UNIVERSITÉ DE MONTRÉAL

DYNAMIC PREDICTION OF TRAFFIC CONDITIONS USING STREAMING DATA AND BAYESIAN APPROACH

MOHAMMAD KIANPOUR
DÉPARTEMENT DES GÉNIES CIVIL, GÉOLOGIQUE ET DES MINES
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE CIVIL)
JANVIER 2017

© Mohammad Kianpour, 2017.

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

DYNAMIC PREDICTION OF TRAFFIC CONDITIONS USING STREAMING DATA AND BAYESIAN APPROACH

présenté par : KIANPOUR Mohammad
en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées
a été dûment accepté par le jury d'examen constitué de :

M. TRÉPANIER Martin, Ph. D., président
M. FAROOQ Bilal, Ph. D., membre et directeur de recherche
M. PATTERSON Zachary, Ph. D., membre

DEDICATION

To my wife. That without her support I could not continue.
And to my older son, that without his understanding I had stopped. . . .

ACKNOWLEDGEMENTS

I am very grateful to my supervisor, Prof. Bilal Farooq, whose encouragement, guidance, hours of assistance, sound advices, and dedication immeasurably improved this thesis, and indeed were indispensable to its completion. I am thankful for these thoughtful and insightful comments to ensure that a quality document will result from this endeavor.
I really appreciate the valuable help of Prof. Catherine Morency and the City of Montréal for providing me with the data. Also, I must thank Dr. Anae Sobhani and Mr. Hamzeh Alizadeh, PhD candidate, for the time that they spent on reading this thesis and their comments for better writing the document.

RÉSUMÉ

La diffusion des données peut être définie par son volume remarquable, la génération de vitesse, la richesse de l'information et la diversité de l'information. Aujourd'hui, de grandes quantités de données, produites par de nombreuses sources nouvelles, peuvent être classées dans ce type de données. Des domaines tels que le transport et l'ingénierie du trafic peuvent bénéficier de ces ensembles de données. L'utilisation de diffusion de données pour créer de nouvelles méthodes de modélisation des comportements de déplacement peut être utile de trois manières. Tout d'abord, cela peut réduire le temps requis et le coût de collecte de données suffisantes avec les méthodes conventionnelles. Deuxièmement, cela peut augmenter le niveau de précision des modèles proposés et les simulations mises en œuvre. De plus la diffusion de données peut réduire la dépendance à l'égard des données traditionnelles.

La prédiction des conditions de circulation a des effets notables sur la diminution des effets nocifs de la congestion dans un réseau urbain. Dans cette thèse, nous définissons une approche pour l'utilisation de la diffusion des données provenant de diverses sources disponibles de façon ubiquitaire, par ex. Les traces de GPS (Global Positioning System), la caméra de trafic, l'approvisionnement de foule, etc. La diffusion de données nous donne la possibilité d'utiliser plus d'informations avec les sources de données traditionnelles pour développer des modèles de transport urbain. Par conséquent, nous élaborons un cadre qui utilise des données en continu pour prédire les conditions de circulation à court terme au niveau du lien dans un réseau urbain. L'approche bayésienne qui définit la probabilité d'un événement basé sur d'autres événements expérimentés a été mise en œuvre pour mettre à jour la probabilité de conditions de circulation possibles. En utilisant l'approche bayésienne qui exploite la diffusion de données en continu, nous créons un nouveau type de modèle de prédiction du trafic dans cette étude.

Nous élaborons un modèle de prévision qui utilise les connaissances existantes sur les conditions du trafic urbain et les données en continu fournies en tant que nouvelles sources non traditionnelles. Les conditions de circulation sont simulées par un modèle antérieur. En développant l'approche bayésienne pour prédire la probabilité d'un état, en utilisant à la fois la simulation des conditions de circulation et les mesures provenant des sources de données en continu, nous considérons la probabilité de notre connaissance préalable des conditions de circulation, la probabilité des mesures et le modèle conditionnel pour chaque état possible des conditions de circulation dans le réseau urbain.

La probabilité d'un état basé sur les données de diffusion disponibles est définie en multipliant la probabilité de l'état défini par le modèle précédent par le rapport de la probabilité définie par le modèle conditionnel sur la probabilité des mesures. Nous définissons la liaison de réseau d'état par liaison. Le résultat de la méthode proposée est l'état le plus probable du système en considérant les données de diffusion disponibles à chaque temps. Par conséquent, nous sommes en mesure de recommander l'état le plus probable à chaque étape de temps ainsi que les paramètres de lien (vitesse, débit et densité).

L'approche proposée est appliquée à un cas réel pour démontrer la faisabilité de l'approche. Deux routes artérielles Nord-Sud vers le centre-ville de Montréal ont été considérées comme l'étude de cas. Les états probables du réseau urbain sont définis en simulant diverses conditions de circulation basées sur différents ensembles de temps de déplacement qui sont disponibles à partir des données fournies par les véhicules flottants qui ont parcouru les routes choisies.

Les résultats démontrent que l'approche bayésienne proposée est capable de nous guider dans la définition de l'état le plus probable du réseau routier urbain. Cette approche peut améliorer les capacités de prévision des conditions à court terme du réseau routier urbain.

De plus, les prévisions de trafic à court terme, en tant que modèle de mise à jour, peuvent aider les opérateurs et les utilisateurs des réseaux routiers urbains à optimiser leurs décisions d'exploitation et d'utilisation de ces réseaux.

Mots-clés: Mise à jour bayésienne, diffusion de données, Prévision de flux de trafic, Simulation.

ABSTRACT

Streaming data can be defined by its remarkable volume, generation of speed, richness of information, and diversity of information. Today, large volumes of data, produced by many new sources, can be classified as this type of data. Domains such as transportation and traffic engineering can benefit from these datasets. Using streaming data to create new methods of travel behaviour modelling can be helpful in three ways. First of all, it can reduce the required time and cost of collecting sufficient data from conventional methods. Secondly, it can increase the accuracy level of proposed models and implemented simulations. Finally, it can reduce dependency on traditional cross-sectional data.

Predicting traffic conditions has noticeable effects on decreasing the harmful impacts of congestion in an urban network. In this thesis, we define an approach for using streaming data coming from various ubiquitously available sources e.g. GPS (Global Positioning System) traces, traffic camera, crowdsourcing, etc. Streaming data provides us with the possibility of using more information along with traditional sources of data to develop urban transportation models. Therefore, we propose a framework that uses streaming data for predicting short-term traffic conditions at link level in an urban network. The Bayesian approach that defines the probability of an event based on other experienced events has been implemented to update the probability of possible traffic conditions. By using the Bayesian approach that exploits streaming data, we create a new type of traffic prediction model in this study.

We develop a prediction model that uses the existing knowledge of urban traffic conditions and streaming data provided as new non-traditional sources. The traffic conditions are simulated by a prior model. In developing the Bayesian approach to predict the probability of a state, using both simulation of the traffic conditions and measurements coming from the source of streaming data, we consider the probability of our prior knowledge of traffic conditions, the probability of measurements, and the conditional model for each possible state of traffic conditions in the urban network.

The probability of a state based on the available streaming data is defined by multiplying the probability of the network state defined by the prior model by the ratio of the probability defined by the conditional model over the probability of measurements. We define the network state link by link. The output of the proposed method is the most probable state of the system considering available streaming data at each time-step. Therefore, we are able to recommend the most probable state at each time-step as well as link parameters (speed, flow, and density).

The proposed approach is applied to a real case to demonstrate the feasibility of the ap-

proach. Two North-South arterial routes to Montréal downtown have been considered as the case study. The probable states of urban network are defined by simulating various traffic conditions based on different sets of travel times that are available from the data provided by floating vehicles that have traveled the chosen routes.
Results demonstrate that the proposed Bayesian approach is able to guide us for defining the most probable state of urban road network. This approach can improve capabilities of short-term conditions prediction of the urban road network.
Moreover, short-term traffic prediction results, as an updating model, can help operators and users of urban road networks to optimize their decisions for operating and using these networks.

Keywords: Bayesian update, Streaming data, Traffic flow prediction, Simulation.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

GPS	Global Positioning System
ITS	Intelligent Transportation System
SM	Social Media
FC	Floating Car
SP	Smart Phone
SC	Smart Card
AVI	Automatic Vehicle Identification
AVL	Automatic Vehicle Location
PF	Particle Filtering
LN-CTM	Link Node Cell Transmission Model
PHD	Probability Hypothesis Density
LWR	Lighthill, Whitham, and Richards
NN	Neural Network
CTM	Cell Transmission Model
CA	Cellular Automata
GHR	Gazis-Herman-Rothery
AP	Action Point
SA	Simulated Annealing
TT	Travel Time
State	Traffic conditions in an urban road network
Status	Speed value for each probe vehicle defines its status between two possible statuses
MPS	Most Probable State
SUMO	Simulation of Urban MObility
FFS	Free Flow Speed
NB	North Bond
SB	South Bond
QGIS	Quantum GIS
GUI	Graphical User Interface
TS	Time-Step
t	Time index
j	Probe vehicle index
J	Total number of probe vehicles

$a_j(t)$	Acceleration of car j at time t
$v_j(t)$	Speed of car j at time t
T	Reaction time of a driver
$\Delta v(t-T)$	Relative speed between j^{th} car and the $(j-1)^{th}$ car located in front
$\Delta x(t-T)$	Space between j^{th} car and the $(j-1)^{th}$ car located in front
c, f, k	Model parameters in GHR model
α_1, α_2, α_3, and α_4	Model parameters in safe distance model
β_1, β_2, β_3 , β_4, and β_5	Model parameters in linear (Helly) model
$d_j(t)$	Desired following distance for car j at time t
w	Width of a car
Θ	Visual angle of leader car from point of view of follower driver
A	Current event in general Bayesian equation
B	Experienced event in general Bayesian equation
i	Link index
I	Total number of links
S_t^i	State of link i at time t
v_t^i	Speed of link i at time t
d_t^i	Density of link i at time t
q_t^i	Flow of link i at time t
S_t	Prior state of an urban road network at time t from simulation
$\widehat{S}_t$	State of an urban road network at time t from measurements
N	Number of possible sets of TT in the urban road network
$\widehat{v}_t^{ji}$	Reported speed of probe vehicle j located in link i at time t
$(S_{t1}, ..., S_{tn}, ..., S_{tN})$	N possible states of traffic conditions at time t
$(SP_1, ..., SP_n, ..., SP_N)$	N states probabilities for N possible states of traffic conditions
PM_t	Probability of measurements at time t
$(CM_{t1}, ..., CM_{tn}, ..., CM_{tN})$	N conditional models for N possible states of traffic conditions at time t
$(BP_{t1}, ..., BP_{tn}, ..., BP_{tN})$	N Bayesian probabilities at time t for N possible states of traffic conditions
t_0^{ji}, x_0^{ji}	The time and location that the speed of probe vehicle j

	decreases to zero in link i
t'^{ji}	Time of entering probe vehicle j into link i
q_x	Rate of change of flow with respect to space
d_t	Rate of change of density with respect to time
Δt	The reference period before time t for defining the probability of measurements
$\sigma_{\hat{v}^i_{t-\Delta T,t}}$	Standard deviation of measured speed between $t-\Delta T$ and t for link i
$\mu_{\hat{v}^i_{t-\Delta T,t}}$	Average of measured speed between $t-\Delta T$ and t for link i
$t^\star$	Dissipation time of the queue located ahead of the probe vehicle
X^{Di}	The coordinate of the intersection located downstream of the link i
L	The effective average length of vehicle
S_r	The uniform service rate of intersection
Δt_r	Remaining time of red time of the signal located at down stream after the time that the probe vehicle starts reporting its speed as zero
t_G	The time that the first next green time changes to red
Q_{Ui}	The volume of arrived vehicles after the probe vehicle from upstream of the link i
λ	Density mean of link i between t_0^{ji} and t
$\sigma_{\hat{v}^i}$	Standard deviation of measured speed in link i

LIST OF APPENDICES

CHAPTER 1 INTRODUCTION

With the increase in car ownership rate in megacities, people have increased their car usage for daily activities. As a result, megacities experience high levels of traffic congestion, on a daily basis. Governments cannot build new infrastructure to cover the increasing demand. Therefore, new methods have emerged to deal with the imbalance in the supply and demand. Finding new policies and solutions is a high priority for decision makers and researchers. Various policies have been proposed to decrease the congestion on urban roads. For example, investing strategically in public transit, optimizing road network performance, eliminating bottlenecks, adding and optimizing new roadway capacity, and road pricing have been recommended (Staley, 2012). Also, real time management and route guidance have been considered as a suitable solution to increase network capacity. However, in order to assess the sustainability of these policies, a precise understanding is required about the spatio-temporal dynamics of traffic conditions on the urban road network.

It is required to know, what are the basic traffic indicators such as speed, flow, and density, in quasi-realtime on link-by-link basis of the road network. Furthermore, it is possible to perform short-term forecasting of the state of the system based on these statistics. Finally, it would enable us to pre-emptively apply various strategies to optimize the flow and maximize the throughout of the network.

Accessing the information produced by new sources, commonly known as Big Data, enables us to understand various phenomena in transportation and traffic engineering. Especially, it helps us for short-term vehicular flow modelling. Using streaming data obtained from the user-centric sources will help us improve the understanding of recently-experienced events in urban road networks. Thus, we can draw a more realistic picture of the short-term traffic conditions by developing new prediction methods. The streaming nature of the data provided by this new source alongside with traditional data that need noticeable finance and effort, can be exploited for better short-term predictions.

In order to make traditional models of demand and activity behaviour, the required data is collected using a small subset of the whole population for a short duration of time every 5 to 10 years. Nowadays, new technologies like GPS, sensors, smart phones, and social media platforms provide us with the opportunity to increase the volume of collected data to develop more accurate transportation models (Vij and Shankari, 2015). Predicting traffic conditions is one of the fundamental parts in Intelligent Transportation System (ITS) (Li et al., 2015). Using more data enables us to develop better models to predict future conditions of the road network more precisely.

Existing traffic simulations are calibrated once using traditional datasets that are predominantly cross-sectional in nature and collected once every couple of years. In these simulations there is no feedback from the recently experienced states of traffic conditions in the network and they are only recalibrated whenever the new cross-sectional dataset is available. These simulations are used to predict the future of the urban network without considering the short-term important feedbacks from local spatial and temporal variations. Therefore, there is a need to consider and apply these feedbacks and update our simulations based on the new information about the short-term state of the urban network. Using recently available streaming data from various sources such as smart phones, traffic cameras and social networks, we can simulate short-term traffic conditions with feedback to incorporate recent conditions.

Streaming data/longitudinal data sources provide us with the ability to develop new methods of predicting traffic conditions. Streaming data are the most recently addressed sources of data that can be used to improve our knowledge of traffic conditions. Moreover, their real-time features give us the opportunity to manage and control traffic in order to increase the performance of transportation systems (Shi and Abdel-Aty, 2015). Streaming data can be collected using network based or user-centric techniques (Danalet et al., 2014). Streaming data collected by user-centric techniques can provide individual level patterns of traffic behaviour.

More formally, Social Media (SM), Floating Car (FC), Smart Phone (SP) and Smart Card (SC) are categorized in the set of user-generated data sources. Mining, managing and analyzing this type of data, in order to understand the complexity and dynamics of trip patterns, is a noticeable job considering the amount, richness, and dynamism. Therefore, we need to create new advanced methods to analyze these data; develop individual-level models; and simulate individual transportation and activity patterns (Gkiotsalitis and Stathopoulos, 2015).

User-centric data generated by the users of the urban road network, enable us to access the required information that we need to define the spatio-temporal fluctuations of traffic indicators at link level. We can define different possible states of traffic conditions in an urban network. By accessing to real-time data about traffic indicators in part of urban road network, we can define which possible state of traffic conditions in whole of urban road network is more probable, regarding available streaming data received form user-centric sources.

In this research, we concentrate on using the Bayesian approach to define the most probable state of an urban network using streaming data provided by user-centric sources as well as the simulation of traffic dynamics. We define the prior model using a traffic microsimulator to create possible states of traffic conditions in the urban road network. Also, we calculate the probabilities of measurements using user-centric data available as streaming data and measure a conditional model for each possible state of urban traffic conditions at each time-step, using the output of the traffic microsimulator and streaming data. The proposed methodology enables us to predict link-by-link traffic conditions at a given time, based on available streaming data.

This research has two main cores: a traffic microsimulator that can generate the traffic conditions in possible states of the urban road network and a proposed Bayesian approach that enables us to decide about the most probable state at each time-step. Therefore, in the following chapter on the review literature we focus on these two main topics to elaborate which traffic modelling frameworks and prediction methods have been used and what are our best options in these two main cores of this study.

The rest of this thesis is organized as following: First, we will present a literature review to describe how various approaches have been used for short-term traffic predictions. Next, our methodology development is presented that explains how the proposed Bayesian approach is used to define the most probable state of the urban road network. A case study of Montréal will be explained and after reviewing and analyzing the available data, some discussion about the results of this study and the sensitivity of the analysis will be presented. The validity of the results of the proposed method will be tested and finally, conclusions and future research directions are presented.

CHAPTER 2 LITERATURE REVIEW

This chapter has been divided into two main sections so as to cover the existing literature dedicated to predicting traffic flow conditions in short-term and different modelling frameworks. The first section presents the methods of short-term prediction of traffic flow. Deciding about the best underlying modelling framework is an important task. Hence, the second section describes advantages and disadvantages of the different traffic modelling frameworks that work at various levels of spatio-temporal aggregation and employ different assumptions in modelling.

2.1 Short-term prediction of traffic flow

Level of complexity of models, spatio-temporal coverage of available data, characteristics of data used for calibration and validation, existence of non-linearity relationship between indicators, and temporal variations in observations for an indicator in an urban road network are some examples of the factors that affect the reliability of traffic condition predictions (Canaud et al., 2013). Consequently, developing new methodologies for overcoming these issues is essential.

Existing studies use different sources of data from various detection technologies. In next part, we present a brief review of available detection technologies. Afterwards, we focus on different method of short-term prediction of traffic flow in existing literature.

2.1.1 Detection technologies

For monitoring urban traffic, in order to understand its macroscopic characteristics such as link level travel time, speed, and flow (Nantes et al., 2013) and to answer the numerous needs of the urban community, traffic information systems play a specific and paramount role (Herring, 2010). Recently, there is a growing endeavor in estimating and forecasting arterial network traffic conditions, by using probe data (Hofleitner et al., 2012). Short-term traffic management and control can greatly benefit by predicting traffic flow indicators such as travel time, speed, and volume.

In order to predict traffic conditions, we need detection/tracking technology to measure traffic indicators in urban road network. There are three types of technologies: in-roadway and over-roadway detectors and off-roadway technologies (Shi and Abdel-Aty, 2015). Loop detectors are examples of in-roadway detectors. Installing and repairing these types of sensors can disturb traffic. In specific situations, they have noticeable failure rates; for instance, in unsuitable weather and road surface conditions. Located over the roadway or alongside and placed at some distance from the nearest traffic lane, the over-roadway sensors have evolved from video image processing to more up-to-date microwave radar sensors. Their most important advantage is minimizing traffic disruptions for installation and maintenance. Probe vehicle is an example of the off-roadway technologies that need in-vehicle devices in addition to fixed infrastructures compared to in - and over - roadway technologies. GPS, cellular phones, Bluetooth, Ground-Based Radio Navigation, Automatic Vehicle Identification (AVI) and Automatic Vehicle Location (AVL) are some examples of this type of detection technologies.

In - and over - roadway technologies have the ability of traffic counting, but speed measurement by these two types of technologies is less accurate than probe vehicle as an example of off-roadway technologies (Martin et al., 2003). While all vehicles do not have in-vehicle devices, access to complete and accurate information about some traffic indicators is not possible using off-roadway technologies (Shi and Abdel-Aty, 2015).

Various existing studies have used different sources of data and have employed different approaches for predicting traffic conditions using available data. In next part, we concentrate on existing methods of short-term prediction of traffic conditions.

2.1.2 Short-term traffic prediction methods

Park and Lee (2004) created travel time models on arterial roads using loop detectors in the network and a dedicated probe vehicle with installed communication devices. They implemented neural networks to calibrate the model and validated their model under limited conditions. Coifman (2002) developed a method using data from a loop detector installed at a single point of a freeway to define link travel time. The proposed method is not able to consider the effect of a delay or a queue resulted by an incident. Coifman and Krishnamurthy (2007) developed a method for detecting congestion in a link located between two loop detectors installed in a freeway by reidentificating long vehicles. In arterials, access to land use is not limited as freeways. Therefore, reidentification of some vehicles that exit arterials between two consecutive loop detectors is not possible between all loop detectors at the end of links. Hence, it can be concluded that their proposed method is not usable for arterials.

Jenelius and Koutsopoulos (2013) used GPS traces to develop an arterial road travel time model in Stockholm. They implemented a maximum likelihood based approach for correlating link characteristics, mean link conditions, and weather conditions to the link speed. Their statistical approach is noticeable that considers the effects of mentioned conditions. But the approach is not able to consider the effects of the traffic conditions from the surrounding network that is possible by implementing a traffic simulation. In their statistical approach, travel time between two points was considered as a combination of stochastic links' travel times and deterministic intersections' delays. Their assumption about delay at intersection cannot reflect what happens in reality.

Pascale et al. (2013) estimated link densities in highway using noisy and sparse measurements of density collected from road-embedded sensors, by proposing recursive Bayesian approach combined by a particle filtering (PF) approach (a sampling method). In traffic simulation, they used a stochastic form of Link Node Cell Transmission Model (LN-CTM). They implemented numerical tests and concluded that even in situations of very sparse sensor deployment, estimating the traffic conditions is possible using mathematical methods. They modeled traffic conditions at a macroscopic level that does not enable their proposed approach for using individual level data.

Canaud et al. (2013) created a new tool for estimating vehicular density on urban roads. They proposed the use of Probability Hypothesis Density (PHD) for solving the state estimation problem. By evaluating their PHD filter against particle filter (PF), they concluded that the PHD filter can be considered as a suitable alternative. The proposed method is not appropriate for developing an approach to use real-time traffic data for defining the short-term state of urban road. The approach is suitable for studying a specific link and cannot enable us to define traffic indicators in the whole urban road network when we have access to measurements in only a part of network.

Herring (2010) proposed a hybrid approach by implementing a general system architecture to analyze traffic data and disseminate accurate, timely traffic information through the Internet. This approach leveraged the advances in the fields of machine learning and traffic theory (using hydrodynamic theory) to estimate arterial traffic conditions. Two different approaches using historical data in modelling traffic conditions applied: Regression and Bayesian. The second approach has been introduced as the better one based on the results of this study. It is noticeable that learning from historical data when there is considerable changes in traffic volumes cannot be accurate. Also, there is a need to combine the Bayesian approach with traffic simulation for achieving more trustworthy results that has not been applied to this study.

Hofleitner et al. (2012) used a dynamic Bayesian network. They used sparsely observed probe vehicles, and proposed a probabilistic modelling framework for approximately defining and forecasting travel time distributions on urban roads. Their method represented the spatio-temporal dependency on the network by considering traffic state as hidden, which is affected by travel time observations. It provided a flexible framework to learn traffic dynamics from historical data and to perform real-time estimations with streaming data. While in the Bayesian updating method the prediction at each time step is updated based on available data, the proposed method in this study considered the dynamism that affects the current traffic state. The sampling rate of probe vehicles in their study did not provided detailed information about the location where vehicles encountered delay. Compared to the baseline approach, their method provided 35 % increase in estimation accuracy. This approach in short-term modelling traffic conditions is notable as it only considers experienced traffic conditions reported by GPS traces. But this approach is not able to provide predictions about the part of the network that is not covered by probe vehicles at all.

Nantes et al. (2013) captured the spatio-temporal relationship between volume and travel time. They proposed an approach based on a simple Bayesian network for analyzing and predicting the complex dynamics of flow or volume that was based on travel time observations from Bluetooth sensors. Their proposed Bayesian approach was simple as it considered the current state dependent only on the previous state and the travel time indicator dependent only on current state (first-order transition model). They considered distribution of observations by sensors Gaussian. The complex dynamics of arterial volume and travel time was estimated and predicted effectively by their proposed Bayesian network. They estimated the joint distributions, over sequences of volume values. Their proposed method focuses on the streaming data in part of the urban road network. The method is not able to produce results about part of the network that has not been covered by the sensors.

Polson and Sokolov (2014) developed a PF and learning method to estimate current traffic conditions (density) and key parameters of Lighthill, Williams, Richards (LWR) model. They proposed an algorithm that could improve on existing methods by permitting real-time updating of the posterior distribution of the critical density and capacity parameters through on-line learning. Even at shock waves, their PF algorithm was able to estimate uncertainty of the traffic state by considering mixture distribution. Moreover, by using the Bayesian parameter learning and estimation, improving the estimation bias was possible. By fixing the parameters it led to mis-estimation. If the relationship between flow and density changes over time or there is a queue or bypassing traffic, the proposed method is not realistic even in highways.

