

of a transfer function from the automatic control field, which is typically used to describe the transition of statues of system. Here, we use this concept to describe the discrepancy between the anticipated performance and real performance.

The error transfer function

$$\Phi_{MOE} = \omega_1 G1 + \omega_2 G2 + ... + \omega_9 G9 \tag{5-10}$$

Where, $\omega_1, \omega_2 .....\omega_9$ represent the prediction error of link 1 to link 9 from the METANET model, and $G1, G2.....G9$ represent the mismatch of segment length, segment division, prediction steps and other system fixed-parameter errors.

In this system, the discrepancies between the actual MOE and those predicted via VSL-METANET fall under two types of uncertainties related to the prediction model: structured uncertainties and unstructured uncertainties. The data utilized in this analysis is from loop detectors and DynaTAM for the day of August 17, 2015, during 4PM-7PM peak hours. All traffic state data and VSL data are aggregated at each minute. The data is from one VDS that corresponds to DMS 1 in the field, since on that day only DMS 1 reacted in response to one bottleneck that was activated.

Structured Uncertainties

In the following figure, Structured Uncertainty (StrU) means the difference between model predicted and adjusted model predicted MOE. This category of uncertainty came from $G1, G2.....G9$, and global parameters of METANET which contribute to part of $\omega_1, \omega_2 .....\omega_9$. The adjustment strategy will be presented after figure.

Unstructured Uncertainties

Unstructured Uncertainty (UU) means the difference between adjusted model predicted and field measured MOE. This uncertainty came from the accuracy of METANET and the choice of optimized VSL.

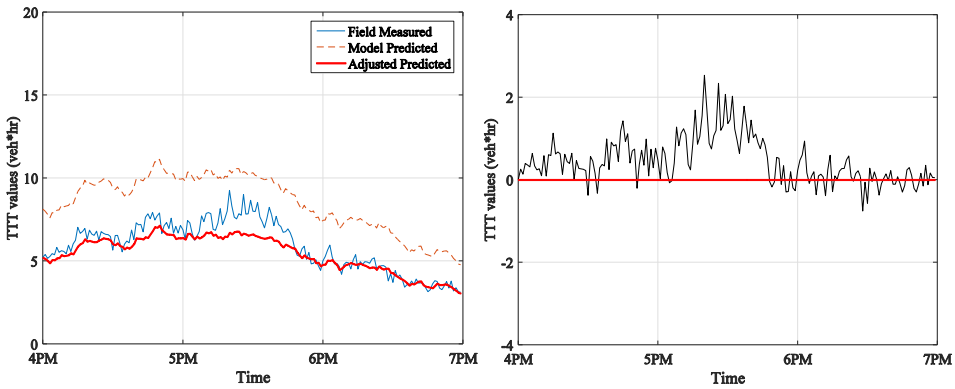


Figure 20 **(a) Comparing field measured TTT, model predicted TTT and adjusted model predicted TTT. (b) Model prediction error using adjusted prediction results.**

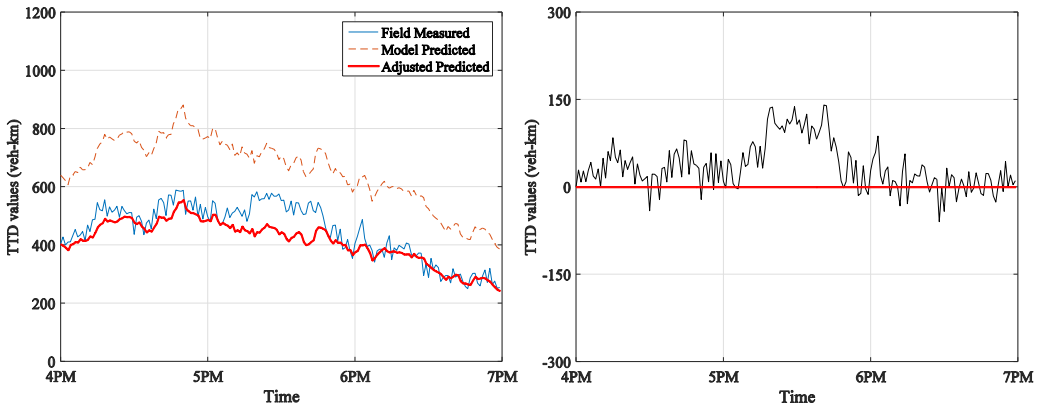


Figure 21 **(a) Comparing field measured TTD, model predicted TTD and adjusted model predicted TTD. (b) Model prediction error using adjusted prediction results.**