Staňková and De Schutter (2010) proposed a PF method for predicting/estimating traffic density using Daganzo's cell transmission model, based on jump Markov linear system (a nonlinear switching Markov system of multiple linear models). They used importance sampling in their PF method to select samples with higher normalized importance weights, the ratio of a sample probability over the probability of an arbitrary proposal distribution. They concentrated on the performance of their method and the system properties by comparing its result with the output of microsimulation by Aimsun 6. Although, they assumed that congestion happens in one segment only, the extension of their approach to the model that contains all possible congestion modes was straightforward. They mentioned that more research is needed to apply the proposed method to the more general problems respecting performance in time, variance, etc.

Chen and Rakha (2014) used a new PF approach to define travel time in highway by applying real-time and historical data. They used a different sampling strategy to filter out invalid or low wighted particles and replace them with historical data that had the same data sequence in real-time measurements.

Modelling urban road traffic using traditional methods of data collection has some disadvantages. To overcome these disadvantages, there are various approaches already proposed, which can be categorized into Statistical approach to Neural Network, Bayesian approach and so on. A summary of these methods is presented in Table 2.1.

Vlahogianni et al. (2014) reviewed the literature related to short-term forecasting methods.

Table 2.1 Applied approaches in short-term prediction of traffic flow

Applied approach	Brief description	Literature used the approach
Neural Network (NN)	A computing system of a number of simple, highly interconnected parts that processes information by dynamic reply to external input	Park and Lee (2004)
Maximum Likelihood based	A method of estimating the parameters of a statistical model given data	Jenelius and Koutsopoulos (2013)
Particle Filtering	A method for estimating the internal states in dynamical systems when there are partial observations and also random perturbations in used sensors and dynamical system	Pascal et al. (2013), Polson and Sokolov (2014), Stankova and De Schutter (2010), Chen and Rakha (2014)
Probability Hypothesis Density (PHD)	A multiple target filter for estimating the number and the state of a set of targets using some observations	Canaud et al. (2013)
Hybrid	A combination of statistical and computational approaches	Herring (2010)
Bayesian Believe Network	Defines probability of an event on the context of a prior knowledge	Hofleitner et al.(2012), Nantes et al. (2013), Polson and Sokolov (2014)

They mentioned 10 important challenges related to this topic. These challenges are related to integration of predicting models; the problem of forecasting traffic and variable choice; data issues and the impact of new technologies on available datasets; developing novel prediction algorithms; and the effect of artificial intelligence model on integrating models into prediction schemes.

Reviewing existing literature, summarized in Table 2.1, it can be concluded that the dynamic prediction of link by link traffic conditions in whole of a complete network using streaming data and Bayesian approach is not possible by the proposed methods of existing literature.

Although, there are some Bayesian approaches have only been used at corridor or neighborhood level, the proposed methods are not able to predict traffic conditions in whole of a network when we have access to streaming data in a part of network. The Bayesian approach in most of the reviewed existing literature has been used to explore historical and real-time data for estimating traffic state in a part of the urban road network. While combination of a simulation based on traditional data and a Bayesian approach that uses streaming data available in a part of network and the simulation outputs has not been applied in the related literature to predict traffic conditions in the whole network. We need to do this combination to increase the ability of the new proposed approach of this study in traffic conditions prediction when we have no information in a part of the network regarding the randomness in space and time of streaming data. Therefore, we aim to study the problem and we will explain our rational for the selected method. In the following part, we focus on reviewing the literature in order to choose a suitable traffic flow modelling framework.

2.2 Traffic modelling frameworks

We need to model traffic conditions in an urban network; therefore, a review of existing traffic modelling framework is presented here.

2.2.1 Cell transmission model (CTM)

In this framework, the network is divided into cells based on the differences in link properties and connections to other links (Alecsandru, 2006). The occupancy of each cell at each time-step is defined based on its occupancy, inflow at the last time-step, and the outflow from the connected downstream links at the current time-step. This approach is reasonably suited and appropriate for planning analysis problems and its macroscopic nature makes it more efficient from a computational point of view. It also requires minimal effort for calibration since it has comparably less parameters. CTM makes traffic flow modelling more realistic [1] in various implementations containing both static and dynamic traffic assignment procedures. This method is suitable for some specific applications, for instance, it is suitable for analysis of disaster evacuation solutions at the planning level. In CTM, the network is divided into cells resulting in possible geometric limitations (Alecsandru, 2006). If the network is divided into many small cells, we need additional memory. Therefore, CTM is not suitable for modelling large-scale urban network. We want to model different facility types. For instance, when we have freeways and intersections on city streets, modelling of large-scale networks will be

1. Because the way it defines the state of each cell at each time-step is more compatible with what happens in reality.

computationally inefficient.

2.2.2 Cellular automata (CA)

It is used in traffic flow microsimulation to model the behaviour of cars traveling on a road network. The basic idea is to discretize space in cells of equal size, each of which can be occupied by at most one vehicle. Cars drive through these cells by choosing their speed according to the space available in front of them. The advantage of cellular automata (CA) is that link capacities are generated from the properties of the links and drivers' behaviour. The major drawback of this method is its computational cost as every agent is simulated in every time step, which is usually a second (Maerivoet and De Moor, 2005). Ignoring the computational cost, developing a model by CA can be useful for modelling special links and is not suitable for modelling a whole network.

2.2.3 Queue-based simulations

In this method, the intersections would be modeled alone where (Charypar et al., 2007):

1. Links are modeled as they "process" cars moving through the network.
2. A queue stores cars coming from each link together with their respective entry times.
3. Each link's capacity and space available for cars are parameters of the model.
4. The related links cooperate in each time-step by considering different link constraints such as capacity, free speed, travel time, intersection precedence and space available at the next link, in order to define the next position of cars in the network.

Queue-based models are faster than CA mainly because the number of simulated units is smaller (links vs. cars). A special form of queuing simulation (event-based) is based on the occurrence of events (Charypar et al., 2007). There is no fixed time-step and the cars are only processed when an associated event (e.g. arrival, departure, etc.) occurs. Entering or leaving a link is also considered an event. It means that event processing rate is related to the flow in all time-steps.

2.2.4 Car-following models

This type of models describe the behaviour of drivers in an urban network (Brackstone and McDonald, 1999). They assume that the behaviour of each car is a function of the proceeding car. To model the reaction of driver to the followed car, there are various types of models that can be classified as car-following models. Gazis-Herman-Rothery (GHR) model (Gazis et al., 1961), safety distance or collision avoidance models (Kometani and Sasaki, 1959), and Linear (Helly) models (Helly, 1959) are classified in this category. Also, there are other approaches in this category, the most important types of which are Psychophysical or Action Point (AP) models (Michaels, 1963) and Fuzzy logic-based models (Kikuchi and Chakroborty, 1992).

GHR model

In the GHR model, acceleration of car j at time t, $a_j(t)$, is defined by (Brackstone and McDonald, 1999):

$$a_j(t) = c(v_j(t))^f \frac{\Delta v(t-T)}{(\Delta x(t-T))^k} \tag{2.1}$$

Where $v_j(t)$ is the speed of car j at time t, $\Delta v(t-T)$ and $\Delta x(t-T)$ are the relative speed and the space between j^{th} car and the $(j-1)^{th}$ car located in front, T is reaction time of a driver, and c, f, and k are the model parameters. This type of car-following models introduced by Chandler et al. (1958), has been adopted and improved by Herman et al. (1959), Gazis et al. (1961) and others.

Safe distance model

In safe distance models, the safe following distance, $\Delta x(t-T)$, is defined by (Brackstone and McDonald, 1999):

$$\Delta x(t-T) = \alpha_1.(v_{j-1}(t-T))^2 + \alpha_2.(v_j(t))^2 + \alpha_3.v_j(t) + \alpha_4 \tag{2.2}$$

Where α_1, α_2, α_3, and α_4 are the model parameters.
This set of car-following models were introduced by Kometani and Sasaki (1959) and improved by Gipps (1981). They were used in some simulating models, for instance, in INTRAS and CARSIM in the USA (Benekohal and Treiterer, 1988).

Linear (Helly) model

In Linear (Helly) model, $a_j(t)$ is defined by (Brackstone and McDonald, 1999):

$$a_j(t) = \beta_1.\Delta v(t-T) + \beta_2.(\Delta x(t-T) - d_j(t)) \tag{2.3}$$

$$d_j(t) = \beta_3 + \beta_4.v_j(t-T) + \beta_5.a_j(t-T) \tag{2.4}$$

Where $d_j(t)$ is desired following distance for car j at time t and β_1, β_2, β_3, β_4, and β_5 are the model parameters. This distance is a linear function of the speed and acceleration of the car by considering the reaction time of the driver. This model proposed by Helly (1959), then calibrated by Hanken and Rockwell (1967) and others. Also, in the mid nineties, a combination of the linear model and GHR model proposed by Xing (1995).

Action Point (AP) model

In AP model that was proposed by Michaels (1963), a driver in approaching a vehicle considers its visible dimension by realizing the relative speed that affects the visual angle of the vehicle (Brackstone and McDonald, 1999). Figure 2.1 shows how the visual angle and its changes over time can be defined using next equations:

$$\Theta = 2\arctan(\frac{w}{2\Delta x}) \tag{2.5}$$

$$\frac{d\Theta}{dt} = -\frac{w\Delta v}{(\Delta x)^2} \tag{2.6}$$

There is a threshold for $\frac{\Delta v}{(\Delta x)^2}$ in the literature (Brackstone and McDonald, 1999). By exceeding this threshold, the driver decelerates and adjusts its speed. Also, the other threshold for adjusting minimum space between two consecutive vehicles is considered. Although, it seems that this model simulates drivers' behaviour correctly and is able to explain what happens in reality, proving or disproving its validity is difficult and calibrating its individual elements and the thresholds has been rarely successful.

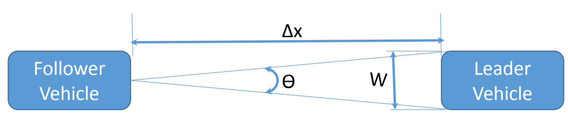


Figure 2.1 Visual angle and the related factors

Fuzzy logic-based model

Fuzzy logic-based models consider some logics in describing the driver behaviour (Brackstone and McDonald, 1999). For example, the term "too close" or "not close" can be defined and quantified. If the time space between two vehicle is less than 0.5 s, the separation is "too close" and its membership number is 1 and if it is more than 2 s, the separation is "not close" and its membership number is 0. Using Fuzzy output set we can define the pattern of driver behaviour by applying the mentioned sets by logical operators. This approach used to "fuzzify" the GHR model (Kikuchi and Chakroborty, 1992), and also the other models.

2.2.5 Comparison between traffic modelling frameworks

In this section, we compare the four frameworks in traffic modelling (CTM, CA, Queue-based, and Car-following) to find the most appropriate simulator to supply its outputs parallel to streaming data, that is at individual level, and feed them to the Bayesian approach that was chosen as the suitable approach in section 2.1.2 for this study.

Hofleitner et al. (2012) used CTM for developing a Bayesian model using streaming data in arterial travel time estimation. Also, Alecsandru (2006) tried to consider the effect of random driving behaviour at macro level and developed a CTM based approach to apply this behaviour. The following results can be mentioned from his study:

1. It is assumed that the randomness of the aggregate behaviour at a macroscopic level is the result of random behaviour of individual drivers. Therefore, random variation in the observed short-term capacity is the result of the randomness in minimum headway; and the random variation in local jam densities is the result of randomness in minimum space headway.

2. In CTM, by the modified flow advancing equation, the effect of randomness can be applied to the free flow speed, flow capacity, space capacity, and the speed of the backward moving wave.

3. Results of the case study by constant and dynamically varying wave speed show that there are some fluctuations in cell occupancies over time. However, in both cases the total travel time was approximately the same.

Stochastic CTM model has been developed using the Bayesian approach to estimate travel time and density in a network. But considering that we need to produce the outputs from the target traffic modelling framework at a microscopic level, using CTM and the stochastic form developed by Alecsandru (2006) is not a suitable choice for a traffic modelling framework.

Cellular automata (CA) considers the behaviour of drivers stochastically. The system that generates this stochasticity has not been developed enough. For instance, it contains some rules for increasing the speed of a vehicle and braking to avoid collisions. The stochasticity in the system, for example in braking events, is defined by a random number drawn from a uniform distribution that is compared with a stochastic noise parameter (called the "slow down probability") that defines the probability of decreasing vehicle speed. Therefore, from the behavioural point of view, CA modelling framework is not able to consider the probabilistic nature of driver behaviour accurately (Maerivoet and De Moor, 2005).

To apply this modelling framework, we need considerable computational time for simulation. Also, we mentioned that it is suitable for simulating parts of a network and special links. It is not computationally suitable for the entire target network. Therefore, using this framework is not recommended for generating the output we need to use with available streaming data. Queue-based modelling framework focuses on simulating events at intersections. Although, this approach gives us outputs at each car level, it can only model the intersections. Therefore, we cannot use it to feed its outputs with available streaming data that are at microscopic level.

Car-following models enable us to consider different patterns of drivers' behaviour in order

Table 2.2 Summary of comparison between four traffic modelling frameworks

Framework	Simulation level	Stochasticity	Computational limitation	Output compatibility
CTM	Dividing a link into cells	Is possible	Is not suitable in large scale	Cannot generate at microscopic level
CA	A cell for one vehicle	Is not realistic	Suitable for part of network	At microscopic level
Queue-based	Vehicle at intersection	Is possible	Suitable for large scale	At microscopic level
Car-following	Vehicle in link	Is possible	Suitable for large scale	At microscopic level

to increase the realism of a chosen model by applying motivational and attitudinal factors. This approach is confirmed by available evidence and it is more realistic framework in modelling traffic conditions due to these various patterns. But it can be noted that relating the mentioned features to observed dynamic behaviour has been rarely addressed in the existing literature (Brackstone and McDonald, 1999). Efforts have been made to develop these types of models. However, it can be concluded that car-following models can better capture the microscopic behaviour in links regarding available streaming data. The summary of these comparisons is presented in Table 2.2.

We propose using car-following modelling framework that is able to provide outputs at a microscopic level as the prior distribution in the Bayesian framework. It does not have the limitations of CTM and CA in computing ; and can consider stochasticity in modelling traffic conditions.

Therefore, applying microscopic car-following modelling framework can provide better results by considering individual behaviour and its effects on link parameters. We need to know what happens at low level in the network in order to consider the effects of streaming data that are at microscopic level on developing more realistic understanding about the urban road network.

2.3 Contribution of this study

It has been demonstrated that traffic conditions in urban network cannot be predicted with naive and deterministic models. For instance, patterns of traffic conditions differ for a given link in the same duration of time for two different working days of a week. Also, the patterns change from a working day of a month to an other month for the same season. These fluctuations cannot be simulated by a deterministic model. In the age of streaming and longitudinal data, the stochastic nature of traffic conditions needs to be captured by new models. Existing models simulate traffic conditions in urban networks using traditional data and their outputs for a network are fixed sets of link properties. The fluctuations in the patterns of traffic conditions dictates considerable uncertainty for using the outputs of a simulation as the urban network state.

Using the user-centric streaming data, we can create a more realistic image of network state in short-term. We can update the existing model that has been developed using traditional data by implementing a feedback from these new sources of non-traditional data. Therefore, the updated prediction of urban network state leads us to recommend a more realistic state. Figure 2.2 shows the contribution of this thesis. Based on Figure 2.2, we try to prove the ability of the proposed methodology for achieving more realistic understanding of traffic conditions in an urban network by fusing the outputs of simulation of various possible states of traffic conditions and streaming data using the Bayesian framework.

We can define link-by-link parameters of the urban road network stochastically. Moreover, the proposed method enables us to assess long term policies. Although, the output of this study is not an operational tool, it can provide us with the opportunity to think about developing one that would help us control and manage congestion in urban road networks dynamically and help users in optimizing their route choice decisions. The main output of this study as a prototype tool, proves us the capability of the proposed method for traffic conditions prediction by using streaming data parallel to the traffic simulation that is based on traditional data.

We propose a Bayesian approach to use the simulation outputs of various possible urban network states and available streaming data for defining the most probable state of the urban network. The next chapter mainly consist of: (1) Streaming data and improvement in traffic prediction methods; (2) Bayesian updating approach; (3) Structure of traffic model using car-following framework.

In the proposed Bayesian approach, the related probabilities of prior simulations and the probabilities of available measurements, streaming data, are defined. Also, the conditional models using the simulation outputs of various possible states and streaming data are mea-

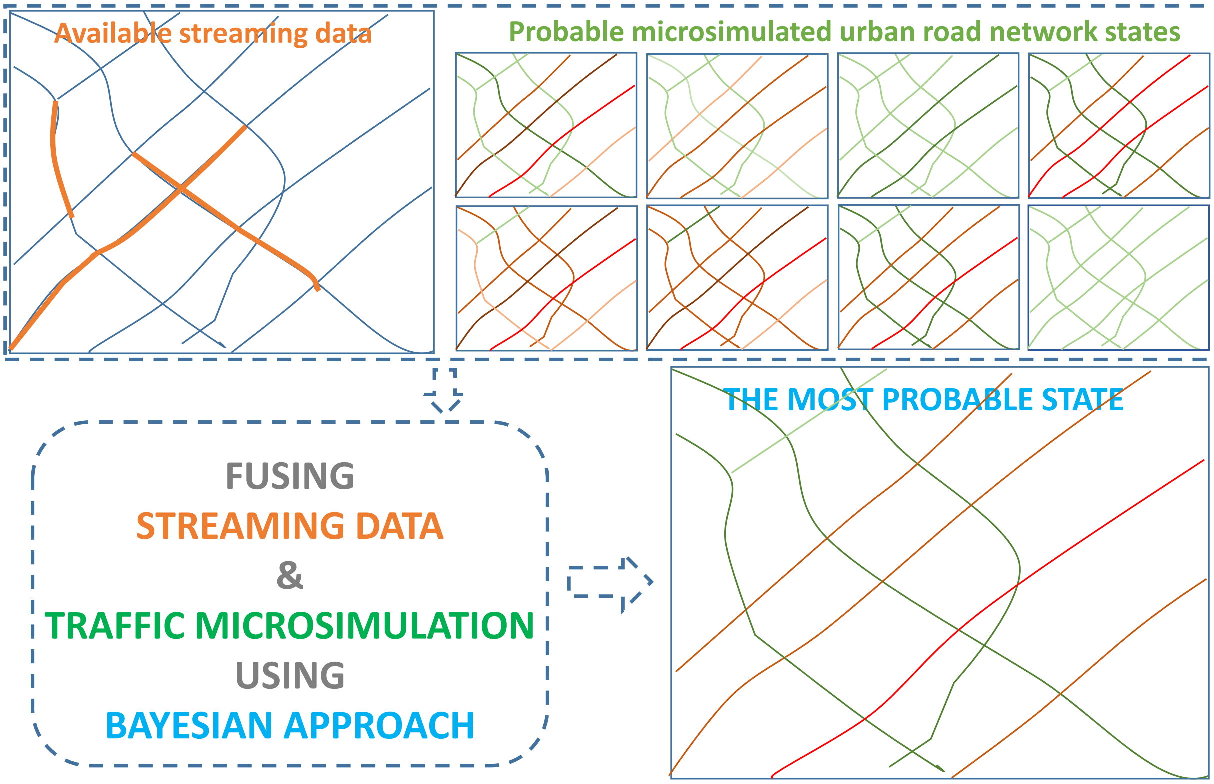


Figure 2.2 Statement of contribution

sured. Finally, using the proposed methodology we can define the probability of each possible state and predict the most probable urban network state at each time-step.

CHAPTER 3 METHODOLOGY

3.1 General framework

Implementing different strategies to manage congestion on urban networks, improving their states, and maximizing their capacities are possible by understanding traffic conditions and short-term forecasting in each link (Farooq, 2013). Figure 3.1 shows the procedure used in a conventional network modelling framework that only uses traditional data. A traffic model is calibrated and validated only once using traditional data and it is implemented for short-term prediction. We know that there are different patterns of traffic behaviour at the same time during different weekdays or different months of a season. Figure 3.2 shows the improved procedure with feedback by implementing recent events in the network that we try to cover it as much as possible on the frame of this thesis. Conventional method of traffic prediction (Figure 3.1) cannot consider the effects of these differences. By implementing streaming data available from sensors (Figure 3.2), we can consider the effect of the mentioned differences on traffic conditions in traffic conditions prediction methods that we propose by using a Bayesian approach and user-centric streaming data.

Streaming data as a non-traditional source of information provides us with the ability to

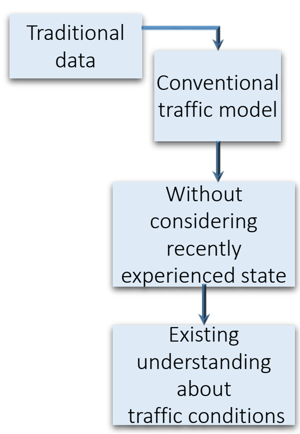


Figure 3.1 Conventional modelling framework without considering impacts of recently experienced traffic conditions, based on Furnish and Wignall (2009)

improve traffic conditions prediction methods. Probe vehicles (vehicles with any form of GPS devices) are able to generate streaming data required to use for this improvement. If there are enough probe vehicles that have in-vehicle devices, they can prepare real-time traffic information at an individual level. However, while only some vehicles have these devices,

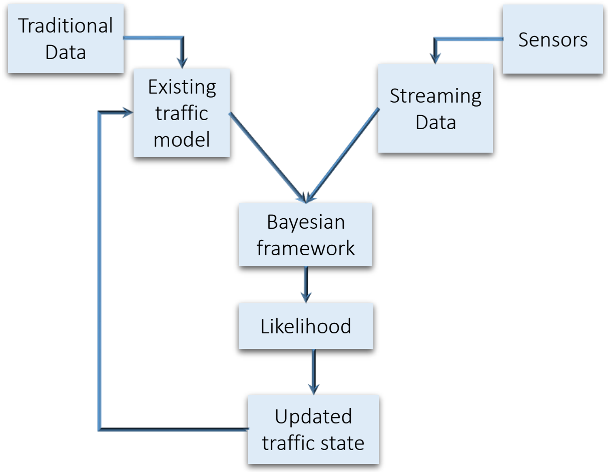


Figure 3.2 Improved network modelling framework using a Bayesian updating framework and user-centric streaming data for considering the impacts of recently experienced traffic conditions

accessing complete traffic indicators by just using this data is not very accurate. Therefore, there is a strong need to efficiently integrate and fuse streaming data with forecasting tools that use traditional data. This integration can help us to estimate the network state more precisely. It noticeably impacts the value of transportation system analysis.

There are direct and indirect benefits from using streaming data. Direct ones can be classified into congestion reduction, incident prediction, and travel time estimation. Strengthening the traffic simulation methodologies by developing, calibrating and validating new generation of models can be categorized as indirect benefits (Shi and Abdel-Aty, 2015).

Previously used tools like embedded sensors, for example loop detectors and camera technology, have considerable disadvantage that were mentioned in section 2.1.1. User-centric sources of data have remarkable advantages in comparison to other sources. For instance, by accessing the streaming data produced by probe vehicles, achieving a sample of speed in the network is possible (Farooq, 2013). Therefore, in this thesis, streaming data available from probe vehicles will be used to improve the prediction of traffic conditions.

Figure 3.2 shows the proposed procedure of traffic prediction with feedback from streaming data. By using streaming data and the outputs of a traffic model, we can consider the impact of recently experienced traffic conditions in a part of network on our understanding about traffic conditions in whole of the network. A Bayesian approach is proposed to use the outputs of a traffic simulation and streaming data. By defining likelihoods of all possible states of an urban road network we can recommend the one with a higher likelihood as the most probable state of the urban network at each time-step. The Bayesian updating approach is explained in the next section.