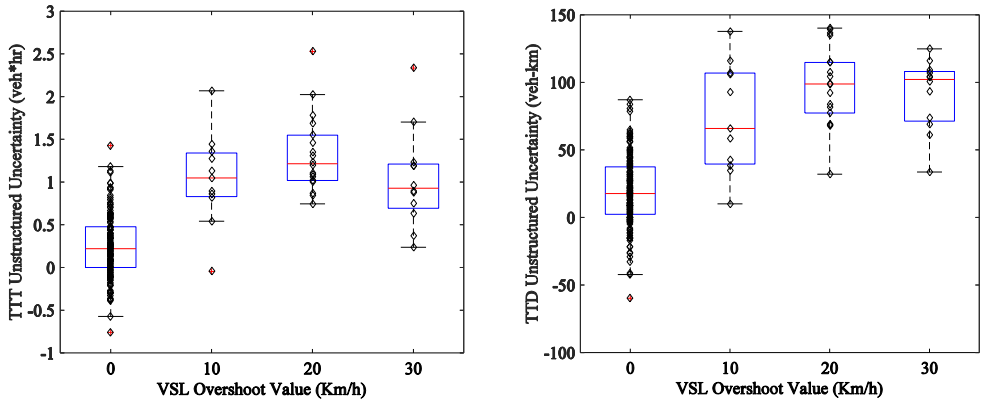
The structured uncertainty of predicted MOE is assumed to be proportional to original predicted MOE, after trying other assumptions such as

proportion to measured MOE, fixed value of uncertainty, and finally the denial of assuming normal distribution. So that the adjusted predicted MOE is calculated by scaling down original predicted MOE, in which the reference point is the beginning of the study period 4:00:00PM, meaning the suggestion that at the start point of this time period no structured prediction uncertainty existed. Based on the above understanding, although the exact values of $\omega_1, \omega_2 .....\omega_9$ and $G1, G2.....G9$ are not known, the overall structured uncertainty can be calculated. After eliminating structured uncertainty by scaling down original predicted MOE, there remains unstructured uncertainty. Figure 21 (b) and (d) show unstructured uncertainty. The expression of structured and unstructured uncertainties of the MOE can be expressed as follows:

$$StrU = \left(1 - \frac{Moe_m(t')}{Moe_p(t')}\Big|_{t'=4:00:00PM}\right) \cdot Moe_p(t) \tag{5-11}$$

$$UU = Moe_m(t) - Moe_p(t) \cdot \frac{Moe_m(t')}{Moe_p(t')}\Big|_{t'=4:00:00PM} \tag{5-12}$$

Where, $Moe_m(t)$ and $Moe_p(t)$ represent field-measured and DynaTAM software predicted TTT and TTD respectively, at time step t before adjustment. From Figure 21 (b) and (d), it can be observed that approximately between 5PM and 6PM, when VSL control was active, the UUs were unnaturally high, and that the UU fluctuated around zero during other times. We then conducted further data mining aimed at finding the quantitative relationship between detailed VSL overshoot values and UU of TTT and TTD during 4PM-7PM.



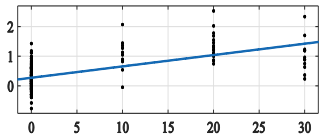
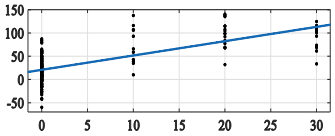
(a) (b)

Figure 22 (a) Boxplot of Unstructured Uncertainty of TTT over VSL Overshoot Value. (b) Boxplot of Unstructured Uncertainty of TTD over VSL Overshoot Value.

In Figure 22, boxplot and original UU data points are overlaid since VSL overshoots are discontinuous. There is positive correlation between VSL overshoot and the value of UU. We assumed that the larger the VSL overshoot is, the lower the $u_i(k)$ in the speed prediction function, the lower of predicted traffic state variables. Theoretically, if drivers comply 100% with speed limit, the MOE prediction accuracy remains the same in different VSL scenarios. However, when VSL is low, the road condition becomes complex and that the compliance rate dropped, indicating that drivers keep driving faster than VSL that displayed. Due to the above reason, the adjusted-predicted MOE is lower than field measured in low VSL scenarios. After the low VSL period, the positive UU decays to zero again. Figure 8 shows the result of mining UU data of MOE, where it is clear that the mean of UU of each box increases with the increase of VSL overshoot. Each

box has similar and small variance, and except for the scenario with zero VSL overshoot, other VSL scenarios have much fewer data points.

Table 13 Statistics of Unstructured Uncertainty of MOE versus VSL Overshoot

Unstructured Uncertainty(UU)	*TTT*				*TTD*			
VSL Overshoot	0	10	20	30	0	10	20	30
Sample N=	139	11	18	12	139	11	18	12
Mean of UU	0.25	1.04	1.33	1.03	19.2	73.7	97.99	91.3
Linear Fit of VSL Overshoot (VSLO) and UU	 UU = p1*VSLO + p2 with 95% confidence bounds, p1 = 0.03826 (0.03121, 0.04532) p2 = 0.2787 (0.206, 0.3514)				 UU = p1*VSLO + p2 with 95% confidence bounds, p1 = 3.079 (2.611, 3.546) p2 = 21.04 (16.23, 25.86)			

Detailed statistical analysis was conducted, and the results are shown in Table 13. Besides providing specific values of classified sample numbers and mean values of UU for both TTT and TTD, a linear fit is used to check the correctness of the positive correlation between VSL overshoot and UU, or in other words, to check the significance of the parameters. The R-square values are not

shown because the data is distributed in four small clusters, which results in a naturally lower R-square, and both R-square are around 0.5. For both TTT and TTD, the parameters of variable VSLO fall between the 95% confidence bounds, which indicate that we have 95% confidence in the VSL overshoot value being closely correlated with the UU of MOE.

5.5 Summary of MPC-VSL Field Test

From the field implementation of MPC-VSL control, firstly the step response analyzing scheme proves to be suitable for evaluating the sensitivity of control system. The analyzing result is that the system is quick and reasonable in terms of speed reaction. Once traffic broke down, or a speed drop occurred, the loop detector immediately upstream of the bottleneck promptly detected that change and successfully reacted with a latency ranging 0.33-6.67 minutes (20-400 seconds); on average, the latency time was around 2-3 minutes. Even without clear criteria to judge the promptness of reaction, this latency is considered acceptable for humans. Moreover, with the time response curve, it can be observed that VSL signals help stabilize the fluctuation of the speed curve, and in some cases, helps with speed recovery. Secondly, the analysis of MOE shows that the MPC-VSL algorithm performs robustly in the field test. The StrU is stable and eliminable. The UU is small in value compared to StrU, and that the UU is able to converge to zero after peak time. The value of UU is positively and closely related to the deviation of VSL values from the regular speed limit 80km/h, and that the UU can be attributed to low compliance rate and chaotic traffic condition, when VSL is low and the road get congested.

The step response analyzing tool and system uncertainty attributing method proposed in this paper provide a framework for evaluating traffic control systems as well as mining traffic flow data in non-control and control cases. For future improvements, considering improving field test performance, the manual check of software provided VSL decisions can be simplified for saving latency time, and the StrU can be eliminated before field implementation. Considering studies about VSL evaluating methods, the step response analyzing method needs to be generalized, and the uncertainty analyzing method should take more factors into consideration. Considering improving MPC-VSL algorithm, the driver compliance rate should be taken into account when optimized VSL value is too low to comply. When traffic is extremely congested, an alternative optimizer may be needed.

CHAPTER 6. CONCLUSION AND DISCUSSIONS

This chapter presents a general conclusion of the thesis and discusses about limitations of this thesis as well as future studies.

6.1 General Conclusions

This thesis is concerned with investigating the problems about field implementation of MPC-VSL traffic control, and the evaluation of the first pilot field test of MPC-VSL control in Edmonton, Canada.

Chapter 3 investigated a real time missing loop detector data imputation method which is specifically designed for the data missing of a lane for a continuous period of time. This type of data missing was observed in loop detector system in Edmonton as well as other disciplines, and this type of missing data significantly impact the performance of real time traffic control such as the MPC-VSL control in this thesis. The imputation algorithm is imputing the missing lane utilizing the information of its neighbor lanes in the same loop detector station via multiple linear regression. The algorithm is based on the assumption that all lanes in the same location are homogeneous in terms of road geometry and traffic volume. The results show that the proposed multiple linear regression method outperforms other commonly used methods and is more convenient to be put into practice.