3.2 Bayesian updating approach

The Bayesian framework, which was introduced by Thomas Bayes is a powerful tool for interpreting a current event in the context of experienced events or recently available knowledge (Stone, 2013). Different studies have concluded that the Bayesian framework is suitable for effectively prediction of the complicated arterial conditions specifically the related hidden dynamism (Herring, 2010; Hofleitner et al., 2012; Nantes et al., 2013; Polson and Sokolov, 2014).
The reason that leads us to use this approach for updating our understanding about traffic conditions in an urban road network is its ability in relating the experienced events to current event. This approach enables us to consider posterior understandings of the network state and its probability regarding available measurements, streaming data, as the recently experienced events. By combining this approach with a traffic simulator that generates the different posterior understandings, we can predict traffic conditions of a whole network while we have no access to complete data of measurements in the network (Nantes et al., 2013).
This framework generally is explained by the following equation :

$$P(A \mid B) = \frac{P(B \mid A).P(A)}{P(B)} \tag{3.1}$$

Where A is the current event and B is the experienced event.
This framework can be used when there are new updates on traffic conditions. We can apply the effect of new information on updating the understanding of reality about traffic conditions as the most probable state of the urban road network.
Urban network combines two types of elements: links and intersections. In a network, at a certain time t, we are interested in determining the most probable state of the system i.e. traffic conditions in a road network link by link. The state of each link consists of three elements: flow, density, and speed. For a link i at time t, we can define the state variables as $S_t^i = (v_t^i, d_t^i, q_t^i)$. This state is the combination of three elements: v_t^i as speed, d_t^i as density, and q_t^i as the flow in link i at time t. The state of the network at time t can be defined as $S_t = (S_t^1, ..., S_t^i, ..., S_t^I)$. We consider S and $\hat{S}$ as the state of the urban network respectively from simulation and measurement that are the combination of the states of all links.
We want to find the probability distribution $P(S_t \mid \widehat{S}_t)$ (equation 3.2). Where $\widehat{S}_t$ is the measurement received as user-centric streaming data. In this condition, applying the Bayesian

framework discussed above is possible. We can write:

$$P(S_t \mid \widehat{S}_t) = \frac{P(\widehat{S}_t \mid S_t).P(S_t)}{P(\widehat{S}_t)} \tag{3.2}$$

Where $P(S_t)$ is the prior knowledge, $P(\widehat{S}_t \mid S_t)$ is the measurement model and $P(\widehat{S}_t)$ is the probability of measurements. This proposed approach mathematically encapsulate the processes in Figure 3.3.

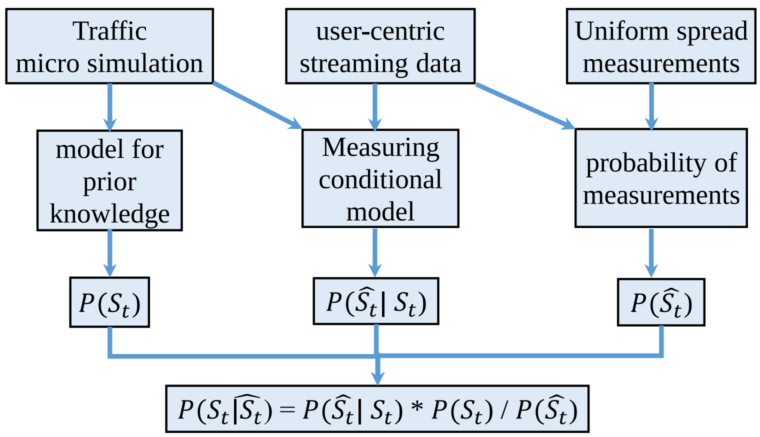


Figure 3.3 Defining the probability of a network state as prior knowledge considering available measurements using Bayesian framework

Based on Figure 3.3, we microsimulate traffic conditions in each possible state of urban network; implement user-centric streaming data; and assume uniform spread in measurements. This means that measurements are distributed by an equal time gap between each two consecutive measurements. This is a realistic assumption when each probe vehicle reports its features at the start and the end of a fixed time duration. Then, we define the probability of different existing traffic conditions in various possible states of urban networks; measure the conditional model; and define probability of measurements.

We follow the same basic approach developed by Danalet et al. (2014) on using Wi-Fi traces to predict pedestrians' activities on a campus. They considered the Bayesian framework for modeling the relationship between the probability of potential activity in the context of available streaming data as measurement about pedestrian activities. They considered that normal distribution is able to describe the errors in latitude and longitude of activity location. Also, they assumed that the errors in latitude and longitude are independent. $P(S_t)$ will be computed using historical datasets that we use to create a new model as traffic microsimulator, while $P(\widehat{S}_t \mid S_t)$ will be computed based on the GPS traces stream giving us the sample measurements at time t and microsimulation outputs.

3.3 Procedure for traffic conditions estimation

Figure 3.4 represents the steps and processes that we need to apply for defining the probability of each possible state of an urban road network and the most probable one as the recommended state at time t. We use this diagram as the main part of the proposed methodology for contribution of this thesis to implement streaming data and Bayesian approach for predicting urban road traffic conditions. We describe each part of this diagram briefly in following lines.

We define prior knowledge of traffic conditions by inputing possible travel time (TT) sets of

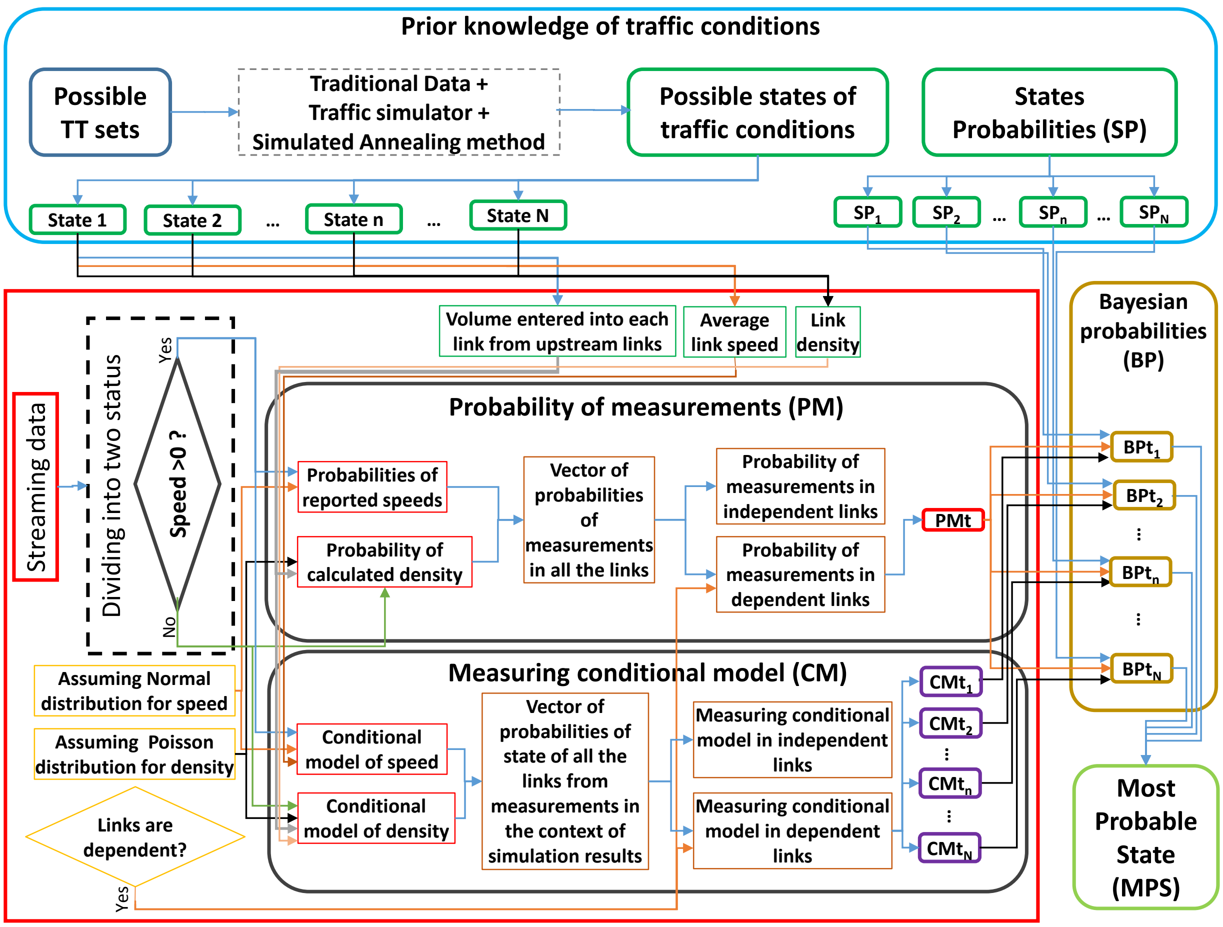


Figure 3.4 Flow diagram describing the steps and processes required for defining the most probable state of urban network at time t

the links in the urban network and traditional data into our traffic simulation and calibration tool. We simulate traffic conditions (using car-following framework as it will be illustrated in section 3.4) and calibrate the simulation by Simulated Annealing (SA) method [1] as it will be explained in section 3.5.

We assume N as the number of possible sets of TT in the urban road network. Therefore, we can simulate N possible states of traffic conditions in the network as our prior knowledge. Each state has a different probability regarding the possible sets of the observed TT in the network as historical dataset.

The part of diagram in Figure 3.4 located in the red box at the bottom left, shows the steps that we need to apply and the conditions we require to check for defining probabilities of measurements considering available streaming data and measuring conditional model related to each possible state.

When we process streaming data, we assume that the value of $\widehat{v}_t^{ji}$ (reported speed by GPS devices installed on probe vehicle j located in link i at time t) has two statuses. We assume that the state of link i can be defined by speed when the reported speed is non-zero and by density when the reported speed is zero. In first status, $\widehat{v}_t^{ji}$ is more than zero. While in second status, $\widehat{v}_t^{ji}$ equals zero.

We assume that in the first status, the flow and density can be defined as a function of speed, as it will be illustrated in more detail in section 3.6. In this status, the state of each link can be defined only by using speed. The speed data reported by an individual vehicle is normally distributed in arterials (Dhamaniya and Chandra, 2013). Therefore, we assume that Normal distribution can describe the reported speeds by each probe vehicle available as an element of the source of streaming data in the first status.

At each time t, we check the value of $\widehat{v}_t^{ji}$. If $\widehat{v}_t^{ji}$ is more than zero, we define the probability of reported speed using streaming data, as it will be explained in section 3.8. Also, we measure the conditional model of speed using streaming data and average link speed available from calibrated simulation in each possible state of traffic conditions in the network. As it will be illustrated in section 3.9.

There are some situations where the value of $\widehat{v}_t^{ji}$ equals zero, but density of link i does not equal to the jam density. More detail about this status will be explained in section 3.7. In these situations, the second status happens and we assume that only definition of density is enough for defining state of link i. In this status, we can define the value of density from measurements and simulation results, as it will be illustrated in section 3.7

We need to consider the rate of arrival to the queue in the second status. We do not access to density data directly from measurements. Therefore, we calculate density and we assume that the calculated density is Poisson distributed. With this assumption for density, we define

1. This method is used for approximately defining the global optimum of a function using probabilistic metaheuristic control. We use this method to be sure about defining global optimum and not choosing a local optimum.

probability of calculated density using streaming data and the volumes entering into each link from the upstream links, as it will be explained in section 3.8. Also, we measure conditional model of density using streaming data, the volume entering into each link from the upstream links, and link density available from calibrated simulation in each possible state of traffic conditions in the network, as it will be illustrated in section 3.9.

There are N possible states of traffic conditions at time t (S_{t1}, ..., S_{tn}, ..., S_{tN}) and each of them has different probabilities (SP_1, ..., SP_n, ..., SP_N) that can be defined using different possible sets of TT in the network from the historical dataset. Considering the dependency between links (section 3.10), the probability of measurements at time t (PM_t) is defined as it will be explained in sections 3.8 and 3.9. Also, we define N conditional models for N possible states of traffic conditions in the network at time t (CM_{t1}, ..., CM_{tn}, ..., CM_{tN}) applying the dependency between the links.

In the right hand side of diagram in Figure 3.4, we define Bayesian probabilities (BP_{t1}, ..., BP_{tn}, ..., BP_{tN}) using N probabilities of possible states (SP_1, ..., SP_n, ..., SP_N), the probability of measurements at time t (PM_t), and N conditional models at time t (CM_{t1}, ..., CM_{tn}, ..., CM_{tN}) by applying the Bayesian framework and considering the effect of dependency between connected links. Finally, we recommend the Most Probable State (MPS) of traffic conditions in the urban road network at time t.

The expanded description of the steps and processes shown in Figure 3.4 briefly explained in this section, will be mentioned in the next sections.

3.4 Prior car-following traffic model

Figure 3.5 shows an urban road network that can be monitored at links or intersections.

For microsimulating the events in an urban network we use the car-following framework that monitors traffic conditions at microscopic level, considering each car behaviour in links. This level of simulation helps us to combine it with the user-centric streaming data that are at an individual level. Using the outputs of this microsimulator for different possible states of the urban network, we can define the speed distribution for all vehicles in each link, the volume of vehicles entering into each link between each pair of time-steps, as it will be explained in section 3.5.2, the average link's speed and link's density available from the traffic simulation-calibration tool in all possible states of traffic conditions.

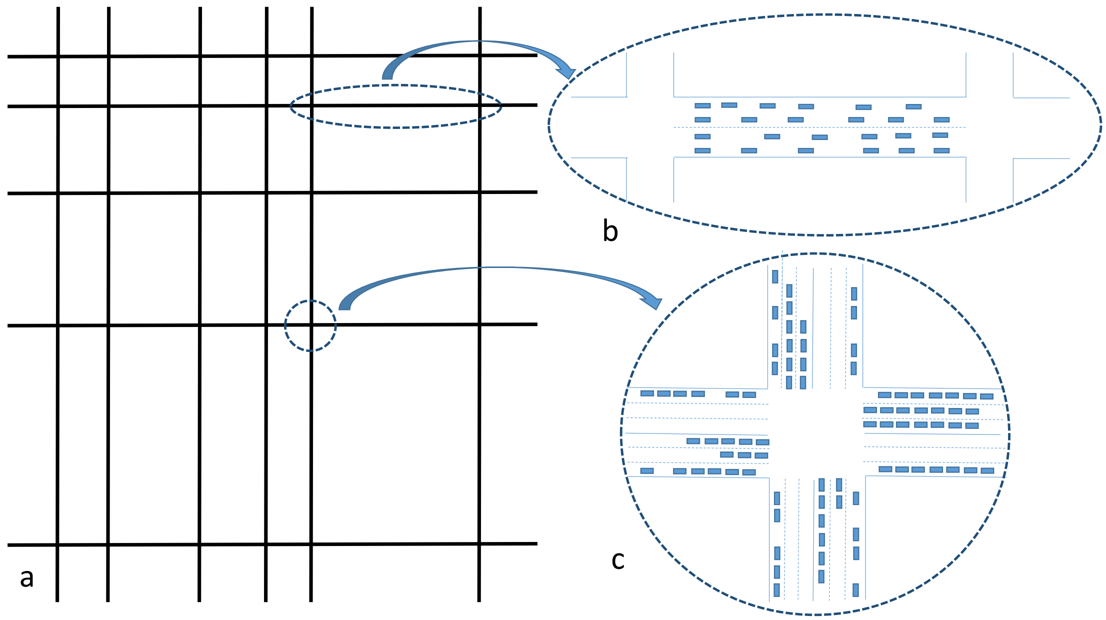


Figure 3.5 Modelling the urban road network (a) at its two elements: Links (b) and intersections (c)

3.5 Traffic simulation and calibration tool

We use SUMO (Simulation of Urban MObility) for traffic simulation that is a free and open traffic simulation package and is available since 2001 (Krajzewicz et al., 2012). As we are interested to define the capability of the proposed Bayesian framework, we adapt the code developed by Jandik (2015).

We develop different states of traffic conditions in the urban network as the possible states and simulate traffic conditions based on different available data in each state using the traffic simulation package (SUMO software). We apply the code developed at Laboratory of Innovations in Transportation, Polytechnique Montréal, to implement SA method for calibrating each possible state based on different travel times experienced in the urban network of the study area.

For more description about implementing SUMO and SA method respectively in traffic simulation and model calibration, refer to Jandik (2015). The steps implemented to feed the possible TT sets to simulation-calibration procedure using SUMO and SA method are presented in Figure 3.6. There are dynamic and static parts for inputing external demand to the traffic simulator. Each simulated state is calibrated by inputting each set of possible TTs in the urban network and the TT outputs of simulation. If the acceptance condition of SA method is not satisfied, the dynamic part of external demand is changed and the procedure

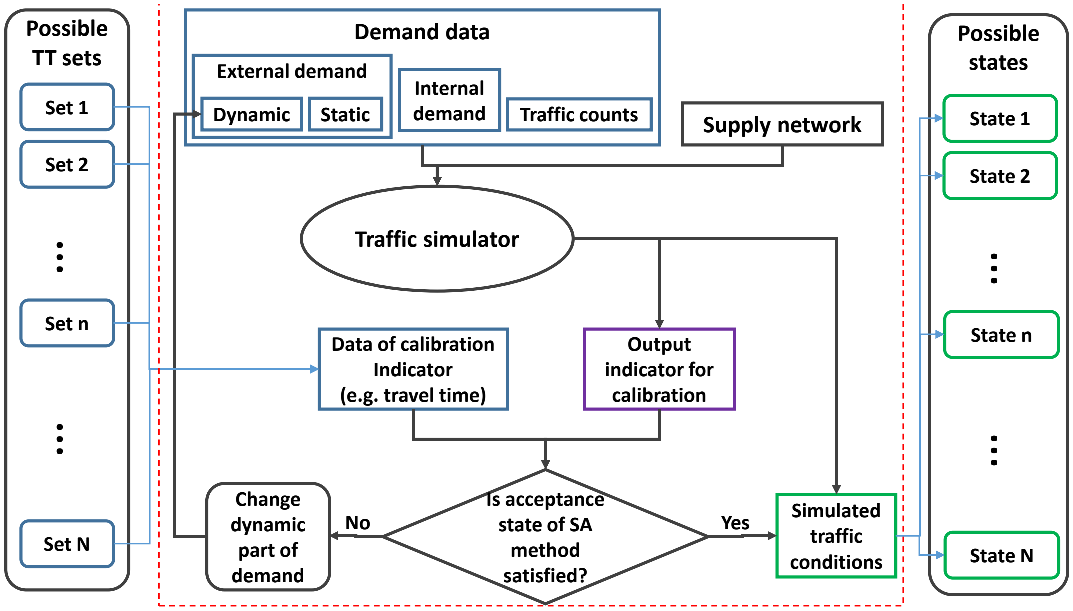


Figure 3.6 The steps used in the process of simulating possible states of traffic conditions by SUMO and calibrating by SA method

continues until the acceptance condition is satisfied. Finally, the simulation results that satisfy the acceptance condition will be introduced as the possible state. The procedure shown in the red box of the diagram in Figure 3.6 is repeated by inputting all N possible sets of TTs and the simulation outputs are N possible states of traffic conditions in the urban road network.

The innovation and novelty in this study is applying the Bayesian framework for defining the most probable state of urban network among the possible states of the network based on available streaming data.

3.5.1 Lane changing effects

To model traffic conditions in an urban network we need to consider the effects of lane changing as well. There are specific models that try to assess the behaviour of drivers in choosing available lanes. We are interested to test the usefulness of the Bayesian framework in predicting traffic conditions and we use SUMO for traffic microsimulation. Lane changing effects are modeled on the frame, limitation and ability of this car-following traffic microsimulator to consider these effects.

3.5.2 Volume entering into each link from upstream

When there is a probe vehicle with the reported speed of zero in link i, we can define the volume of vehicles that enter into link i from upstream links between the first time of reporting a speed of zero and each time before reporting the speed of none-zero by the probe vehicle j using the outputs from simulation.
Using SUMO traffic microsimulator we know the position and speed of each vehicle at each time-step. Also, we can use the outputs of simulation by SUMO to define the volume that enters into each link between each pair of time-steps.
Figure 3.7 shows the link state at three times t_0^{ji}, t'^{ji}, and t. We know the time t_0^{ji} and location x_0^{ji} respectively as the time and position that the speed of probe vehicle j decreases to zero in link i. We can define the time (t'^{ji}) when probe vehicle j enters into the link i using its reported trajectory. Therefore, the volume entering into each link i between t'^{ji} and time t, that is between time t_0^{ji} and the time that the probe vehicle increases its speed from zero, can be defined by SUMO traffic microsimulator.

3.6 Link state definition using speed

LWR models are considered to be the most reasonable approximation of the macroscopic behaviour of traffic flow. They are able to reproduce phenomena like shock waves and queue formation and dissipation in both congested and un-congested regimes. The vehicle conservation equation must be satisfied, such that:

$$q_x + d_t = 0 \tag{3.3}$$

Where q_x is the rate of change of flow with respect to space, and d_t is the rate of change of density with respect to time.
We assume that when the value of speed at each link is more than zero, we can define its density and flow regarding the identifiable relationship as shown in Figure 3.8. In the status that $\widehat{v}_t^{ji}$ is non-zero, considering the relationships between flow (q), density (d), and speed (v), we can estimate the value of d and q for link i. We consider q and d as a function of v in this status. Hence, we can assume that the state of link i can be defined only by $\widehat{v}_t^{ji}$ when its value is more than zero. Therefore, we can assume that the probability of state of link i equals to the probability of its speed:

$$P({S_t}^i) = P({v_t}^i) \tag{3.4}$$

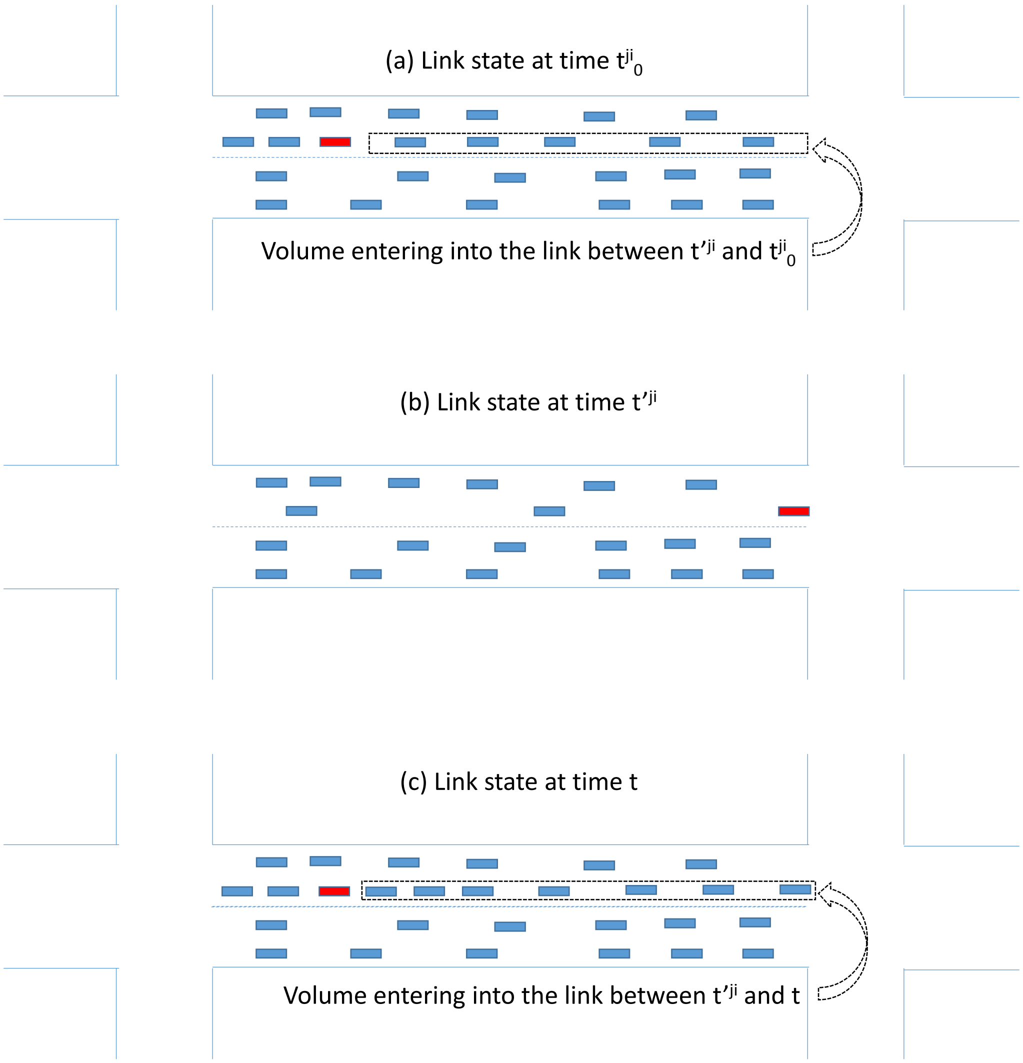


Figure 3.7 Link state at three times t_0^{ji} (a), t'^{ji} (b), and t (c) and the volume entering into each link between t'^{ji} and t_0^{ji} (a) and t'^{ji} and t (c)

3.7 Link state definition using density

In the status that $\widehat{v}_t^{ji}$ equals zero, flow equals zero and density can be between zero and jam density (Figure 3.9). Therefore, only density definition can help us to define the state of link i at time t. In this status, we can define d_t^i, density of link i at time t, using following procedure:

1. Using map-matching, we can define to which link at time t each probe vehicle relates.
2. When $\widehat{v}_t^{ji}$ equals zero, we can define the number of vehicles in front of the probe vehicle j in the queue using its latitude and longitude information.
3. To define the number of vehicles that will come after the probe vehicle j in link i

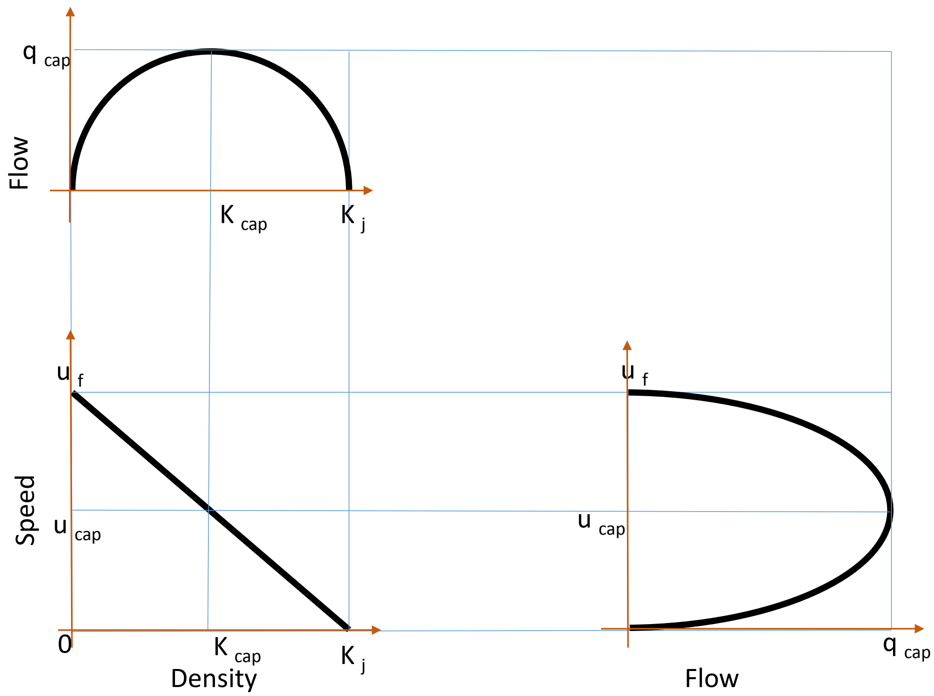


Figure 3.8 Fundamental diagram when the value of speed is non-zero.

between time t'^{ji} and time t, we can assume that during remaining red time of the signal, the turning volumes from the upstream links to link i is used from simulation. Therefore, we can define the link density during red time of signal.

4. We can postulate that the state of each link when the speed of its related probe vehicle is zero, is defined by one independent parameter: the density that varies during red time of the signal.

5. The probability of state of link i equals to the probability of the calculated density:

$$P({S_t}^i) = P({d_t}^i) \tag{3.5}$$

3.8 Defining probability of measurements

First of all, we assume that measurements at time t reported by each probe vehicle j are independent of what were reported by other vehicles.
We can consider ΔT period before time t as the period that we call it the reference period for defining the probability of measurements. If during this period the position of probe vehicle j that is reported, shows that the reported trajectory is related to link i, we use it to define the probability of measurements in link i.

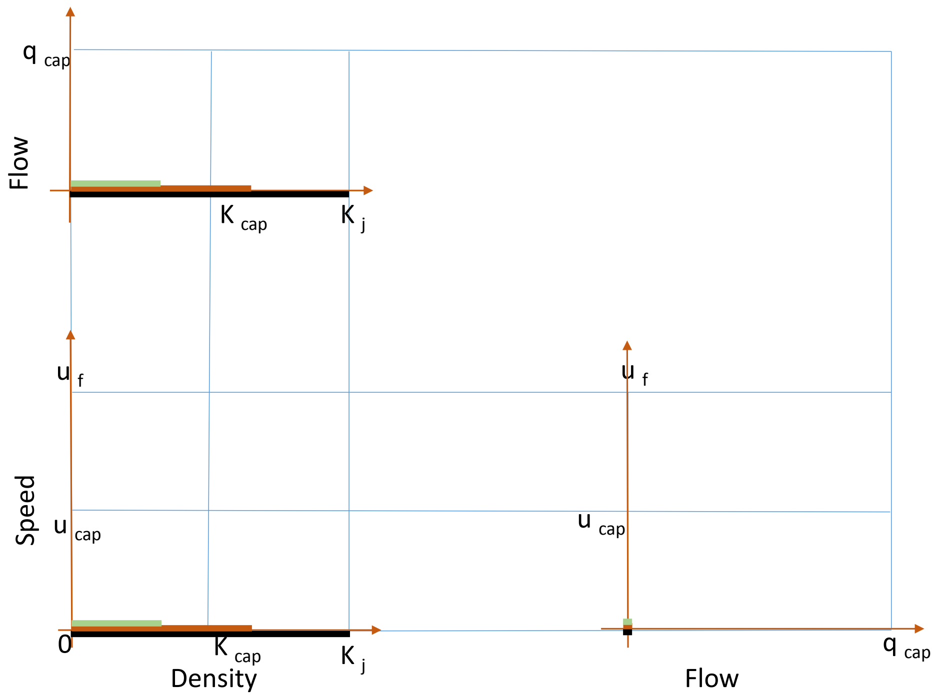


Figure 3.9 Fundamental diagram when the value of speed is zero.

If the trajectory categorizes as the first status, we use only the reported speed. Otherwise, for the second status, we use the calculated density to define the probability of measurements. In the first status, with the assumption of Normal distribution for speed reported between $t - \Delta T$ and t that described in section 3.3, we can write:

$$P(\hat{v}_t^i) = \frac{1}{(\sigma_{\hat{v}^i_{t-\Delta T,t}})\sqrt{2\pi}} e^{-\frac{(\hat{v}_t^i - \mu_{\hat{v}^i_{t-\Delta T,t}})^2}{2(\sigma_{\hat{v}^i_{t-\Delta T,t}})^2}} \tag{3.6}$$

Where $\sigma_{\hat{v}^i_{t-\Delta T,t}}$ is the standard deviation of measured speed and $\mu_{\hat{v}^i_{t-\Delta T,t}}$ is the average of measured speed between $t - \Delta T$ and t for link i.

We are interested in dense urban networks. In this type of network, the link sizes are small (50 m to 500 m). So the speed variance within link is small and this average ($\mu_{\hat{v}^i_{t-\Delta T,t}}$) can be representative of the link average. If there is more than one probe vehicle in link i, we can consider average of their speeds as $\hat{v}_t^i$ and the rest of calculation is the same as above.

In the second status, with the assumption of Poisson distribution for the calculated density between $t - \Delta T$ and t that described in section 3.3, we can write:

$$t^\star = \frac{x_0^{ji} - X^{Di}}{L * S_r} + t_0^{ji} + \Delta t_r \tag{3.7}$$

If $t_G < t < t^\star$ then $\hat{d}_t^i = \frac{x_0^{ji} - X^{Di}}{L} + Q_{Ui} - S_r.(t_G - t_0^{ji} - \Delta t_r)$
If $t < t^\star \& t < t_G$ then $\hat{d}_t^i = \frac{x_0^{ji} - X^{Di}}{L} + Q_{Ui} - S_r.(t - t_0^{ji} - \Delta t_r)$

$$P(\hat{d}_t^i) = \lambda^{\hat{d}_t^i}.\frac{e^{-\lambda}}{\hat{d}_t^i!} \tag{3.8}$$

Where $t^\star$ is the dissipation time of the queue located ahead of the probe vehicle, x_0^{ji} is the location of the probe vehicle when its speed decreases to zero, X^{Di} represents the coordinate of the intersection located downstream of the link i, L is the effective average length of vehicle, S_r is the uniform service rate of intersection, t_0^{ji} is the time that the probe vehicle starts reporting its speed as zero, Δt_r is the remaining time of the red time of the signal located at downstream after the time that the probe vehicle starts reporting its speed as zero, t_G is the time that the first next green time changes to red, Q_{Ui} is the volume of arrived vehicles after the probe vehicle from upstream of the link i, and λ is the density mean of link i between t_0^{ji} and t. We need to calibrate value of λ between t_0^{ji} and t.
If there is more than one probe vehicle, in the second status, for simplification, we only consider the first probe vehicle that has started to report its speed as zero.
Regarding the fact that probe vehicles are randomly distributed, we assume that the reported trajectories from probe vehicles at time t are independent. Therefore, we can write:

$$P(\hat{S}_t) = \Pi_{i=1}^{I} P(\hat{S}_t^i) = \Pi_{i=1}^{I} P(\hat{v}_t^i) or P(\hat{d}_t^i) \tag{3.9}$$

3.9 Measuring conditional model

When speed is none-zero, the values of density and flow are dependent upon the value of v. Considering the fundamental relationship between d and v and also between q and v, the state of each link from car-following based microsimulation can be defined only by v. In this situation, that is the first status, we can define the probability of state of link i from measurements given the state of link i from simulation by defining the speed probability of link i from measurements given the speed probability of link i from simulation. Therefore, we can write:

$$P(\hat{v}_t^i \mid {v_t}^i) = \frac{1}{(\sigma_{\hat{v}^i})\sqrt{2\pi}} e^{-\frac{(v_t^i - \hat{v}_t^i)^2}{2(\sigma_{\hat{v}^i})^2}} \tag{3.10}$$

Where $\sigma_{\hat{v}^i}$ is the standard deviation of the measured speed in link i that is defined by analyzing the level of confidence of x and y coordinates of probe vehicles.
When speed is zero, the values of density may or may not be zero but flow is zero. In

this situation, that is the second status, we can define the state probability of link i from measurements, given the state of link i from simulation by defining the probability of density of link i from measurements, given the probability of density of link i from simulation. Therefore, we can write:

$$P(\hat{d}_t^i \mid {d_t}^i) = \lambda^{{d_t}^i} . \frac{e^{-\lambda}}{{d_t}^i!} \tag{3.11}$$

If we can assume that the measured and simulated speeds in two connected links are independent, we can write:

$$P(\hat{S}_t \mid S_t) = \Pi_{i=1}^{I} P(\hat{S}_t^i \mid S_t^i) = \Pi_{i=1}^{I} P(\hat{v}_t^i \mid v_t^i) or P(\hat{d}_t^i \mid d_t^i) \tag{3.12}$$

This assumption leads us to predict the probabilities to be less than what it is in reality. We know that the probability of two dependent events is less than their joint probabilities if we assume them independent. The difference between these two probabilities varies from each pair of events to the other pair. In this problem, we need to consider probabilities of links' states in urban network where each of them is connected to at least two other links. Also, each link has some effects from the links that are not connected to them directly. For instance, the impact of an accident in a link can be seen in an indirectly connected link after some seconds. Therefore, we drop this assumption and consider the effect of this dependency.

3.10 Effect of dependency between connected links

The state of each link at each time t can be considered dependent upon other links. For example, a link can be crowded at time t, if one of the connected links has been crowded at the time before t. Considering a simple example (Figure 3.10), we have a link i_1 that is connected to other links i_2, i_3, i_4, i_5, i_6, i_7, i_8, and i_9 through 4 intersections.

We consider a matrix of dependency degree between links. As a simplifying assumption,

.

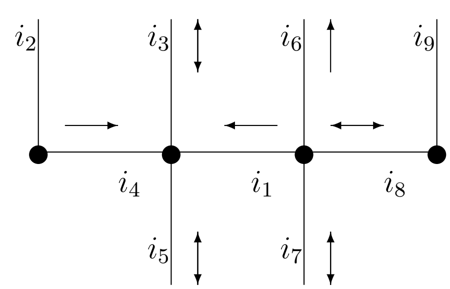


Figure 3.10 A simple demonstration of link i_1 connected to the other links through 4 intersections

when two links are connected directly (e.g., link i_1 and i_3), indirectly with a link (e.g., link

i_4 and i_8), indirectly with two links (e.g., link i_2 and i_6), and indirectly with more than two links (e.g., link i_2 and i_9), their degrees of dependency will be respectively 1, d_1, d_2, and 0. Where d_1 and d_2 are between 0 and 1.

To generalize, a dependency matrix based on direction of traffic flow can be developed. If

$$\begin{array}{cc} & \begin{array}{ccccccccc} i_2 & i_3 & i_4 & i_5 & i_1 & i_6 & i_7 & i_8 & i_9 \end{array} \\ \begin{array}{c} i_2 \\ i_3 \\ i_4 \\ i_5 \\ i_1 \\ i_6 \\ i_7 \\ i_8 \\ i_9 \end{array} & \begin{bmatrix} 1 & 0.5 & 1 & 0.5 & 0.5 & 0.25 & 0.25 & 0.25 & 0 \\ 0.5 & 1 & 1 & 1 & 1 & 0.5 & 0.5 & 0.5 & 0.25 \\ 1 & 1 & 1 & 1 & 1 & 0.5 & 0.5 & 0.5 & 0.25 \\ 0.5 & 1 & 1 & 1 & 1 & 0.5 & 0.5 & 0.5 & 0.25 \\ 0.5 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0.5 \\ 0.25 & 0.5 & 0.5 & 0.5 & 1 & 1 & 1 & 1 & 0.5 \\ 0.25 & 0.5 & 0.5 & 0.5 & 1 & 1 & 1 & 1 & 0.5 \\ 0.25 & 0.5 & 0.5 & 0.5 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0.25 & 0.25 & 0.25 & 0.5 & 0.5 & 0.5 & 1 & 1 \end{bmatrix} \end{array}$$

Degree

$$\begin{array}{cc} & \begin{array}{ccccccccc} i_2 & i_3 & i_4 & i_5 & i_1 & i_6 & i_7 & i_8 & i_9 \end{array} \\ \begin{array}{c} i_2 \\ i_3 \\ i_4 \\ i_5 \\ i_1 \\ i_6 \\ i_7 \\ i_8 \\ i_9 \end{array} & \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \end{array}$$

Direction

$$\begin{array}{cc} & \begin{array}{ccccccccc} i_2 & i_3 & i_4 & i_5 & i_1 & i_6 & i_7 & i_8 & i_9 \end{array} \\ \begin{array}{c} i_2 \\ i_3 \\ i_4 \\ i_5 \\ i_1 \\ i_6 \\ i_7 \\ i_8 \\ i_9 \end{array} & \begin{bmatrix} 0.5 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0.25 & 0.25 & 0.25 & 0 & 0 & 0 & 0 \\ 0.5 & 0.25 & 0.25 & 0.25 & 0.25 & 0 & 0 & 0 & 0 \\ 0 & 0.33 & 0.33 & 0.33 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.2 & 0 & 0.2 & 0.2 & 0 & 0.2 & 0.2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.33 & 0.33 & 0.33 & 0 \\ 0 & 0 & 0 & 0 & 0.25 & 0.25 & 0.25 & 0.25 & 0 \\ 0 & 0 & 0 & 0 & 0.25 & 0.25 & 0.25 & 0.25 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \end{bmatrix} \end{array}$$

Effect

Figure 3.11 Related matrices for the simple network presented in Figre 3.10

based on the direction of traffic flow in two directly connected links, there is a possibility of moving traffic between links, direction dependency will be 1, otherwise it is 0.

We need to consider the effect of the number of connected links to a specific link. For example, traffic conditions experienced in link i_1 can be affected by experienced conditions in link i_1, i_3, i_5, i_7, and i_8. Although, we know that this effect is not exactly the same for all contributing links, for simplification, we can consider the same effect. Therefore, we can consider the

adjustment matrix.
To define the state probability of each link we need to multiply the probability vector (that is defined using different probabilities for all the links using procedures illustrated in sections 3.8 and 3.9) by each matrix. For instance, Figure 3.11 presents related matrices for the simple network shown in Figure 3.10. In this example, we assumed the value of d_1 and d_2 respectively 0.5 and 0.25.

3.11 Algorithms for implementing the proposed methodology

Algorithms that are used in implementing the proposed methodology in order to define the probability of measurement and measure the conditional model in each link at each time t are shown in next two pages.
We use the following two algorithms in the case study that is described in the next chapter. These algorithms implement the discussion in the previous sections of this chapter.

```
Data: GPS streaming data that contain latitude, longitude and speed of probe vehices at each
      time t
Result: Probability of measuremnts at time t for each link
initialization;
for each time t do
    for each probe vehicle do
        if speed of probe vehicle > 0 then
            for each link do
                if latitude and longitude of the probe vehicle equal latitude and longitude of the
                begining of link then ;
                consider t as the time of entering the probe vehicle into the link;
                if latitude and longitude of the probe vehicle are between the latitude and
                longitude of link then ;
                define the average speed of the probe vehicles located in each link;
            end
        else
            for each link do
                consider the probe vehicle that is near the end of each link, as the probe vehicle
                to which we need to define the latitude and longitude of the vehicle when there
                is a reported speed of zero;
                if latitude and longitude of the probe vehicle equal the latitude and longitude of
                the begining of the link then
                    consider t as the time of entering the probe vehicle into the link;
                else
                    define the time of entering the probe vehicle into the link by considering its
                    average speed as half of the link's free flow speed (FFS);
                end
            end
        end
    end
    define the average and standard deviation experienced for the average reported speed of
    non-zero between t − Δt and t in each link;
    define the probability of measurements at time t for each link considering the Normal
    distribution for average reported speeds of none-zero between t − Δt and t in each link;
    run SUMO to define the volume entered into each link between the time of entering the
    probe vehicle into the link and time t;
    define the calculated density in each link at time t and mean density between the time of
    entering the probe vehicle into the link and time t, that is less than the time of increasing the
    speed of the probe vehicle from zero;
    define the probability of measurements at time t for each link considering the Poisson
    distribution for rate of arrival;
    define the vector of probability of measurements at time t for all the links;
end
```

Algorithm 1: Defining the probability of measurements at time t for each link

```
Data: GPS streaming data that contain latitude, longitude and speed of probe vehices at each
      time t and simulation results
Result: Conditinal model at time t for each link
initialization;
for each time t do
    for each probe vehicle do
        if speed of probe vehicle > 0 then
            for each link do
                if latitude and longitude of probe vehicle equal latitude and longitude of the
                begining of the link then ;
                consider t as the time of entering the probe vehicle into the link;
                if the latitude and longitude of the probe vehicle are between the latitude and
                longitude of the link then ;
                define average speed of the probe vehicles located in each link;
            end
        else
            for each link do
                consider the probe vehicle that is near the end of each link, as the probe vehicle
                to which we need to define the latitude and longitude of the vehicle when there
                is a reported speed of zero;
                if the latitude and longitude of probe vehicle equal the latitude and longitude of
                the begining of the link then
                    consider t as the time of entering the probe vehicle into the link;
                else
                    define the time of entering the probe vehicle into the link by considering its
                    average speed as half of the link FFS;
                end
            end
        end
    end
    run SUMO to define the average speed and density in each link at time t;
    define average and standard deviation experienced for average reported speed of non-zero
    between t − Δt and t in each link;
    measure conditional model at time t for each link considering the Normal distribution for
    average reported speeds of none-zero between t − Δt and t in each link;
    run SUMO to define volume entered into each link between the time of entering the probe
    vehicle into the link and time t;
    define the calculated density in each link at time t and mean density between the time of
    entering the probe vehicle into the link and time t, that is less than the time of increasing the
    speed of the probe vehicle from zero;
    measure the conditional model at time t for each link considering the Poisson distribution for
    rate of arrival;
    define the vector of probabilities of states of all the links at time t from measurements in the
    context of simulation results;
end
```

Algorithm 2: Measuring the conditional model at time t for each link

CHAPTER 4 CASE STUDY, DATA AND DESCRIPTIVE ANALYSIS

4.1 Case study

As a case study, a subsample of Montréal's road network is developed. The case study area of this research is located in the city of Montréal. This area consists of two parallel routes. The first route, in the west, combines Boulevard de l'Acadie and Avenue du Parc. These two arterials are connected by Avenue Beaumont as well as Rue Jean-Talon-Ouest (Brown lines in Figure 4.1). The second route contains two parallel streets (Boulevard St. Laurent and a combination of Boulevard St. Laurent, Rue Clark, and Rue St. Urbain) (Red and blue lines in Figure 4.1). These streets act as two-way arterials between the Autoroute Métropolitaine and Rue de Castelnau and as two one-way streets, after Rue de Castelnau in the south. The motivation for selecting this area is that we can test the usage of the proposed method in a more complex traffic state. Also, we can study the routes that are connected to Montréal downtown area and experience recurrent traffic jams during peak hours.

Mont Royal hill has blocked the access to Montréal downtown from Chemin de la Cote-

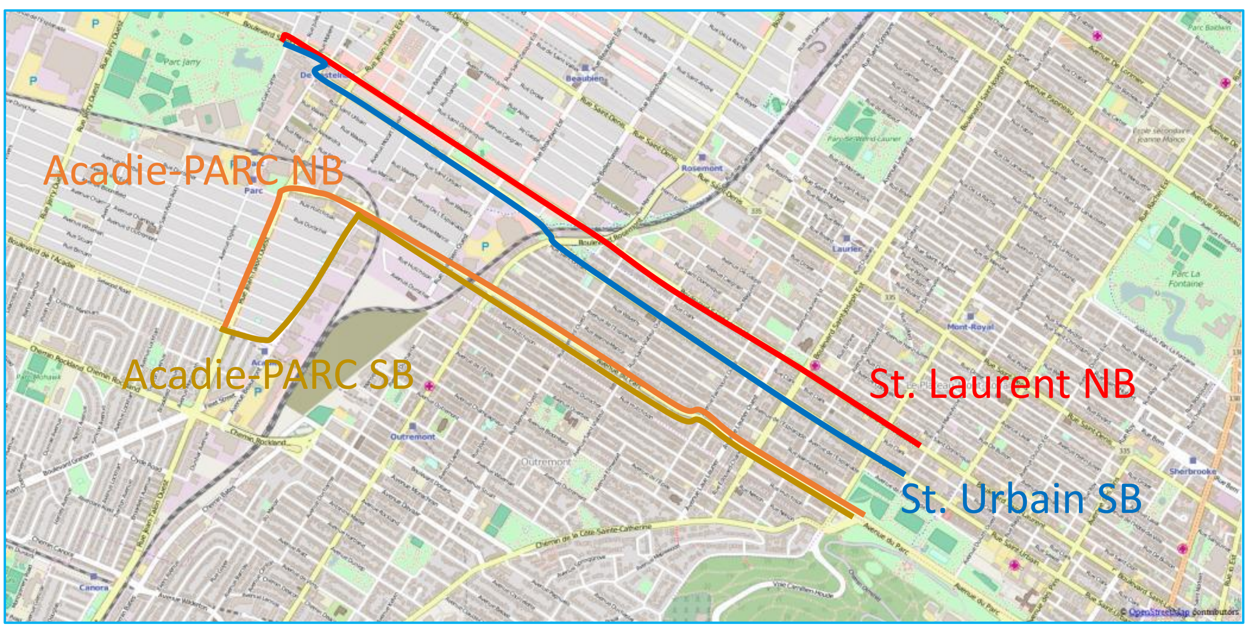


Figure 4.1 Case study area

des-Neiges to Avenue du Parc (Figure 4.2). The chosen routes provide the shortest access to downtown area at the east of the Mont Royal hill. The high volume of vehicles commuting to and from downtown, experience considerable traffic jams during the morning and afternoon peak hours. Therefore, the selected routes are interesting for traffic study and testing the applicability of our proposed methodology.