Chapter 4 investigated the feasibility of modifying one critical term called "desire speed" in METANET prediction model for better prediction accuracy. In this case study, weather factors are introduced into the desire speed term to improve prediction accuracy under unsatisfying weather conditions. The mathematical form of weather factor is linear in terms of weather variables, and that the key weather variables are filtered using real weather and loop detector data of November 2013. It is found that the difference between conventional fixed parameter prediction and weather specific prediction is larger when weather condition is worse. This is proved that desire speed term is critical in representing driver's expected speed, and this term is reasonable to be modified for other uses such as MPC-VSL control that will be described in next chapter.

Chapter 5 first describes the scheme of MPC-VSL system, and then evaluates the performance of field test that conducted in Edmonton. The field feedback shows that the MPC-VSL system reacts quickly with the speed drop of bottle neck area, and is able to catch the trend of speed change. In terms of measure of effectiveness, both total travel time and total travel distance show discrepancy between model predicted values and accrual values. This discrepancy or model uncertainty can be classified into two kinds, structured uncertainty and unstructured uncertainty. The structured uncertainty is caused by fixed METANET parameters which were calibrated before and may not be suitable for all situations. Another source of structured uncertainty is inaccurate measured segment length. Fortunately the structured uncertainty can be eliminated since it is proportional to the original predicted total travel time and total travel distance. The unstructured uncertainty is caused by randomness of METANET model, the optimized variable speed limit value and so on. The unstructured uncertainty is found closely related to the VSL overshoot, which indicate that the lower VSL value is, the larger the unstructured uncertainty will be.

6.2 Limitation of this Thesis

The missing data imputation method proposed in this thesis has the following limitations. Firstly, it is suitable for lane missing type instead of random missing type which seems to be more common in sensor data collection. When random missing data imputation is needed, more complex statistics based imputation method should be used instead. Secondly, the proposed multiple linear regression technique is not granted to perform better than pairwise linear regression model in

the example of this thesis. However, multiple linear regression is more easy to apply into practice since when doing pairwise linear regression, the "pair" of lane should be defined before, and in multiple linear regression, the imputation is based on all other lanes in the same station. Thirdly, the performance of imputation relies heavily on the quality of training data. In practice it is better to choose training days that are closest to missing days.

When checking the feasibility of modifying desire speed term by showing case study of introducing weather factors, the limitation is the quality of weather data. The weather data frequency is one day, which is thought to be insufficient when weather changes drastically within one day. The weather variables in Edmonton are mostly related to snow, but in other countries of the world more weather variables such as rainy variables and foggy variables should be added and be filtered again.

In field test evaluation, the limitation lies in the following aspects. Firstly, the driver compliance rate was not measured during field test. The sensor corresponding to one specific DMS is on the bottleneck location while the DMS is located 500m upstream of the bottleneck. There is no sensor at the location of DMS to measure how people slow down responding to the variable speed limit. Secondly, the effectiveness of MPC-VSL is not directly measured since during the field test the before and after compare is not conduct. The compare requires that the system to be turned off for a while during test period. In this thesis the most direct indicator of the effectiveness of MPC-VSL is the speed profile.

Thirdly, the source of unstructured uncertainty is not fully investigated. The current finding is that this type of uncertainty is related to the VSL overshoot.

6.3 Recommendations for Future Researches

In the future, the loop detector missing data method can be expanded to be able to address more missing types. Imputation methods those are suitable for small scale random missing pattern, such as multiple imputation method and expectation maximization method. Currently the missing data imputation is an algorithm in research papers, to implement it into real time actual traffic control more programing work is needed. The multiple linear regression imputation method can be implemented to not only online traffic control, but also off line database that suffers from data missing.

The modification of the desire speed term in METANET model can serve different purposes. In the case study in Chapter 4, the modification goal is improving the model prediction accuracy under bad weather conditions. In the future if higher resolution weather data is available, the weather factor parameters can be more reliable and more accurate. What is more, different categories of weather variables should be included and modelled with real world data for building a complete weather factor modelling system that helps improving traffic state prediction models not only in winter cities such as Edmonton, but in cities all over the world. The desire speed can be modified to other uses such as different road geometry and different vehicle types if data is accessible. The modification of METANET model is not restricted to modifying desire speed. Speed dynamic as the only dynamic function in METANET that involve approximation and

parameter calibration, new terms may be introduced to improve the accuracy of speed prediction, and that requires the Payne model to be expanded again into higher orders.