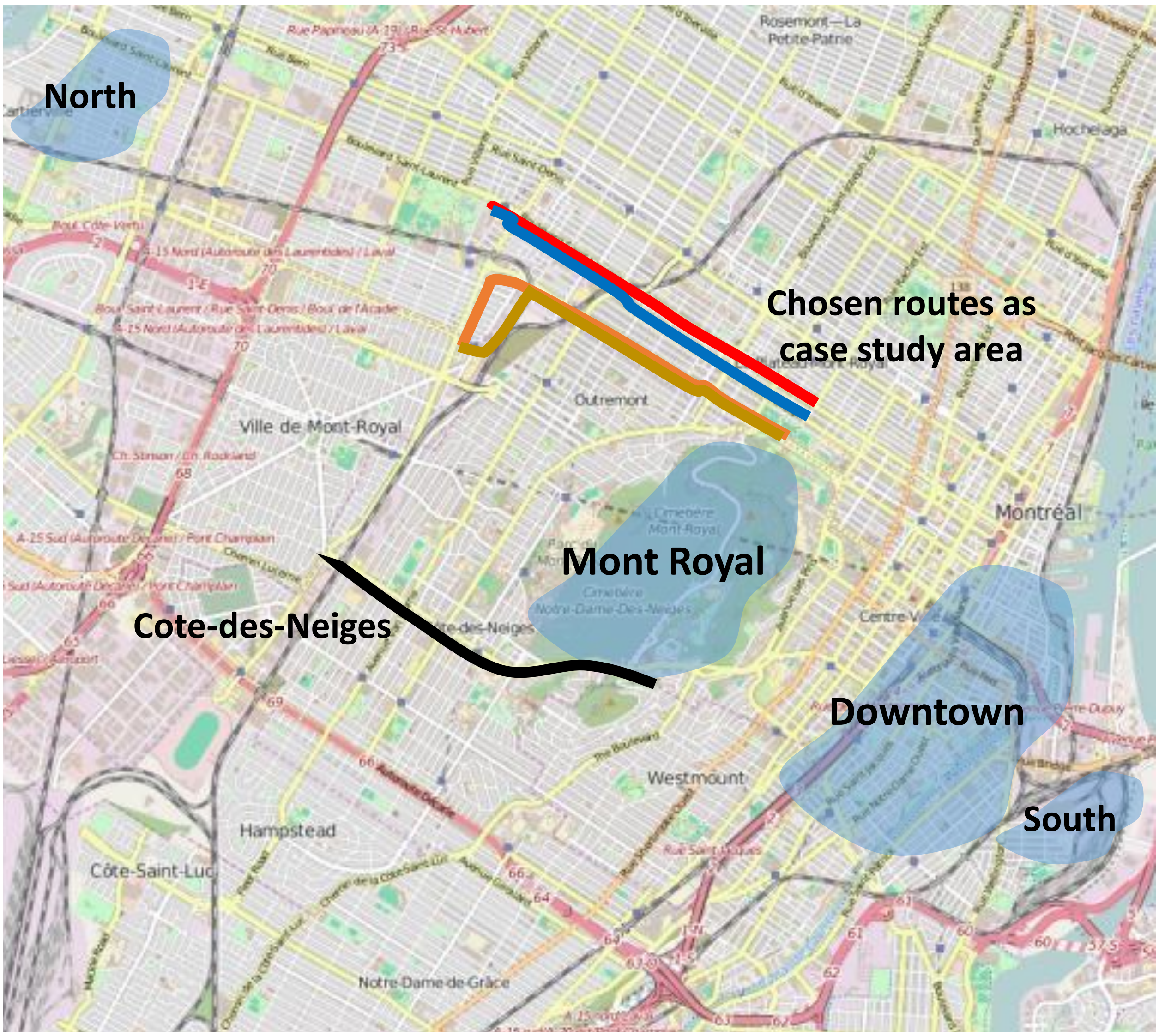


Figure 4.2 Chosen routes as the east access to Montréal downtown considering the blockage by Mont Royal hill

4.1.1 Detailed characteristics of the chosen routes

We divide the chosen routes into different links in traffic simulation. Therefore, we need to define their characteristics (length, number of lanes, parking facility, and Free Flow Speed

(FFS) [1]). Table 4.1 shows a summary of the important characteristics.
In Acadie-Parc SB direction, the minimum and maximum link lengths are respectively 30 m and 539 m and the average length is 263 m. Also, average FFS is 52 km/h. In Acadie-Parc NB direction, the minimum length is twice, while the maximum and average length are around the values of these parameters in Acadie-Parc SB direction. Also, average FFS is 48 km/h. Generally, it can be concluded that while distances between intersections in Acadie-Parc SB direction are less than the related values in Acadie-Parc NB direction, the average value of FFS is higher in SB direction in comparison to NB direction. It can be related to different patterns of the signalized intersections in two directions in Acadie-Parc route.
In St. Laurent NB direction, the average link length is 241 m while in St. Urbain SB direction, it is 270 m. In St. Urbain SB direction, all the minimum, average, and maximum link lengths are higher than these values for St. Laurent NB direction. Average FFS in St. Urbain SB direction is 47 km/h while it is 54 km/h in St. Laurent NB direction. These characteristics are provided in detail in 4 tables in Appendix A.

Table 4.1 Summary of links' characteristic in the chosen routes

Direction	Node		Length			No. of Lanes			FFS	
		Min	Ave	Max	1	2	3	Min	Ave	Max
Acadie-Parc SB	21	30	263	539	20	10	70	33	52	71
Acadie-Parc NB	22	74	267	529	43	43	14	32	48	62
St. Laurent NB	22	81	241	566	0	100	0	45	54	60
St. Urbain SB	20	131	270	676	0	89	11	27	47	56

4.2 Data description

We have access to data about a part of Montréal network. Therefore, regarding the limitation in available data we try to test the usefulness of the proposed method in last chapter for using the streaming data in a part of Montréal network to predict traffic conditions in the same part of Montréal network.
The set of required data needed to test our methodology are: 1) GPS data from probe vehicles, 2) traffic volume data, 3) travel time datasets. In this section, we focus on describing the available datasets.

1. That was defined by the maximum average speed on the link measured by a test car.

4.2.1 GPS dataset

As streaming data we use the data provided by synthetic probe vehicles that are modeled by available real floating cars using SUMO. Prof. Catherine Morency provides us with the GPS data produced by real floating cars (probe vehicles). GPS data is selected for the cars that traveled our study routes during January, February, March, and April 2005. Figure 4.3 shows all the data points that we have access to information using this available GPS dataset. Based on this Figure, we have access to information in all point of the chosen routes for the case study.

The dataset provides information on speed (Appendix B for detailed information on the dataset). The availability of GPS data during different months and on various days are summarized in Figures 4.4, 4.5, 4.6, and 4.7. The dataset is used for developing our models and performing simulations.

Based on Figures 4.4, 4.5, 4.6, and 4.7, we have access to speed data during three time spans: morning, noon, and afternoon. Traveling start times of the floating cars are between 3:30 A.M. and 9:00 A.M. for morning time span, 11:30 A.M. and 1:30 P.M. for noon time span, and 3:30 P.M. and 7:00 P.M. for afternoon time span. The data was collected in Acadie-Parc NB and SB directions during morning and noon only on 2005-01-11 and during morning time on 2005-04-27. In all four directions of the study area, we have access to the speed data on 2005-02-18, 2005-02-23, 2005-02-24, 2005-02-25, 2005-03-07, and 2005-03-11.

The dataset provides information on 2005-02-23, 2005-02-24, and 2005-02-25 in all three time spans (morning, noon, and afternoon). We do not have access to any speed information during morning time span on 2005-02-18, 2005-03-07, and 2005-03-11. Based on Figures 4.4, 4.5, 4.6, and 4.7, morning traffic conditions can be studied on 2005-01-1, 2005-02-21, 2005-02-22, 2005-03-07, 2005-03-11, and 2005-04-27 in Acadie-Parc NB and SB directions, and on 2005-02-23, 2005-02-24, and 2005-02-25 in St. Laurent NB and St. Urbain SB directions. We can use this dataset to model afternoon traffic conditions on 2005-02-23 and 2005-02-24 in Acadie-Parc NB and SB directions, and on 2005-02-18, 2005-02-24, and 2005-02-25 in St. Laurent NB and St. Urbain SB directions. Availability of data during noon time spans enables us to study traffic conditions on 2005-02-23, 2005-02-25, and 2005-03-07 in St. Laurent NB and St. Urbain SB directions, and on 2005-01-11, 2005-02-18, 2005-02-21, 2005-02-22, and 2005-02-24 in Acadie-Parc NB and SB directions. Detail of start times of recording trajectories in the chosen routes is presented in 4 tables in Appendix C.

This dataset provides us with 27 different travel time series for the chosen routes during morning rush hour (07:30 to 08:30). Each car is modeled by 10 probe vehicles in our simulation. Therefore, there are 270 synthetic probe vehicles in our simulation from this source of data that we use as a sample of streaming data to test the suitability of the proposed method.

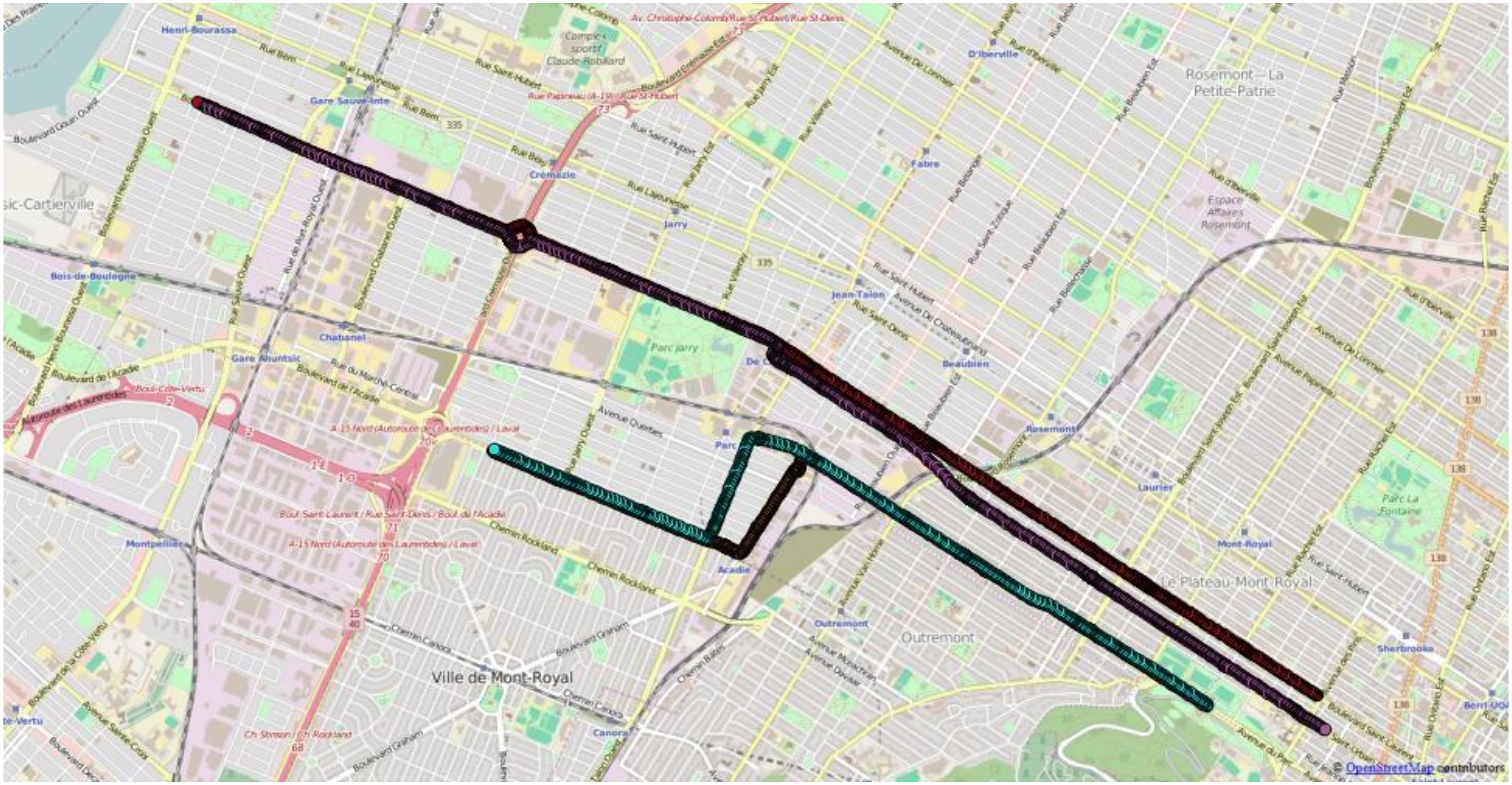


Figure 4.3 Defining the availability of recorded trajectories in the chosen routes using QGIS

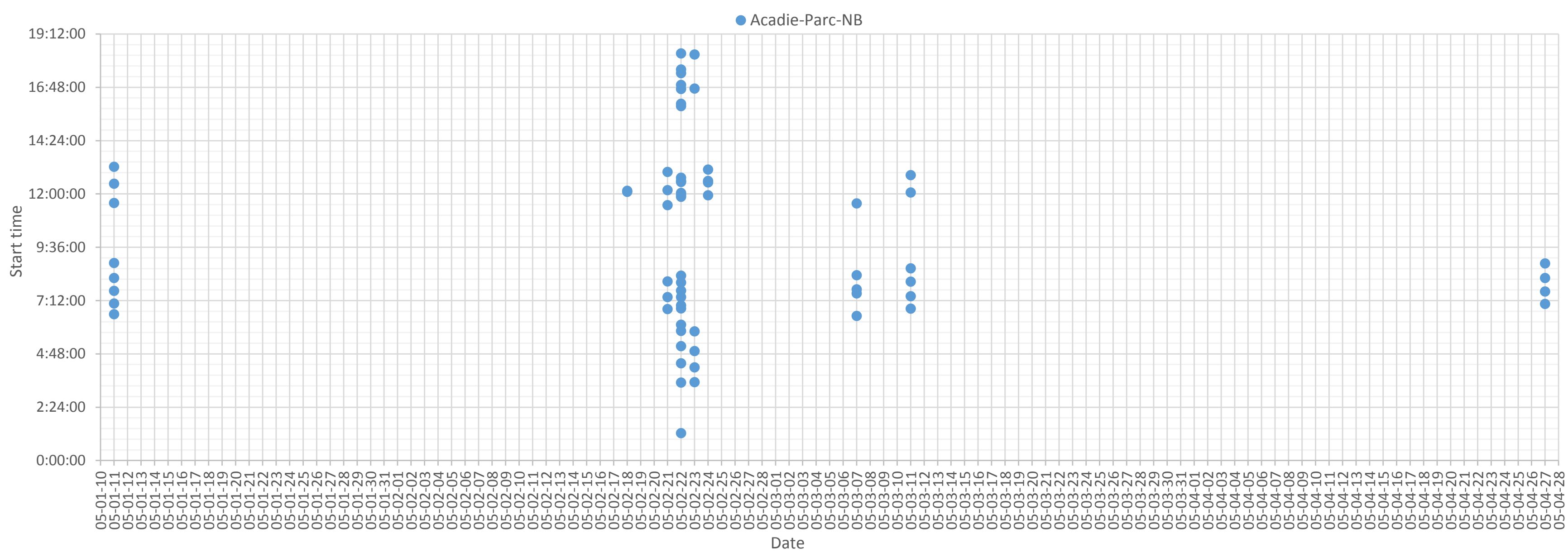


Figure 4.4 Summary of data availability on base of start time of recording trajectories on different months and days in Acadie-Parc NB direction

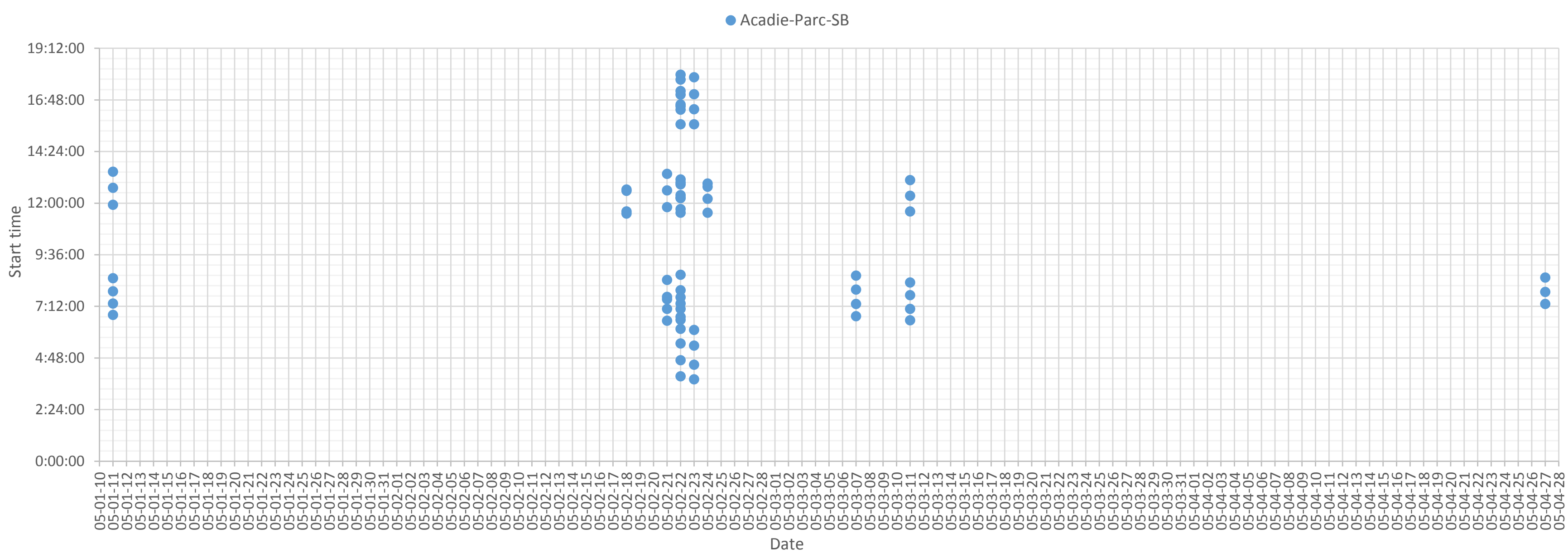


Figure 4.5 Summary of data availability on base of start time of recording trajectories on different months and days in Acadie-Parc SB direction

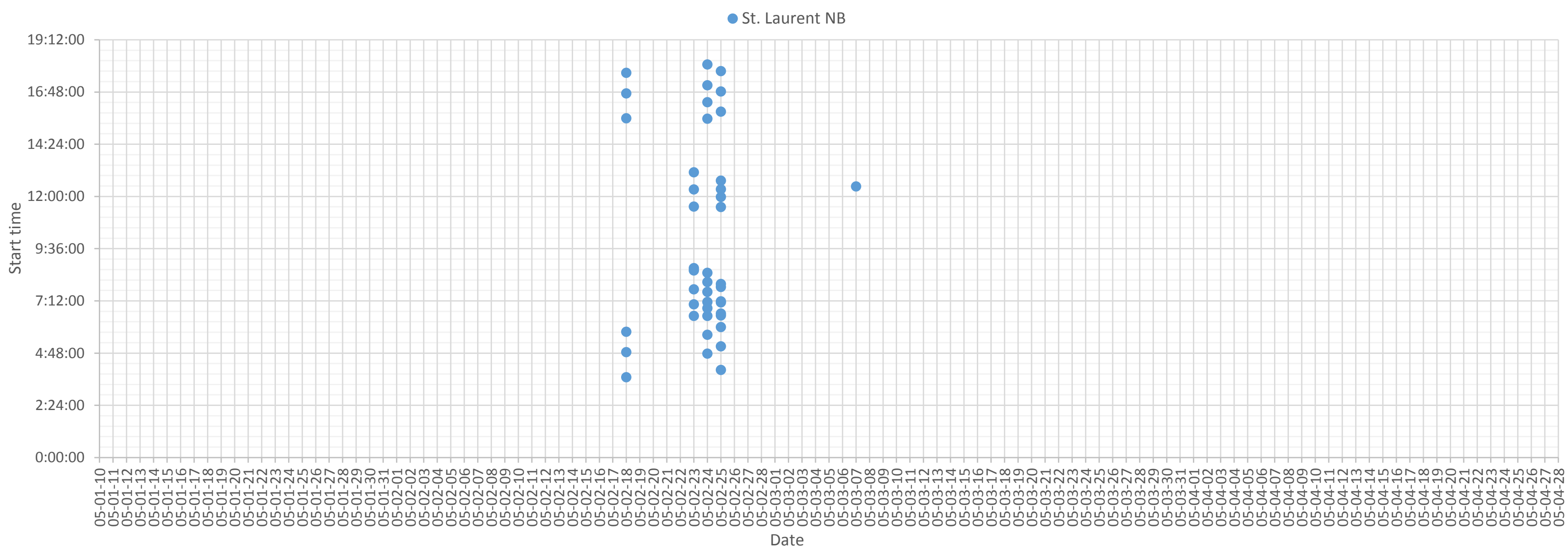


Figure 4.6 Summary of data availability on base of start time of recording trajectories on different months and days in St. Laurent NB direction

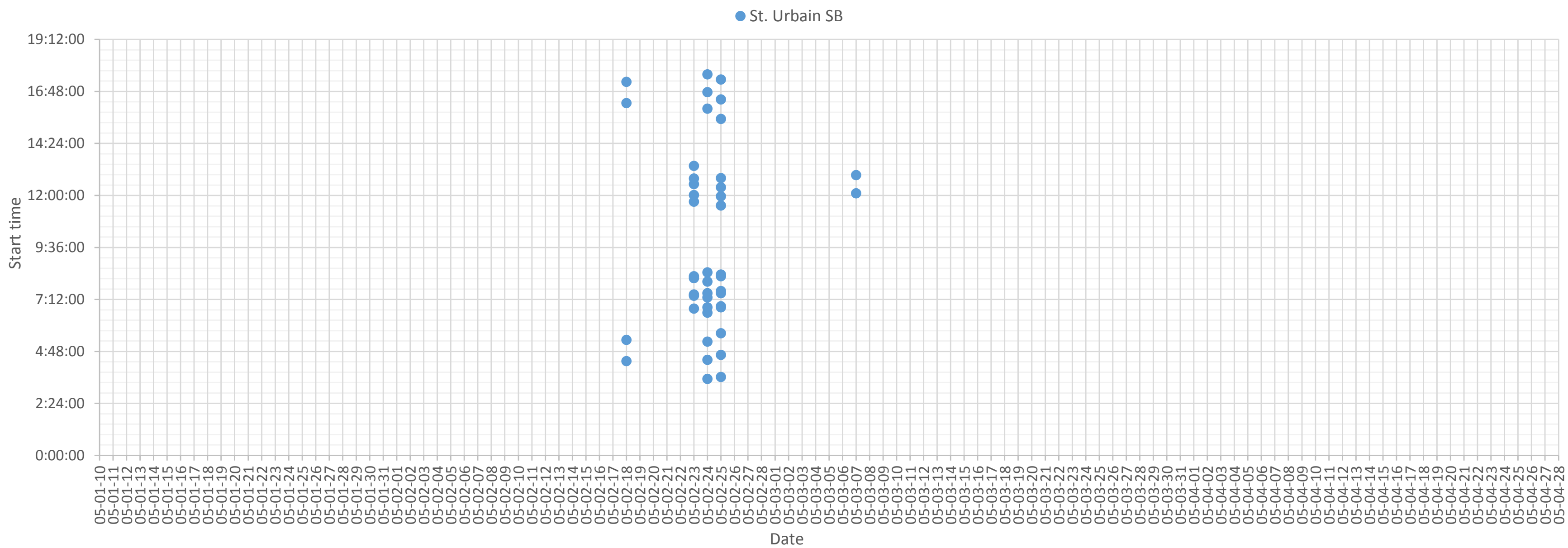


Figure 4.7 Summary of data availability on base of start time of recording trajectories on different months and days in St. Urbain SB direction

4.2.2 Traffic volume dataset

We need to simulate different possible states of traffic conditions using SUMO. To develop simulated states of traffic conditions in the chosen routes, the major and minor sources of traffic volumes in target routes are considered. Therefore, another dataset is acquired from the City of Montréal. The data was collected manually in 15 minutes intervals at all main intersections around Montréal city during September, October, and November 2003. The dataset is freely available at "Comptage des véhicules et des piétons" (City of Montréal, 2016). We use this dataset to consider the effect of traffic volume through the study area (Appendix D shows a sample table of available information by this dataset. Also, it contains a figure for description of the possible turns at each intersection).

Figure 4.8 shows the volume of traffic in morning rush hour (07:30 A.M. to 08:30 A.M) at intersections of the studied routes. Based on this figure, we know that four crowded intersections in Acadie-Parc route are Acadie/Jean Talon, Parc/Van Horn, Acadie/Jarry, and Mont Royal/Parc; in St. Laurent NB direction are Jarry/St. Laurent, Jean Talon/ St. Laurent, Faillon/St. Laurent, and Villeray/ St. Laurent; and in St. Urbain SB direction are Clark/Jean Talon, St. Joseph/ St. Urbain, Mont Royal/ St. Urbain, and St. Urbain/St. Viateur.

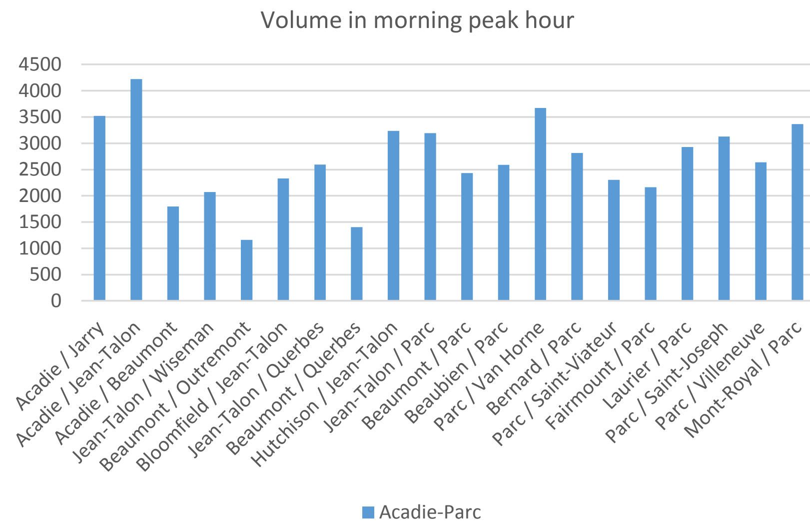
Figure 4.9 shows the volume of main turn [2] at each intersection. This figure shows that the main turn action at most of the intersections in Acadie-Parc route is from north to south. Main turns in St. Urbain SB direction and St. Laurent NB direction are respectively from south to north and from north to south that is expected considering that they are one-way roads.

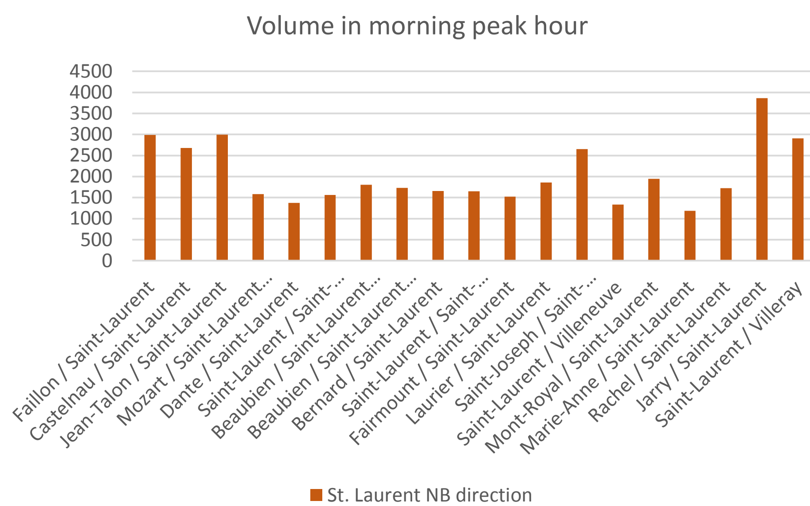
4.2.3 Travel time dataset

Ideal travel time dataset

We need to consider the effect of different patterns of traffic behaviour at the same time as it was mentioned in section 3.1. Different real traffic conditions at morning or evening rush hours are the result of various behaviour by the users in using and different actions by the operators in operating the urban road network. There are remarkable changes in traffic conditions that can be the results of important changes in traffic behaviour. Traffic behaviour by the users change mainly considering the following events and times: the days in the first weeks of schools, snowfall seasons, snowy days, rainy days, other days, and different day of a week. Also, supply networks can be under some constructions at different times.

2. The turn that has the greatest volume between available turns at each intersection.





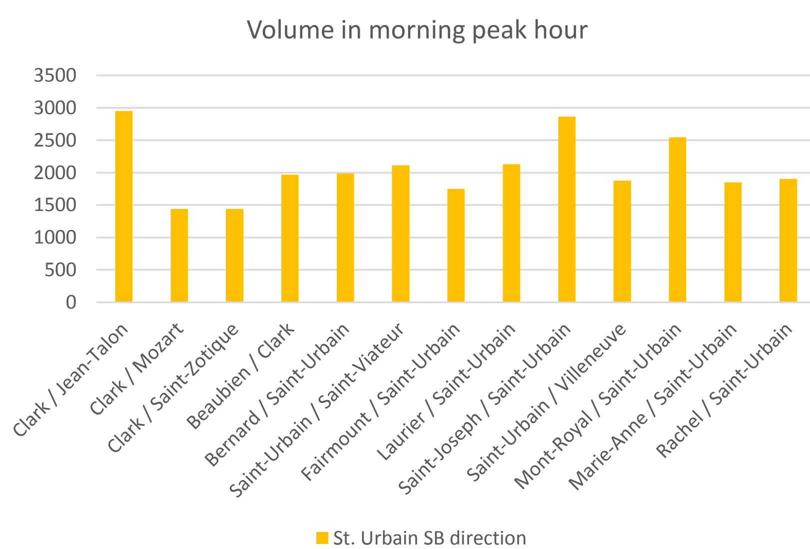


Figure 4.8 Traffic volume at each intersection in morning rush hour

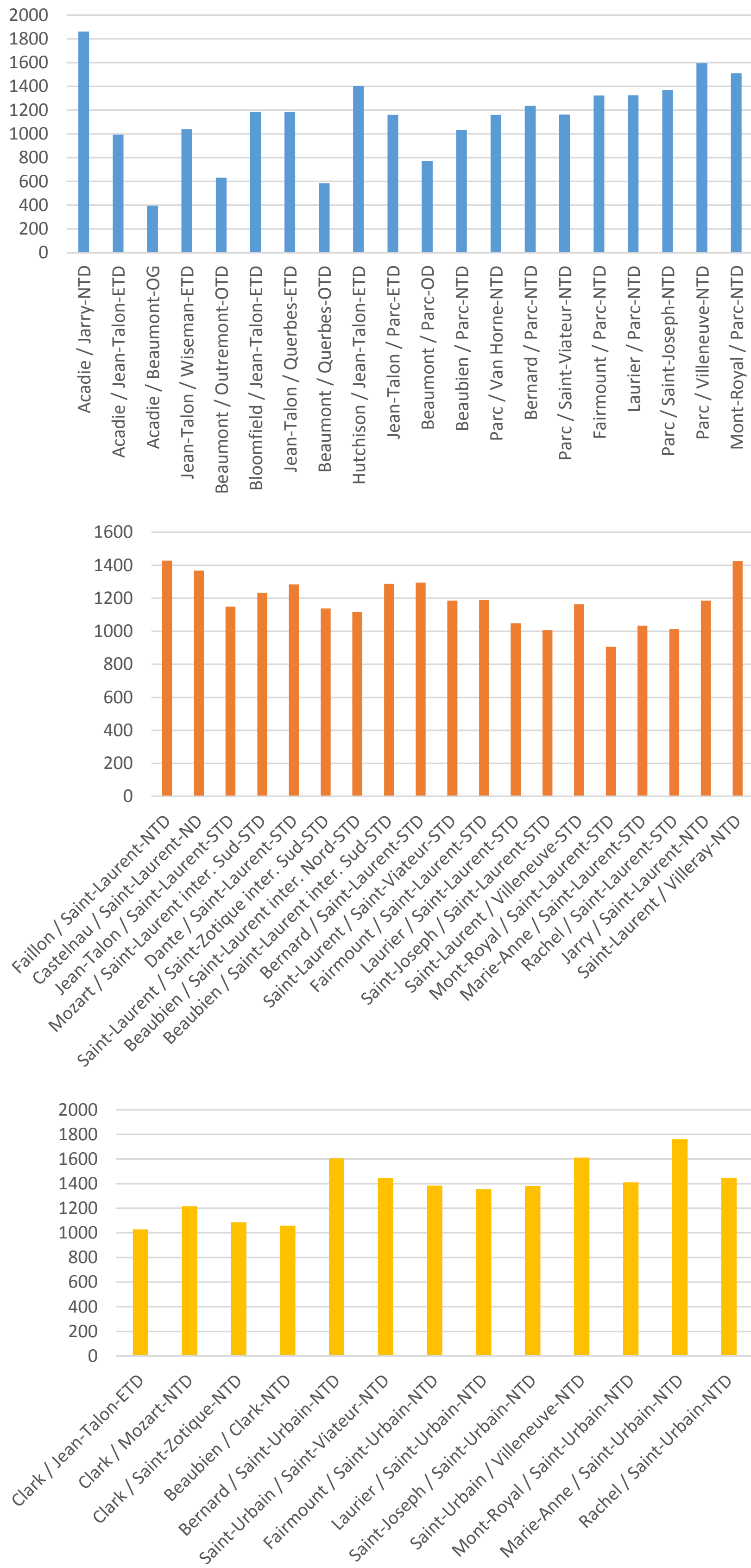


Figure 4.9 Main turn and its volume at each intersection in morning rush hour

This construction can affect traffic conditions. But, for simplification, we only consider the effects of different users' behaviour on traffic conditions of the whole network. We can propose the following categories for the fluctuations in traffic demand and supply that affect traffic conditions in urban road network:

1. First group (Summer - School and University - 23rd of June until 31st of August): 4 different patterns (weekend/non-weekend sunny/rainy days).
2. Second Group (Summer time - University - 1st of May until 22nd of June): 4 different patterns (weekend/non-weekend sunny/rainy days).
3. Third group (First weeks of schools - 1st of September until 21st of September): 6 groups (weekend/non-weekend days of the first week of school, weekend/non-weekend sunny/rainy days of the 2nd or 3rd weeks of school).
4. Fourth group (School time - 22nd of September until 30th of April): 8 groups (weekend/non-weekend sunny/rainy/snowy/after-snow days).

Therefore, ideally we have 22 groups of different days during a year.
We know that traffic counts are collected mainly in the first season of schools (every 5 or 10 years) to help the designers of the urban road networks to plan the networks for the high level of congestion in this duration. With this type of traditional data we need to have access to TT data to calibrate our simulation using SUMO and SA method. Using floating cars we can define TT for part of the network in 22 different groups of days in a year. Therefore, using TT datasets for part of the network we can calibrate 22 different states of traffic conditions in the network.

Available travel time dataset

We have access to travel time data of the studied routes. But the available dataset has not collected as an ideal dataset described in last section. The data has been collected by the same floating cars mentioned in section 4.2.1 by the same source that used for GPS data. The data covers travel times of route segments during morning peak hours for Acadie-Parc-NB direction on 02-21-2005, 02-22-2005, 03-07-2005, and 04-27-2005; for Acadie-Parc-SB direction on 02-22-2005, 03-07-2005, and 04-27-2005; for St. Laurent-NB direction on 02-23-2005, 02-24-2005, and 02-25-2005; and for St. Urbain-SB direction on 02-23-2005, 02-24-2005, and 02-25-2005. Tables 4.2, 4.3, 4.4, and 4.5 provide us with the segments' travel times through 27 travel time series.

Table 4.2 Travel time (sec) in Acadie-Parc -SB direction

Date From/To	02-22 07:37	02-22 07:57	03-07 07:58	04-27 07:52
Beaumont/Beaubien	243	133	131	156
Beaubien/Van-Horn	22	54	19	51
Van-Horn/Bernard	31	44	33	48
Bernard/Bernard-St Vlateur	12	15	23	12
Bernard-St Viateur/St Viateur	34	23	13	17
St Viateur/St Viateur-Fairmont	19	17	16	17
St Viateur-Fairmont/Fairmont	17	21	16	17
Fairmont/Laurier	17	34	34	31
Laurier/St Joseph	61	60	57	11
St Joseph/Villeneuve	38	34	33	27
Villeneuve/Mont-Royal	93	88	23	79

Table 4.3 Travel time (sec) in Acadie-Parc -NB direction

Date From/To	02-21 07:38	02-22 08:00	02-22 08:18	03-07 07:41	03-07 08:20	04-27 07:36	04-27 08:12
Mont-Royal/Villeneuve	68	28	97	76	78	23	110
Villeneuve/St Joseph	56	109	121	56	123	44	60
St Joseph/Laurier	26	26	25	16	19	55	67
Laurier/Fairmont	50	49	15	52	14	20	20
Fairmont/St Viateur	18	17	32	18	18	14	16
St Viateur-Fairmont/St Viateur	35	41	20	41	36	15	27
St Viateur/Bernard-St Viateur	13	17	45	17	13	13	14
Bernard-St Viateur/Bernard	47	46	58	40	24	45	36
Bernard/Van-Horn	55	52	54	52	57	44	49
Van-Horn/Beaubien O	21	23	19	18	75	20	23
Beaubien O/Beaumont	34	73	31	37	34	55	37
Beaumont/Jean-Talon	25	28	43	39	41	34	30
Jean-Talon/Hatchson	55	52	54	52	57	44	49

Possible states of traffic conditions

We are interested in defining the most probable state of the studied network among available simulation results. We need to create possible states of traffic conditions using available data. Based on Tables 4.2, 4.3, 4.4, and 4.5 experienced link travel times are different from a day to day and a month to month during morning rush hours.
In section 4.2.3, the ideal travel time dataset was described. Accessing the ideal dataset needs development of data collection technologies (e.g. DataMobile http://www.datamobileapp.ca/) that was not possible in this thesis. We do not have a complete travel time data for a same day in the chosen routes. We have access to travel time information in one of the chosen routes on a day while we have no information about the other chosen route on the same day. Therefore, we decided to create the possible states of traffic conditions by grouping the

Table 4.4 Travel time (sec) in St. Urbain (SB direction)

Date From/To	02-23 08:10	02-23 08:16	02-24 07:29	02-24 08:01	02-24 08:26	02-25 07:29	02-25 07:35	02-25 08:16	02-25 08:21
Jean-Talon/Mozart O	39	43	43	46	37	42	47	37	45
Mozart O/St Zotiqe	24	60	59	32	58	25	21	62	27
St Zotiqe/Beaubien O	28	40	41	28	39	28	30	41	28
Beaubien O/Bernard	74	93	77	82	76	75	83	86	141
Bernard/St Viateur	29	94	32	91	26	31	25	90	50
St Viateur/Fairmont	107	51	34	43	47	33	32	123	107
Fairmont/Laurier	27	67	22	68	15	15	73	74	80
Laurier/St Joseph	60	28	58	28	57	67	16	24	65
St Joseph/Villeneuve	23	20	23	24	22	17	17	20	23
Villeneuve/Mont-Royal	30	72	35	73	30	21	30	86	33

Table 4.5 Travel time (sec) in St. Laurent (NB direction)

Date From/To	02-23 07:43	02-23 07:47	02-24 07:36	02-24 08:03	02-24 08:29	02-25 07:50	02-25 07:58
Mont-Royal/Villeneuve	48	48	48	48	48	48	48
Villeneuve/St Joseph	78	28	17	19	80	17	16
St Joseph/Laurier	12	18	12	11	19	12	11
Laurier/Fairmont	58	67	13	12	23	59	11
Fairmont/St Viateur	32	41	33	31	40	36	30
St Viateur/Bernard	27	30	25	19	23	23	22
Bernard/Beaubien Est	67	49	47	41	96	48	41
Beaubien Est/Baubien O	45	50	49	6	10	10	6
Baubien O/St Zotiqe	25	31	27	24	28	24	21
St Zotiqe/Dante	13	16	16	18	61	15	15
Dante/Mozart O	14	14.67	14.67	26.33	16.33	26.67	25.33
Mozart O/Mozart Est	14	14.67	14.67	26.33	16.33	26.67	25.33
Mozart Est/Jean-Talon	14	14.67	14.67	26.33	16.33	26.67	25.33

available data and combining them as following. The studied routes are grouped into: 1) Acadie-Parc SB and Acadie-Parc NB directions, and 2) St. Laurent NB and St. Urbain SB directions. Regarding the data, we consider the effect of different behaviour during February, March, and April in the first group and various patterns of traffic flow during 23rd, 24th, and 25th of February in the second group and combine them as shown in Table 4.6.

The difference in experienced travel times suggests the consideration of 9 possible sets of travel time data that we feed them as input to traffic simulation and calibration tool in order to create 9 possible states of traffic conditions based on available data.

These 9 possible states will be used to consider different behaviour of cars on these routes during different days and at different times of our target span of simulation (07:30 A.M. to 08:30 A.M.). With this integration of available travel times data, we are able to produce possible states of the network as the input of our proposed Bayesian approach. Based on

Table 4.6 Nine various inputs (series of travel times) for defining nine different possible states of the network of the case study

State	Acadie-Parc SB	Acadie-Parc NB	St. Laurent NB	St. Urbain SB
1	02-22	02-21+02-22	02-23	02-23
2	02-22	02-21+02-22	02-24	02-24
3	02-22	02-21+02-22	02-25	02-25
4	03-07	03-07	02-23	02-23
5	03-07	03-07	02-24	02-24
6	03-07	03-07	02-25	02-25
7	04-27	04-27	02-23	02-23
8	04-27	04-27	02-24	02-24
9	04-27	04-27	02-25	02-25

the data provided by floating cars it is observed that there is a considerable variation in the reported travel time in each segment during the days and the chosen span of morning peak hours. This means that deciding about the most probable state of network regarding the streaming data that we receive from other sources is important and we need to consider the different patterns of behaviour based on various recorded series of travel times.

By existing limitation in available data, there is no more information about the probability of measured travel time in the segments of the chosen routs of this study. Therefore, we consider the same probability for all 9 possible states that are defined as input of our proposed Bayesian approach and are simulated by inputting segments travel times to the simulation and calibration tool.

Nine possible states of traffic conditions in our case study are simulated and calibrated separately. The simulation outcomes provide details on traffic conditions. Therefore, we can define the probability of measurements and also conditional model using the data provided by 270 synthetic probe cars. The most probable state of system among the 9 possible states will be the state with the highest probability using the proposed Bayesian approach that will be recommended as the prediction of the network state.

4.3 Data analysis

4.3.1 Analysis of available GPS data

Available recorded GPS data can be divided into 3 different groups based on the time of recording data during each day. As shown in Figures 4.4, 4.5, 4.6, and 4.7, it is clear that the first group relates to the morning time before 9 A.M., the time span of the second group is around noon, and the third group is dedicated to the evening. The data provides us with the ability to develop models for each of the different time spans of the day to study traffic

behaviour.
We are interested to test the ability of the proposed Bayesian approach, and we focus on peak hours which are important for both users and operators of urban network. Therefore, we concentrate on specifically one hour during the morning peak hours (07:30 A.M. to 08:30 A.M.) to test the proposed method of this study.
Figure 4.10 shows the availability of start times of trajectories recorded between 07:30 A.M. and 08:30 A.M. for: a) Acadie-Parc NB direction, b) Acadie-Parc SB direction, c) St. Laurent NB direction, and d) St. Urbain SB direction. There are 2, 1, 2, 3, and 1 probe vehicles in Acadie-Parc NB direction, and 1, 2, 1, 2, and 3 probe vehicles in Acadie-Parc SB direction, respectively on the 27th of April, 11th of March, 7th of March, 22nd of February, and 21st of February. Also, there are 2, 2, and 1 probe vehicles in St. Laurent NB direction, and 3, 2, and 2 probe vehicles in St. Urbain SB direction respectively on 25th, 24th, and 23rd of February.
Using the latitude and longitude of the GPS data, we can draw Figure 4.11, and see which one relates to which route regarding the difference between their patterns. This is more clear by inputing the GPS data to QGIS software as it has been provided by Figure 4.3.
Considering the average speed experienced during morning time (See Appendix E), it is clear that at proximity of the most important arterials that cross the two studied routes there are considerable decreases in speed.
If we plot average speed in all the 4 paths (See Appendix F), it can be concluded that although speed patterns are different during the 3 time spans (morning, noon, evening) of the days, the average speed fluctuates approximately in the same region.

4.3.2 Analysis of available counted data

Using the available traffic volume data at intersections, traffic conditions are simulated. Appendix G shows results of defining turning ratios at all intersections. Summary of probabilities of possible turns at each intersection, which we use in simulation, has been presented in Appendix G.

4.3.3 Analysis of travel time data

Analyzing the fluctuations of experienced travel times, it is observed that at noon time span, travel times have less fluctuations (see Appendix F). Also, it is clear that travel time increases from the early morning and it is approximately in the same region of the end of morning time span at noon time span. But it increases during evening time span.

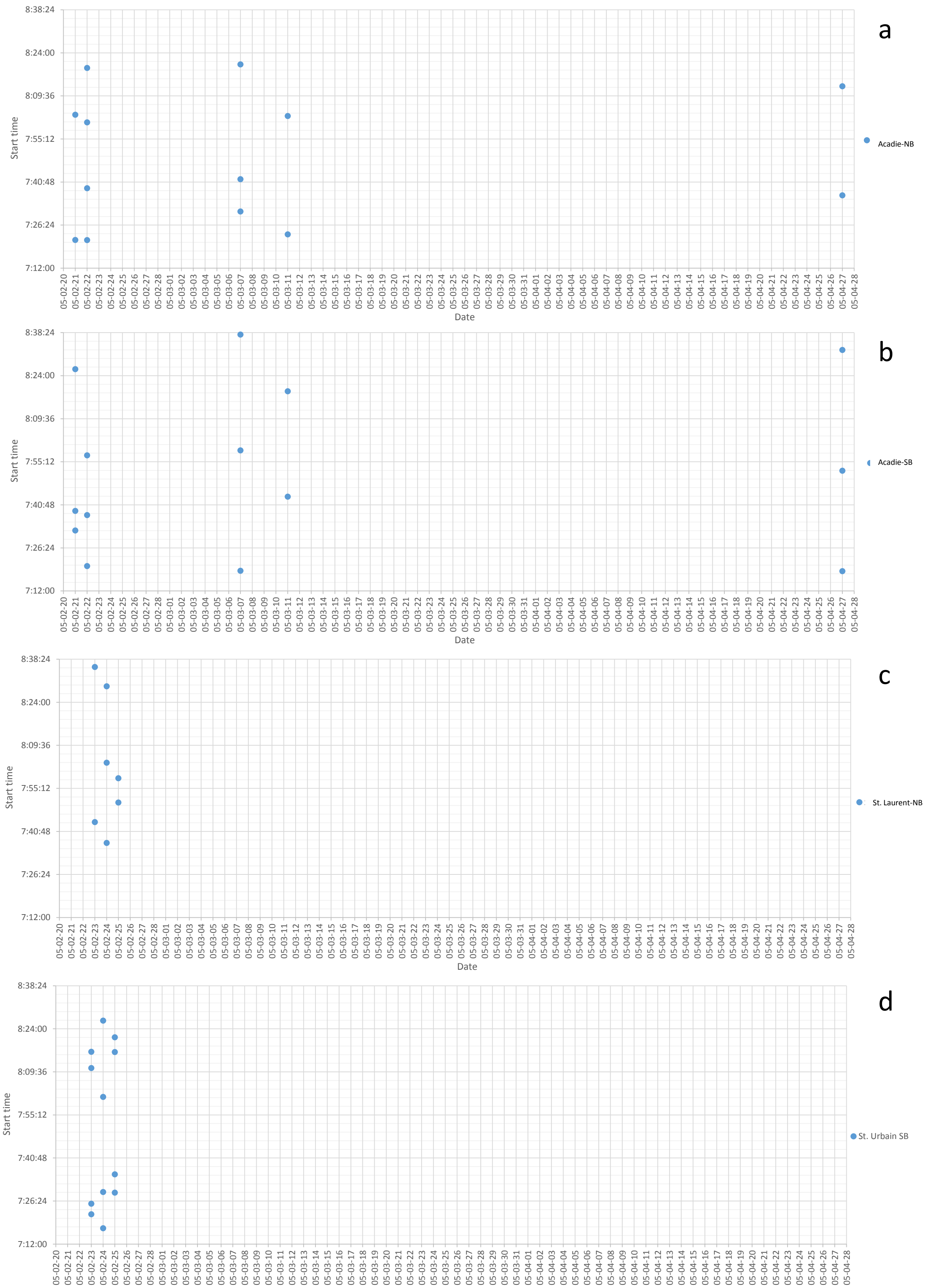


Figure 4.10 Available GPS recorded trajectories during morning peak hour (07:30 A.M. to 08:30 A.M.)

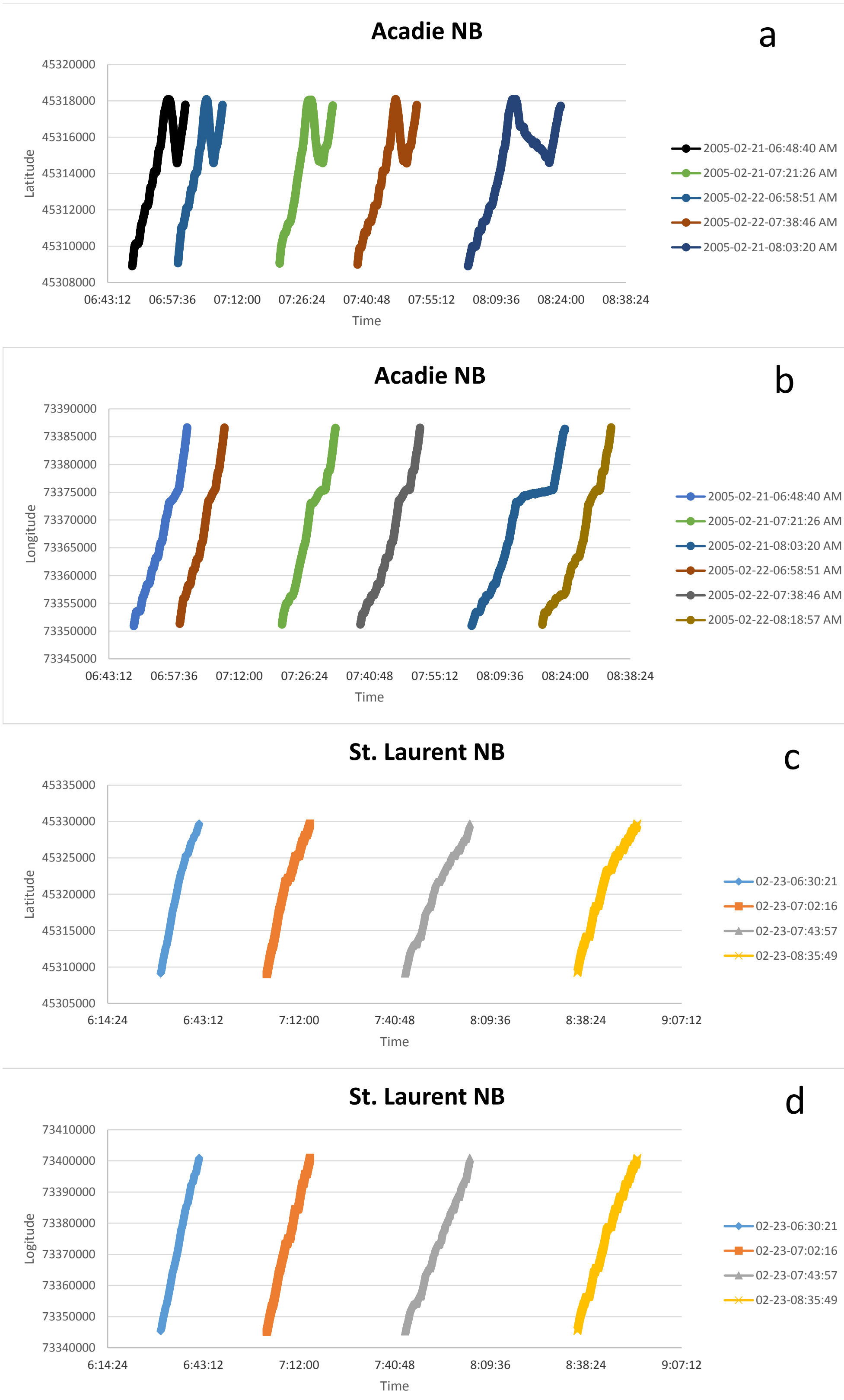


Figure 4.11 Various patterns of the routes as it is clear from latitude (a and c) and longitude (b and d) of GPS traces

4.4 Simulating the possible states of traffic conditions

To feed the available data to simulation and calibration tool we use all the 27 series of TT data regarding the limitation in data source. We know that in proper way, we should simulate and calibrate possible states of traffic conditions by the bigger part of data and consider the rest (smaller part) of data for validating the model results. In our case study, regarding the mentioned limitation, we validate the result of the model by speed data of the same floating cars that is highly correlated to travel time indicator.