Through this thesis, the major finding in MPC-VSL field test is that there are discrepancies between expected and actual measure of effectiveness, for both TTT and TTD. The expected measure of effectiveness under estimate the increase of TTT and TTD, and that under-estimate is correlated with VSL overshoot. This finding in field test is a result, and different component in the system contribute to that discrepancy. Part of the reasons can be the improper missing data imputation method since in current system the imputation algorithm imbedded in software is average of surrounding detectors which is proved not the best method so far. Another part of reason is the modification of desire speed term in METANET, the optimal VSL value may be hard for drivers to follow during short period of time, and that the current optimal VSL value does not consider driver compliance rate, so that when desire speed term value is far from 80km/h, prediction error become large.

In this thesis, the missing data problem and METANET modification problem are investigated separately under the title of MPC-VSL implementation. In the future, those two problems are to be investigated further and together in terms of how they eventually cause prediction error of VSL-METANET model and then cause the discrepancy between expected and actual measure of effectiveness. More sensors will be installed on exact DMS locations. The before-and-after study will be done to enable direct observation of the effectiveness of

MPC-VSL control. All those plans will be carried on in the next stage of MPC-VSL field test in the future.

REFERENCES

[1] B. Smith, W. Scherer, and J. Conklin, "Exploring Imputation Techniques for Missing Data in Transportation Management Systems," *Transp. Res. Rec.*, vol. 1836, no. 03, pp. 132–142, 2003.

[2] C. Chen and K. Petty, "Freeway Performance Measurement System Mining Loop Detector Data," *Transp. Res. Rec. J. Transp. Res. Board*, no. January, pp. 96–102, 2001.

[3] H. Al-Deek, C. Venkata, and S. Ravi Chandra, "New Algorithms for Filtering and Imputation of Real-Time and Archived Dual-Loop Detector Data in I-4 Data Warehouse," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1867, pp. 116–126, Jan. 2004.

[4] B. Smith and S. Babiceanu, "Investigation of Extraction, Transformation, and Loading Techniques for Traffic Data Warehouses," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1879, pp. 9–16, Jan. 2004.

[5] L. Qu, L. Li, Y. Zhang, and J. Hu, "PPCA-based missing data imputation for traffic flow volume: A systematical approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 3, pp. 512–522, 2009.

[6] L. Li, Y. Li, and Z. Li, "Efficient missing data imputing for traffic flow by considering temporal and spatial dependence," *Transp. Res. Part C Emerg. Technol.*, vol. 34, pp. 108–120, 2013.

[7] M. T. Asif, N. Mitrovic, L. Garg, J. Dauwels, and P. Jaillet, "Low-dimensional models for missing data imputation in road networks," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, pp. 3527–

3531, 2013.

[8] B. D. Ni, A. Guin, and C. Feng, “A Multiple Imputation Scheme for Overcoming the Missing Values and variability issues in ITS data," *Journal of transportation engineering*, vol 131, No.12, pp: 931-938, 2005.

[9] K. Henrickson, Y. Zou, and Y. Wang, “Flexible and Robust Method for Missing Loop Detector Data Imputation,” *Transp. Res. Board 94th Annu. Meet.*, vol. No. 15–580, no. 206, 2015.

[10] B. M. Williams and L. a. Hoel, “Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results,” *J. Transp. Eng.*, vol. 129, no. 6, pp. 664–672, 2003.

[11] C. Antoniou, M. Ben-Akiva, and H. Koutsopoulos, “Online Calibration of Traffic Prediction Models,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1934, pp. 235–245, Jan. 2005.

[12] S. Afandizadeh and J. Kianfar, “A Hybrid Neuro-Genetic Approach to Short-Term Traffic Volume Prediction,” *Int. J. Civ. Eng.*, vol. 7, no. 1, pp. 41–48, 2009.

[13] B. L. Smith, B. M. Williams, and R. Keith Oswald, “Comparison of parametric and nonparametric models for traffic flow forecasting,” *Transp. Res. Part C Emerg. Technol.*, vol. 10, no. 4, pp. 303–321, Aug. 2002.

[14] J. Laval, Z. He, and F. Castrillon, “Stochastic Extension of Newell’s Three-Detector Method,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2315, pp. 73–80, Dec. 2012.