To perform the simulations, we must create 22 groups of different input files for all 22 ideal possible states using ideal travel time dataset based on section 4.2.3. But by existing limitation in available data based on section 4.2.3, we have 9 groups of different input files and possible states. All states are simulated using the same network and mainly the same demand conditions. But one of the input demand files (internal demand and dynamic and static external demand) that defines the dynamic part of demand is different.

To define network input files, we create node characteristics, edges characteristics, connection between edges, and traffic light definitions for signalized intersections.

To define demand conditions, five input files are created. The number of nodes and edges used in defining input files are presented in Appendix H. One of the five input files that we need to define the demand conditions is created to define the time and the path of assigning probe vehicles (See Appendix I as an example). While there are 9 possible states, we need to define 9 different files to consider different probe vehicles in each state.

Also, the other file is created to define the target travel times of different links in the study area that were defined using the data provided by floating cars (See Appendix I as an example). There are nine possible states; therefore, we need to define nine various files to consider the different experienced travel times of the links.

Figure 4.12 is the case study network created for performing SUMO simulations using SUMO Graphical User Interface. Each state has been calibrated using the code developed for applying SA method and the outputs of each state have been used to apply the proposed methodology.

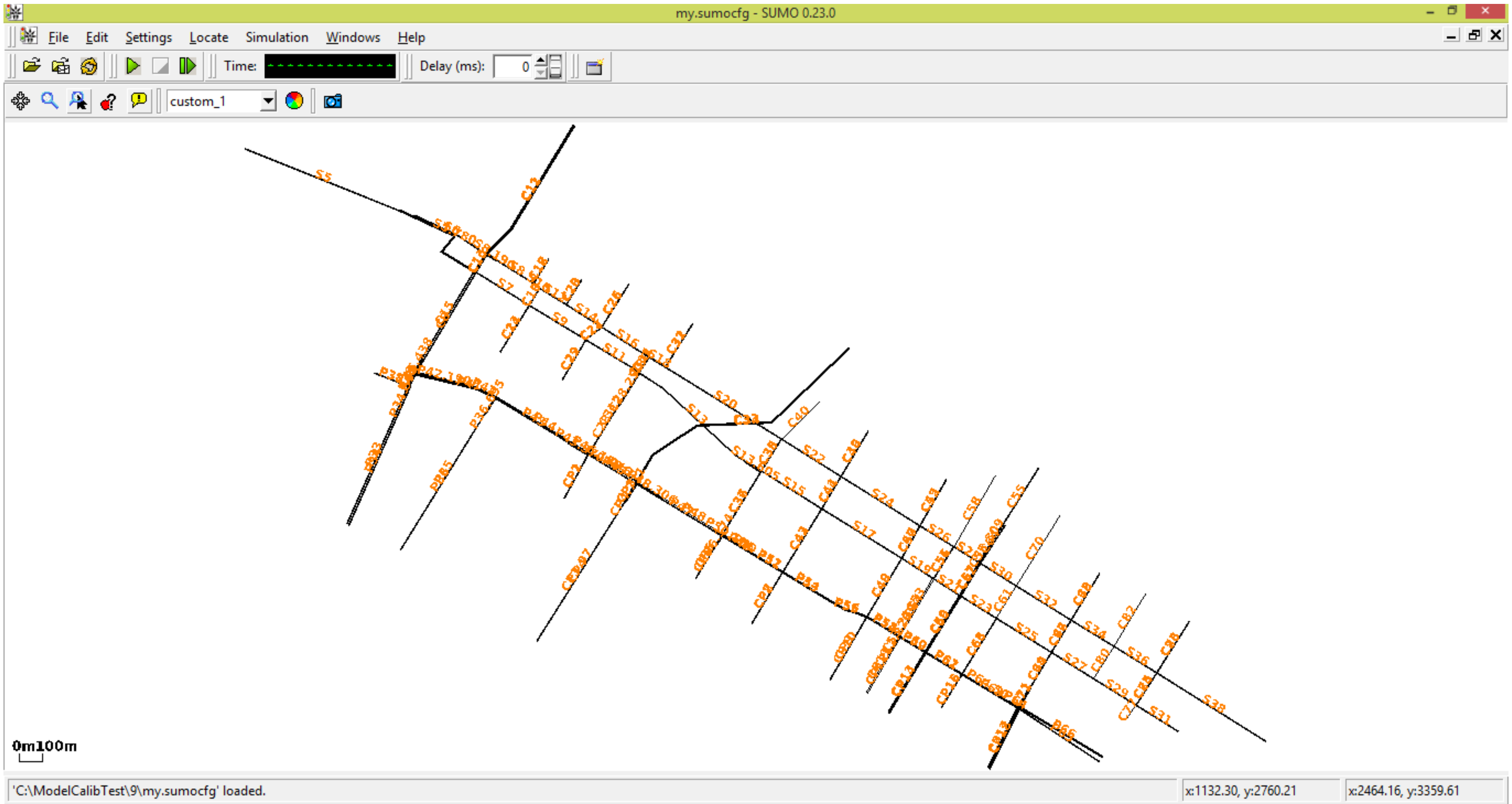


Figure 4.12 Modeled network using SUMO (Output by SUMO GUI)

CHAPTER 5 IMPLEMENTING THE PROPOSED METHOD, RESULTS AND DISCUSSION

5.1 Implementing the proposed methodology

The probability of each state at each time-step of morning peak hour is calculated. Using the output file of each simulated state that contains information about position and speed of each simulated probe vehicle, the average and standard deviation of experienced speed at each time-step for reference time, Δt (10 seconds before each time-step), for each probe vehicle are calculated. Then the amount of probability for the related speed for each probe vehicle at each time-step is defined by applying equation 3.6.

The value of probability of state of each link at each time-step when the related speed is zero, is then defined using the outputs of simulation and by applying equation 3.8.

To define probability of state of each link when the reported speed is zero, we need to define the volume of vehicles located after the probe vehicle. We define a unique code for all reported trajectories using the related link, vehicle name and the related time-step. We define the rows of information that are related to the specific combination of the links and probe vehicles that have speed of zero. Then, we can define the time-step (and the related information) that probe vehicles with reported speed of zero have entered into the related links. We combine the produced table in last two steps and produce a new table. We sort the new table by link name, probe vehicle name, and then by time-step. We have defined the input file we need to produce the volume of vehicles that enter into each link during the time of starting and ending zero speed and the time that the probe vehicle has entered into the link. We use the characteristics of intersection for serving each connected link to calculate the density of each link at each time-step.

In order to define the volume of entered vehicles into each link, performing SUMO is time-consuming and must be managed regarding the limitation of SUMO in producing output files and the limitation of Excel that we use for processing the output files produced by SUMO. The number of arrived and entered vehicles into each link during the related time spans are defined using SUMO. Then, we can define the probability of calculated density when the reported speed of the probe vehicle is zero. Until now, we have defined the probability of measurements. To define probabilities related to conditional model, we need to produce the other outputs files from the simulation.

We define conditional model in the status that the reported speed is non-zero using the average speeds of the links that is produced as the output of simulation and available synthetic streaming data by applying equation 3.10. In the status that the reported speed is zero, we use link densities and the volume entering into each link (during the time of starting and ending zero speed and the time that the probe vehicle has entered into the link) as the outputs of simulation using equation 3.11.

The state probability for each link at each time-step is then summarized. We define and use three different dependency matrices (Degree, Direction, and Effect) which we need to consider the dependency between links. We assumed that if two links are connected indirectly with a link and two links, their degrees of dependency are respectively 0.5 and 0.25. We define the depended state probabilities of each link. Then, we determine the final calculated value of the related probabilities.

Using the proposed Bayesian framework, we can define different probabilities for 9 possible states of traffic conditions as the output of different simulation results in the context of available streaming data using 9 probabilities of the prior model (that in our case are the same), probability of measurements (available streaming data), and 9 conditional model at each time-step. The most probable state of the studied network introduces as the prediction of the proposed method at each time-step.

5.2 General results and discussion

Figure 5.1 shows the probability of each state at each time-step from the beginning of the period of study (07:30 A.M.) to the end (08:30 A.M.). Most of the calculated probabilities are less than 0.3. The most probable state at each time-step has been defined and the output of the Bayesian model as the most probable state (MPS) is shown in Figure 5.2. Based on this output, the proposed Bayesian model can define the most probable state of urban road network among the 9 considered possible states. State 1 is the most probable state. Results show that the model is able to consider the effect of applying the most recent streaming data to improve the predicted state of a modeled urban road network. Also, the model provides enough feedback from the recent experienced state of network.

We are using the data provided by synthetic probe vehicles; therefore, it is clear that in some time-steps there is no possibility for some states. Because there is no input to the proposed Bayesian model from synthetic vehicles at these time-steps. State 1, state 4, and state 7 have no possibility in time-steps at the beginning of the studied period.

Moreover, results in Figure 5.1 demonstrate that state 1 and state 2 have higher possibility compared to other possible states. State 3 has a possibility of less than 0.1 at most of the time-steps. Also, the other states (states 4, 5, 6, 7, 8, and 9) have approximately the same pattern in calculated probability and are less probable than states 1 and 2.

Figure 5.2 explains that state 2 is the most probable state from 07:30 A.M. to 07:45 A.M. and 08:00 A.M. to 08:15 A.M. Also, state 1 is the most probable state between 07:45 A.M. and 08:00 A.M., and 08:15 A.M. and 08:30 A.M. State 3 is the less probable state. States 4, 5, and 6 can be categorized in other group. They are more probable than state 3 but less probable than states 7, 8, and 9. This leads us to conclude that stochastic nature of traffic conditions prediction can be captured using the proposed method while in nearly 80 % of time-steps, states 1 and 2 that are approximately 20 % of possible states are more probable in comparison to the rest (80 %) of possible states.

State 1 has been simulated using the travel time data on 2005-02-22 for Acadie-Parc SB, on 2005-02-21 and 2005-02-22 for Acadie-Parc NB, on 2005-02-23 for St. Laurent NB, and on 2005-02-23 for St. Urbain SB. State 2 is the same as state 1 in using the travel time data for the first route (Acadie-Parc SB and Acadie-Parc NB). The difference between state 2 and state 1 is in applying the travel time data for the combination of St. Laurent NB and St. Urbain SB (second route). We used the data that is available on 2005-02-24 in state 2 for the second route.

The difference between state 3 and states 1 and 2 is in applying different travel time data for the second route. We used the data that is available on 2005-02-25 in state 3 for the second route. Therefore, it can be mentioned that the data provided on 2005-02-21 and 2005-02-22 for the first route in combination by the available data for the second route on 2005-02-23 and 2005-02-24 are able (and on 2005-02-25 are not able) to produce the simulation results that are more compatible with the available probe data at most of the seconds from 07:30 A.M. to 08:30 A.M.

This finding leads us to conclude that the the simulated state using the travel time information on same day or on the days by same pattern of traffic behaviour is more able to realistically model traffic conditions using streaming data.

Also, it is worth mentioning that combining the travel time data from different days of different months (2005-03-07, 2005-04-27, 2005-02-23, 2005-02-24, and 2005-02-25) as used in defining states 4, 5, 6, 7, 8 and 9 is less able to produce the simulation results that are nearer to the data provided by probe data.

5.3 Results for sample link

We can draw the fluctuations of speed and density based on the most probable state at each time-step. The results of applying the proposed methodology in order to define the most probable state of urban road network has been shown, as instance, for link P45 in Figure 5.3 and Figure 5.4. These figures show the predicted (blue line) speed and density at each

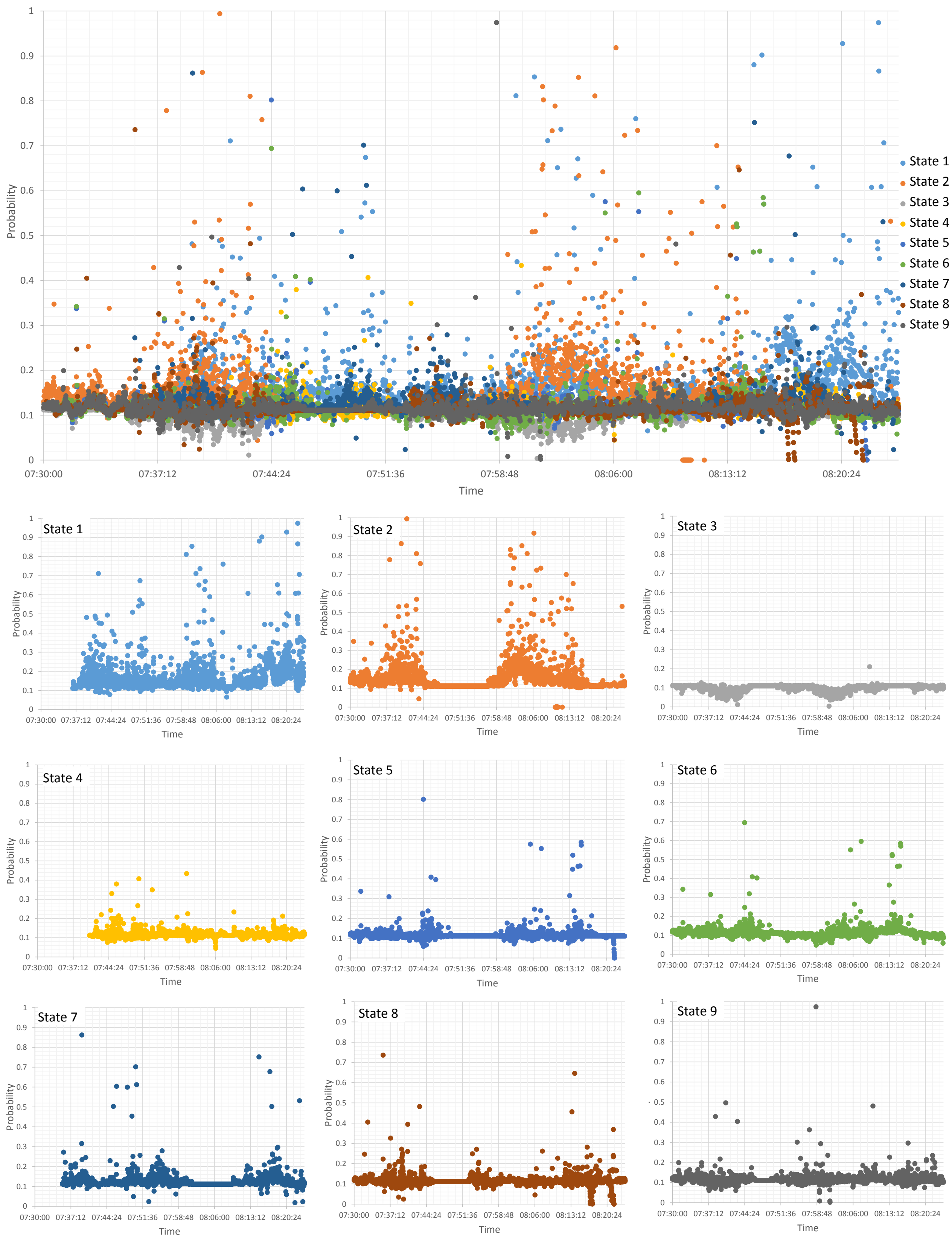


Figure 5.1 Distribution of calculated probability of each available state

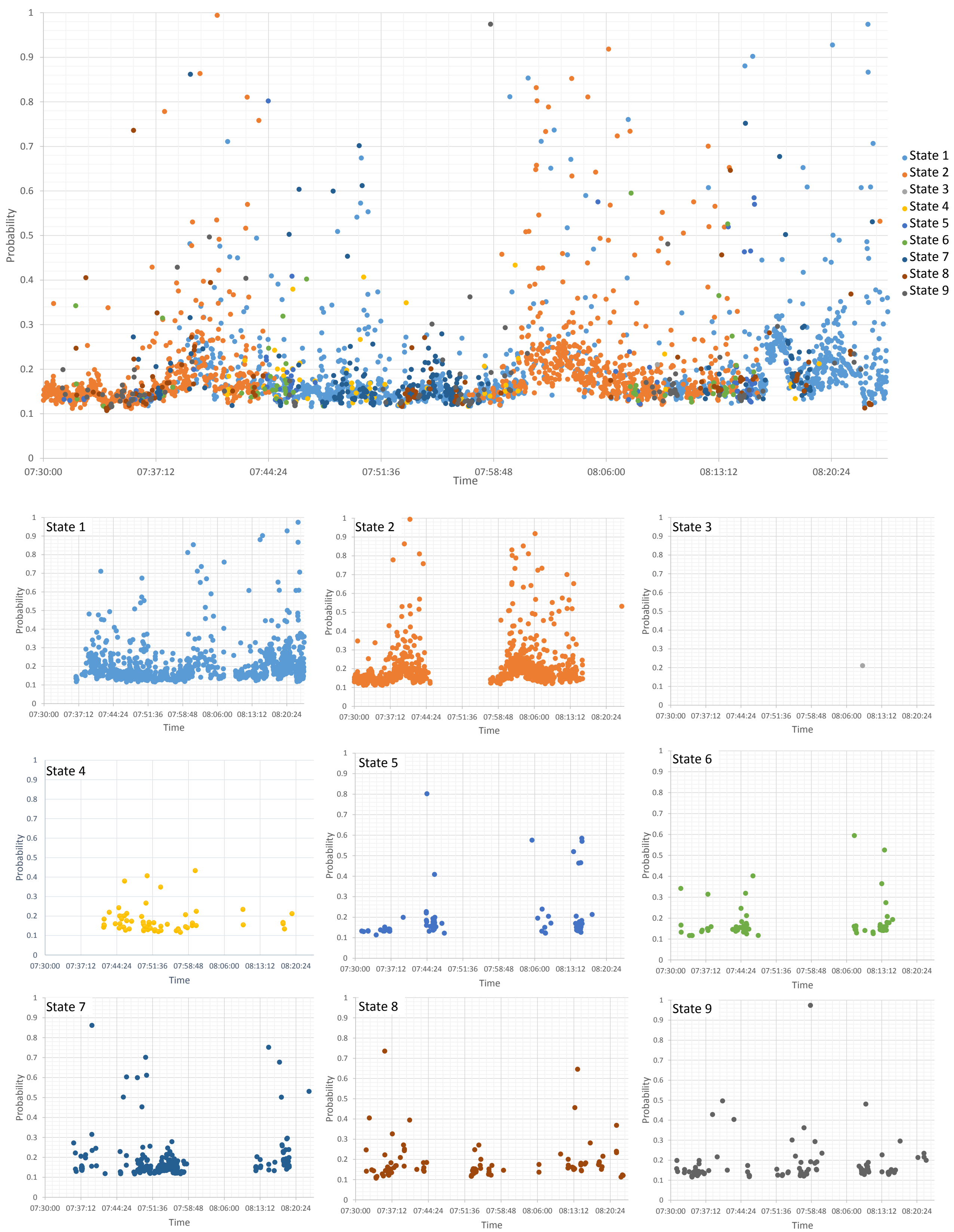


Figure 5.2 The most probable state at each time

time-step that can be compared by minimum (gray line) and maximum (red line) possible values at each time-step.

Comparing the output of the proposed Bayesian approach with the probable maximum and minimum speeds and densities (available from different possible states using simulation) shows that deciding about the most probable state of system can be complicated. Although, we tried to define the most probable state of urban road network in the case study using synthetic streaming data, it is mentionable that the wide span of possible speeds and densities guides us to use more data in predicting the state of urban road network more precisely. For instance, at time-step 2420, the difference between maximum and minimum possible densities of link P45 (0 and 15) shows us that deciding about the most probable state of urban road network needs to be studied using more data. Therefore, we can conclude that predicting traffic conditions is complicated problem that by using sufficient data we are able to decide about the most probable state of network more precisely.

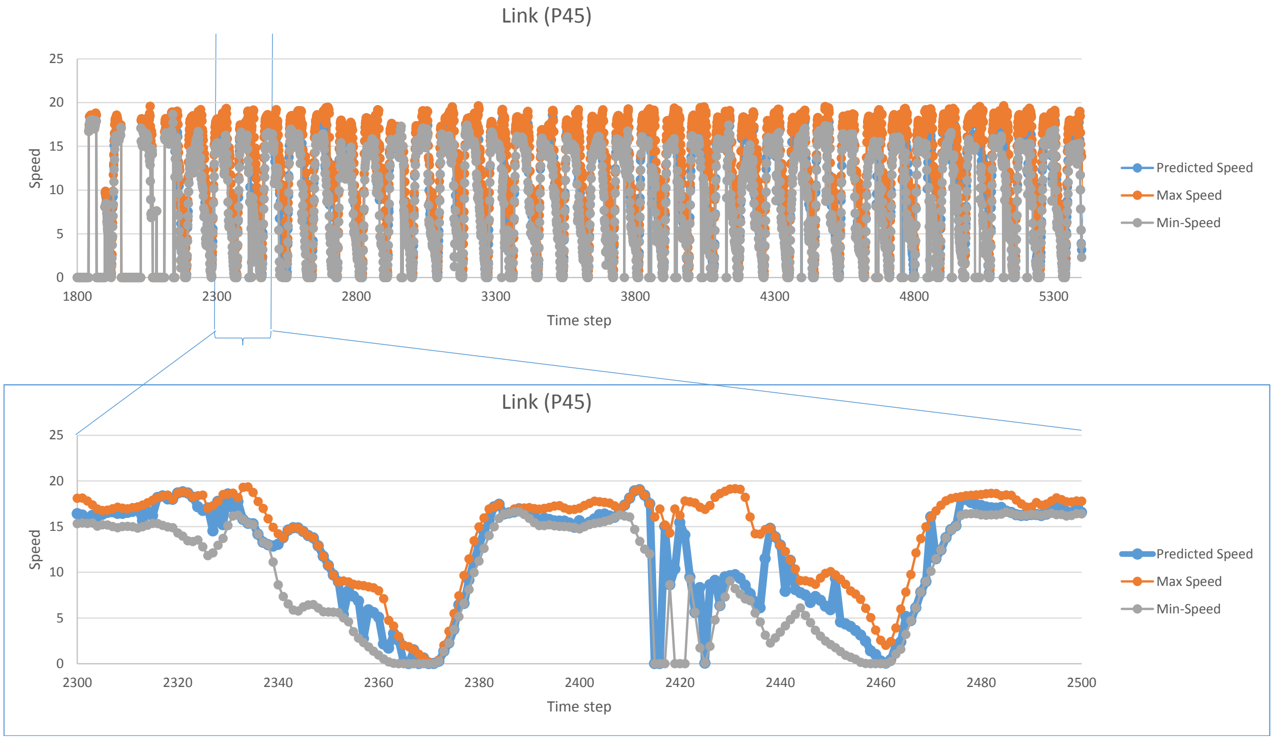


Figure 5.3 Predicted speed in link P45

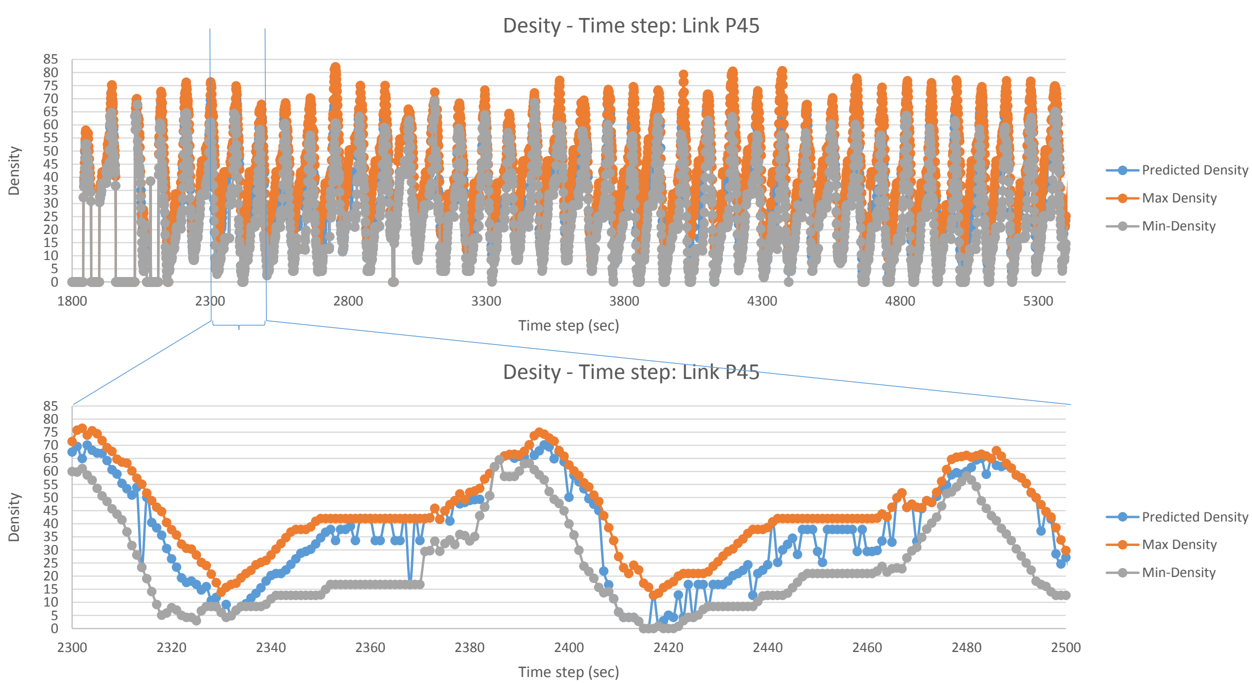


Figure 5.4 Predicted density in link P45.