[15] A. Spiliopoulou, M. Kontorinaki, M. Papageorgiou, and P. Kopelias, "Macroscopic traffic flow model validation at congested freeway off-ramp areas," *Transp. Res. Part C Emerg. Technol.*, vol. 41, pp. 18–29, Apr. 2014.

[16] M. J. Lighthill and G. B. Whitham, "On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads," *Proc. R. Soc. London A Math. Phys. Eng. Sci.*, vol. 229, no. 1178, pp. 317–345, May 1955.

[17] P. I. Richards, "Shock Waves on the Highway," *Oper. Res.*, vol. 4, no. 1, pp. 42–51, Feb. 1956.

[18] C. F. Daganzo, "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transp. Res. Part B Methodol.*, vol. 28, no. 4, pp. 269–287, Aug. 1994.

[19] C. F. Daganzo, "A behavioral theory of multi-lane traffic flow. Part I: Long homogeneous freeway sections," *Transp. Res. Part B Methodol.*, vol. 36, pp. 131–158, 2002.

[20] C. F. Daganzo, "A behavioral theory of multi-lane traffic flow. Part II: Merges and the onset of congestion," *Transp. Res. Part B Methodol.*, vol. 36, no. 2, pp. 159–169, 2002.

[21] H. J. Payne, "Models of freeway traffic and control.," *Math. Model. public Syst.*, 1971.

[22] J. A. Laval and C. F. Daganzo, "Lane-changing in traffic streams," *Transp. Res. Part B Methodol.*, vol. 40, no. 3, pp. 251–264, 2006.

[23] Y. Wang and M. Papageorgiou, "Real-time freeway traffic state estimation

based on extended Kalman filter: a general approach," *Transp. Res. Part B Methodol.*, vol. 39, no. 2, pp. 141–167, 2005.

[24] L. Mihaylova, R. Boel, and A. Hegyi, "Freeway Traffic Estimation within Recursive {Bayesian} Framework," *Automatica*, vol. 43, no. 2, pp. 290–300, 2007.

[25] J. C. Herrera and A. M. Bayen, "Incorporation of Lagrangian measurements in freeway traffic state estimation," *Transp. Res. Part B Methodol.*, vol. 44, no. 4, pp. 460–481, 2010.

[26] G. F. Newell, "A simplified theory of kinematic waves in highway traffic, part II: Queueing at freeway bottlenecks," *Transp. Res. Part B*, vol. 27, no. 4, pp. 289–303, 1993.

[27] G. F. Newell, "A simplified theory of kinematic waves in highway traffic I: General theory. II: Queuing at freeway bottlenecks. III: Multi-destination flows.," *Transp. Res. Part B*, vol. 27, no. 4, pp. 281–313, 1993.

[28] M. Papageorgiou, J.-M. Blosseville, and H. Hadj-Salem, "Macroscopic modelling of traffic flow on the Boulevard Périphérique in Paris," *Transp. Res. Part B Methodol.*, vol. 23, no. 1, pp. 29–47, Feb. 1989.

[29] M. Papageorgiou, H. Hadj-Salem, J. M. Blosseville, and N. Bhouri, *Control, Computers, Communications in Transportation*. Elsevier, 1990.

[30] T. Maze, M. Agarwai, and G. Burchett, "Whether Weather Matters to Traffic Demand, Traffic Safety, and Traffic Operations and Flow," *Transp. Res. Rec.*, vol. 1948, pp. 170–176, 2006.

[31] J. Asamer and M. Reinthaler, "Estimation of road capacity and free flow

speed for urban roads under adverse weather conditions," *IEEE Conf. Intell. Transp. Syst. Proceedings, ITSC*, pp. 812–818, 2010.

[32] HCM, *Highway capacity manual 2010*. 2010.

[33] T.-J. Kwon, L. Fu, and C. Jiang, "Effect of Winter Weather and Road Surface Conditions on Macroscopic Traffic Parameters," *Transp. Res. Board Annu. Meet.* 2013.

[34] H. Alhassan and J. Ben-Edigbe, "Extent of Highway Capacity Loss due to Rainfall," *World Academy of Science, Engineering …*, vol. 6, no. 12. pp. 127–134, 2012.

[35] T. Hou, R. Alfelor, J. Kim, and M. Saberi, "Calibration of Traffic Flow Models under Adverse Weather and Application in Mesoscopic Network Simulation Procedures," *Transportation Research Record*, no. November 2012. pp. 23, 2013.

[36] W. H. K. Lam, M. Asce, M. L. Tam, X. Cao, and X. Li, "Modeling the Effects of Rainfall Intensity on Traffic Speed, Flow, and Density Relationships for Urban Roads," *Journal of Transportation Engineering*, vol. 4, no. July. pp. 758–770, 2013.

[37] S. Smulders, "Control by Variable Speed Signs - The Dutch Experiment," *IEE Conf. Publication*, pp. 99–103, 1992.

[38] C.-C. Chien, Y. Zhang, and P. A. Ioannou, "Traffic Density Control for Automated Highway Systems", *Automatica*, vol. 33, no. 7, pp. 1273–1285, Jul. 1997.

[39] M. Hadiuzzaman, T. Z. Qiu, and X.-Y. Lu, "Variable Speed Limit Control

Design for Relieving Congestion Caused by Active Bottlenecks," *J. Transp. Eng.*, vol. 139, no. April, pp. 358-370, 2012.

[40] E. van den Hoogen, "Control by variable speed signs: results of the Dutch experiment," *Seventh Int. Conf. `Road Traffic Monit. Control.*, vol. 1994, no. 391, pp. 145–149, 1994.

[41] A. Hegyi, B. De Schutter, and H. Hellendoorn, "Model predictive control for optimal coordination of ramp metering and variable speed limits," *Transp. Res. Part C Emerg. Technol.*, vol. 13, no. 3, pp. 185–209, 2005.

[42] A. W. Sadek, M. J. Demetsky, and B. L. Smith, "Case-Based Reasoning for Real-Time Traffic Flow Management," *Comput. Civ. Infrastruct. Eng.*, vol. 14, no. 5, pp. 347–356, Sep. 1999.

[43] J. Cuena, J. Hernández, and M. Molina, "Knowledge-based models for adaptive traffic management systems," *Transp. Res. Part C Emerg. Technol.*, vol. 3, no. 5, pp. 311–337, Oct. 1995.

[44] A. Alessandri, A. di Febbraro, A. Ferrara, and E. Punta, "Nonlinear optimization for freeway control using variable-speed signaling," *IEEE Trans. Veh. Technol.*, vol. 48, no. 6, pp. 2042–2052, 1999.

[45] E. F. Camacho and C. Bordons, *Model predictive control*. 2007.

[46] C. E. García, D. M. Prett, and M. Morari, "Model predictive control: Theory and practice—A survey," *Automatica*, vol. 25, no. 3, pp. 335–348, May 1989.

[47] J. Maciejowski, *Predictive Control: with Constraints*, Pearson education, 2000.

[48] A. Kotsialos, M. Papageorgiou, and F. Middelham, “Optimal coordinated ramp metering with advanced motorway optimal control,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1748, no. -1, pp. 55–65, 2001.

[49] A. Kotsialos, M. Papageorgiou, M. Mangeas, and H. Haj-Salem, “Coordinated and integrated control of motorway networks via non-linear optimal control,” *Transp. Res. Part C Emerg. Technol.*, vol. 10, no. 1, pp. 65–84, 2002.

[50] A Hegyi, M. Burger, B. De Schutter, J. Hellendoorn, and T. J. J. Van Den Boom, “Towards a practical application of model predictive control to suppress shock waves on freeways,” *Proc. Eur. Control Conf. 2007*, vol. 19, pp. 1764–1771, 2007.

[51] C. Chen, “Detecting Errors And Imputing Missing Data For Single Loop Surveillance Systems,” *Transp. Res. Board*, vol. 1855, no. January, pp. 160–167, 2003.

[52] A. Messner and M. Papageorgiou, “METANET: a macroscopic simulation program for motorway networks,” *Traffic Eng. Control*, vol. 31, no. 8–9, pp. 466–470.

[53] P. K. Munjal and L. A. Pipes, “Propagation of on-ramp density perturbations on unidirectional two- and three-lane freeways,” *Transp. Res.*, vol. 5, no. 4, pp. 241–255, 1971.

[54] L. Elefteriadou, S. S. Washburn, Y. Yin, V. Modi, and C. Letter, “Variable Speed Limit (VSL) – Best Management Practice,” *Florida Dep. Transp. Res. Center, USA*, vol. 11, no. July, 2012.