CHAPTER 6 SENSITIVITY ANALYSIS

As mentioned in the previous chapter, we need to be cautious in deciding about the most probable state of the urban road network. Therefore, we need to analyze sensitivity of the proposed Bayesian approach for predicting the most probable state of the urban road network. If we analyze the output of the proposed method based on the possible states at each time-step, we can conclude that in 98.64 % of the time-steps there is a recommended state as the most probable state of the urban road network of our case study. While only in 1.36 % of the time-steps, the proposed approach results in two possible states, and there is no time-step with the possibility of more than two probable states as the recommended state of urban road network based on the applied case study. It is noticeable that when we have no streaming data to feed the proposed Bayesian approach we cannot decide among the possible states. This point underlines the sensitivity of the proposed approach in providing a unique state as the most probable state of urban road network.
The applied method based on the implemented case study proposes state 1, state 2, state 7, state 9, state 6, state 8, state 5, and state 4 respectively in 44.3 %, 34.1 %, 7.4 %, 3.5 %, 3.3 %, 3.1 %, 2.4 %, and 2 % of time-steps. This means that in nearly 80 % of the time-steps, the proposed approach guides us to two specific states (state 1 and state 2) as the most probable state.

6.1 Sensitivity of the results to sampling rate

We can test sensitivity of the proposed method to different sampling rates. We test the reaction of the proposed method by using more or less streaming data and feeding it to the proposed Bayesian approach.
We applied 100 % of the available streaming data in the main part. Here we are interested in the sensitivity to the network coverage and amount of available data. Using two different strategies of sampling rate, we implemented 50 % and 75 % of the available streaming data in the proposed method. Figure 6.1 shows the result of using 50 % of the available data. Also, Figure 6.2 represents the result of implementing 75 % of the data. We can compare the results of these two types of sampling rate to the results of the main model with the entire data. It is clear that by increasing the volume of input data when we use more data (Figures 5.2 and 6.2), we can achieve a model that recommends the state of urban road network with more probable states at different time-steps. Using 50 % data (Figure 6.1) results in only a state at most of the time-steps while we have 2 possible states at most of the time-steps

using 75 % of the data and the entire data.
By using more data, we can detect the probabilistic feature of the problem of traffic conditions prediction in an urban road network. Different values of travel time experienced by real floating cars highlight the point that the pattern of traffic state and drivers' behaviour cannot be described by a deterministic state. We need to consider different patterns and behaviour to detect the probabilistic feature of the problem of traffic conditions prediction in the urban road network. Therefore, the main conclusion of sensitivity analysis is that by increasing the network coverage and volume of steaming data we are able to better stochastically predict traffic conditions.

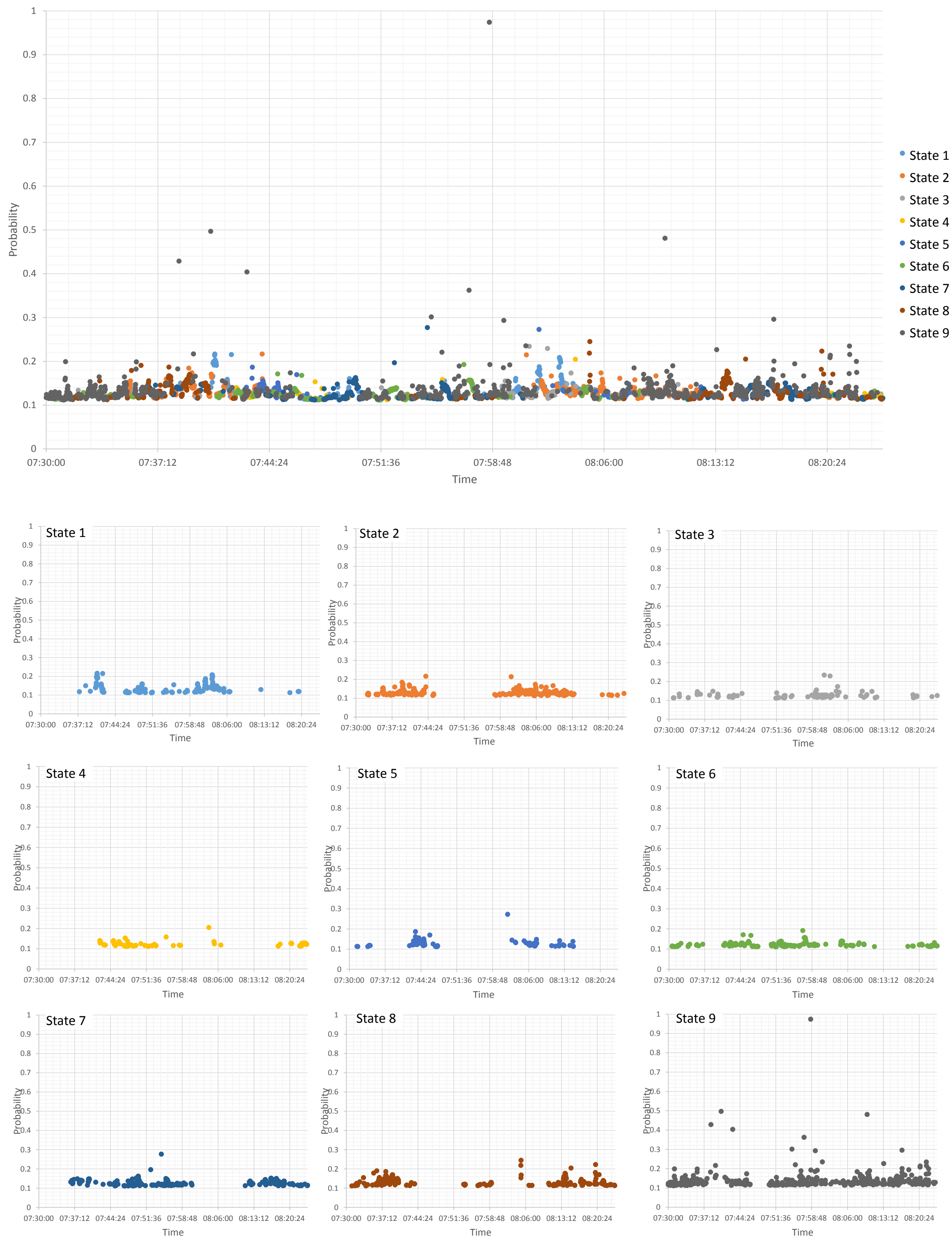


Figure 6.1 The most probable state at each time using 50 % of available streaming data

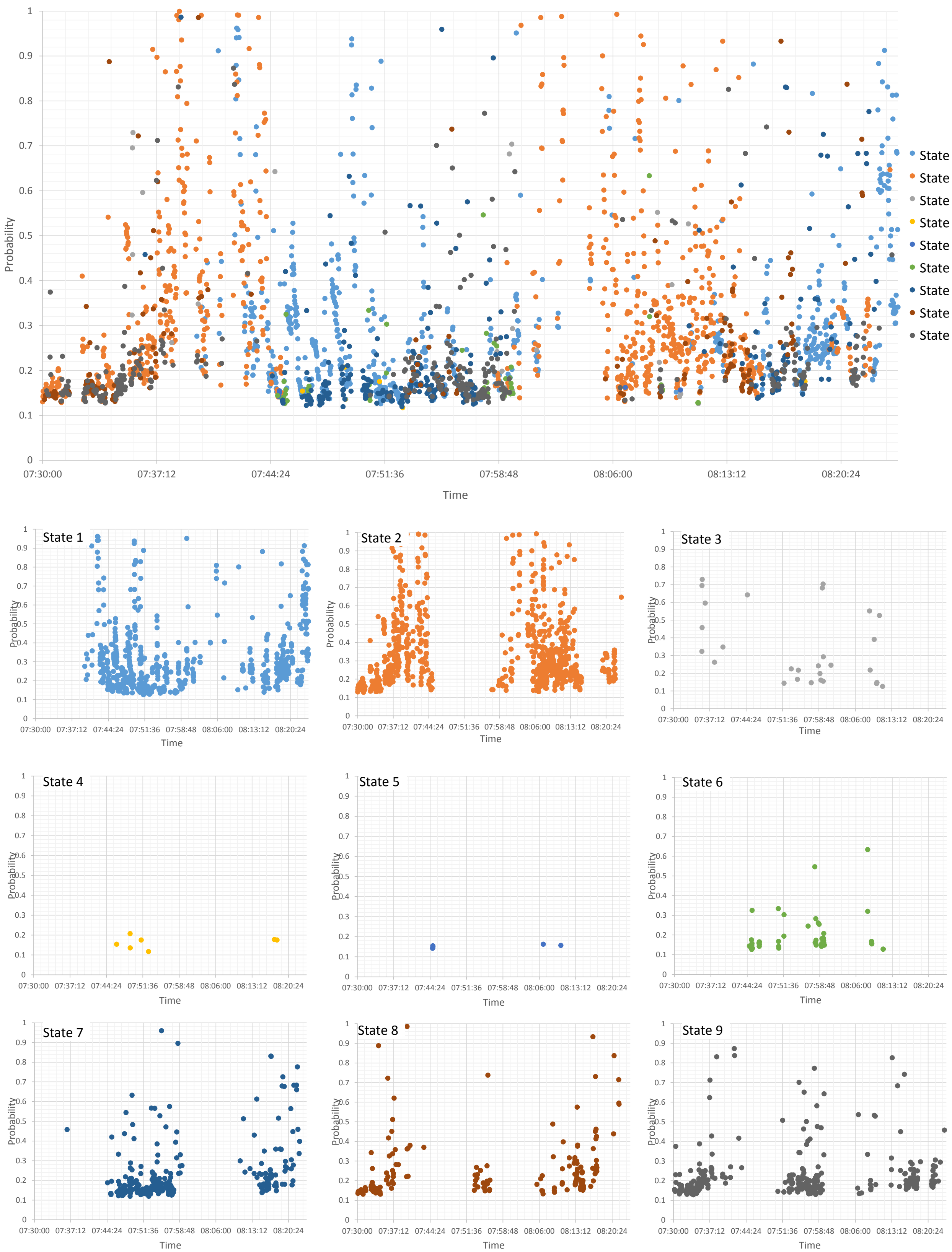


Figure 6.2 The most probable state at each time using 75 % of available streaming data

CHAPTER 7 VALIDATION

We divide the available data into two parts. As it was mentioned in section 4.4, in order to create a model with the data, we use the smaller part of the data in validating the results of the model that is created by the main (bigger) part of the data. In this thesis, regarding the limitation in data availability, we used entire travel time data in creating the model. To validate the result of the model created by the proposed method, we use the available data for the other indicator (speed) that is highly correlative with the travel time indicator that we used in creating the main model.

We can use the data from available 27 real probe vehicles and compare the output of the proposed method in recommending the most probable speeds to the real experienced speeds that are available by real vehicles in the links. We know that comparing the speed of a car, by GPS, to the average speed on a link that is the result of a model cannot help us to accurately validate the result of the proposed method of this study. But regarding the limitation in available data, we use this part of data for validation.

Available GPS data for real probe vehicles define the position of the real vehicles based on their latitude and longitude. We need to define the latitude and longitude of the links. After defining the latitude and longitude of the links, we identify where each real probe vehicle has been located at each time-step (See Appendices J and K).

Figures 7.1, 7.2, and 7.3 show the result of validating the output of the proposed method by comparing to available GPS reports for Acadie-Parc SB direction. These results show that the span of speed change predicted by the proposed Bayesian method is the same as what was reported by GPS devices.

We have access the GPS reports of 4 real probe vehicles for St. Urbain SB direction. The

Table 7.1 Result of validation

Link	Error in predicting speed (%)
P45	36
P47	42
P49	20
P51	26
P53	111
P55	260
P57	170
P59	47
P61	21

validation results have been summarized in Table 7.1. The data confirm the results of the proposed method by considerable error. The reason of this error is that the number of available vehicles that reported their status in this direction is low. Also, the distribution of the real probe vehicles is important in assessing the validity of the proposed method. In our case study, regarding the small dataset that is available for validating the results of the proposed model, the volume and distribution of real probe vehicles are not sufficient and suitable.

This validation method expectingly produces weak result, because due to the limitation in available data we were forced to integrate the travel time data for a day with the data for another day. Also, we recommend the most probable state of network in our case study among 9 possible states at each time-step. This leads the results in the chosen arterial for validation by considerable error, because we are comparing the real speed data for 4 probe vehicles that traveled Acadie-Parc SB direction on 2005-02-22 by the result of our model as the most probable state among 9 possible states that are combination of different data from different days and different months. We mentioned that regarding the limitation in available data we decided to create 9 possible states of this case study. This type of defining possible states leads us to this weakness in validation that we mentioned that it was expected.

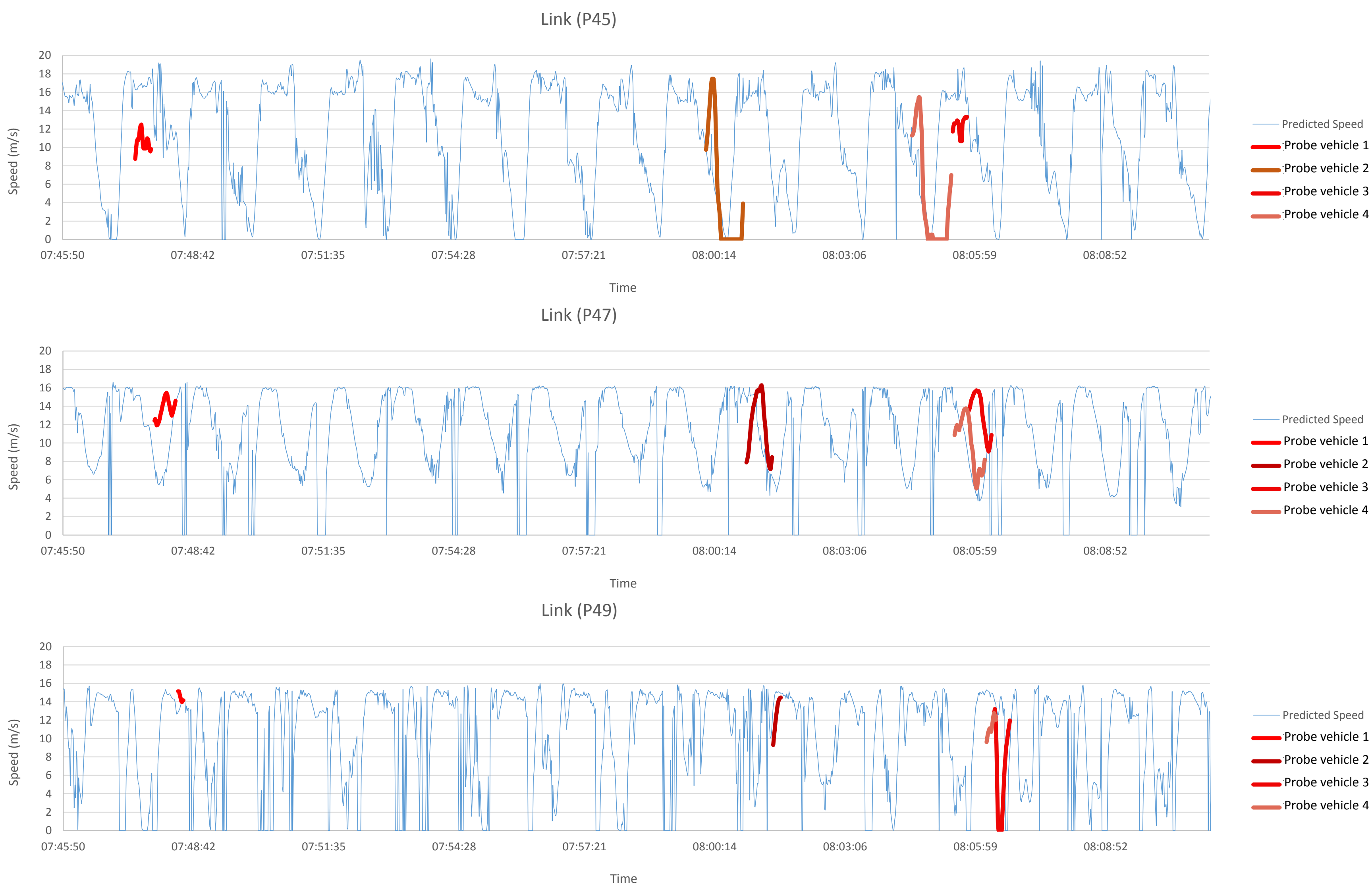


Figure 7.1 Comparing the results of the proposed method by real speeds reported by GPS traces for links P45, P47, and P49

Figure 7.2 Comparing the results of the proposed method by real speeds reported by GPS traces for links P51, P53, and P55

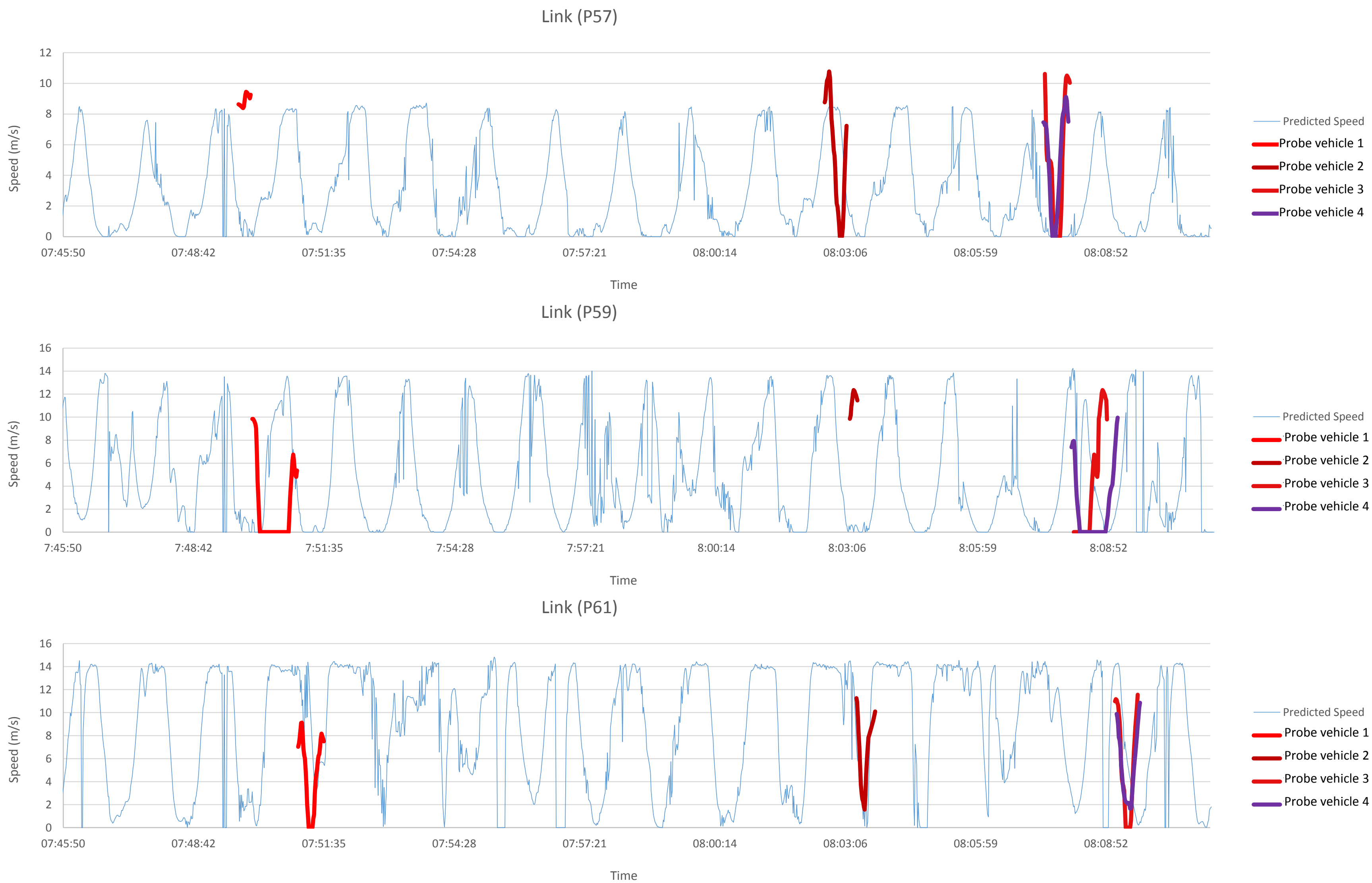


Figure 7.3 Comparing the results of the proposed method by real speeds reported by GPS traces for links P57, P59, and P61

CHAPTER 8 CONCLUSION

In this study, we highlighted the necessity of short-term prediction of traffic conditions in urban road network and its usefulness for user and operators of the network. We reviewed existing literature for short-term traffic prediction and traffic modelling frameworks and concluded that a combination of a Bayesian approach and car-following modelling framework will help us to consider and implement important feedback from the network received as user-centric streaming data provided by probe vehicles that are at individual level. Our proposed method considers two different statuses of streaming data based on the recorded speed of the probe vehicle: zero speed and non-zero speed statuses.

We assumed that Normal distribution in zero speed status can describe the fluctuation of the speed of probe vehicle. Also, we assumed that Poisson distribution is able to explain density change when the recorded speed of probe vehicle is zero. We proposed a Bayesian approach to use the vector of probabilities of measurements and the vectors of probabilities related to conditional model. Also, we assumed the dependency between links and defined the probability of each possible state of traffic conditions simulated by SUMO and calibrated by SA method.

In our case study that is a sub-model of Montréal city, we defined the ideal data that we need to study the problem in the case study. Regarding the limitation in available data we defined 9 possible states of traffic conditions and implemented our proposed method on the available data. The results show us that feeding more data enables us for better modelling the stochastic nature of the problem of traffic conditions prediction in urban road network. Although, the validation in our case study could not confirm our proposed method strongly, but we mentioned that this type of weakness in validation is the result of our limitation in available data that we used in the case study.

This method defines the most probable state at each time-step by applying the streaming data as the experienced conditions. The method is able to apply received data and define the state of urban road network based on recently experienced conditions.

Various patterns of available travel times experienced during morning peak hour make the prediction of the urban road network state challenging. We know that the data availability and the implemented method of predicting the state of urban road network are two important issues in the success of the proposed method. Therefore, the following sections describe the benefits and limitations of the used method and data source.

8.1 Benefits

Traffic conditions prediction method developed in this study, helps us to estimate the state of the urban road network and improves it by new streaming data. We know that the simulated state of the network is affected by the limitations in collecting traditional data. By using streaming data that are available from different sources, we are able to update our models by implementing recently experienced states. Using streaming data provides the possibility of estimating traffic conditions by new prediction models which better simulate and predict what happens in the network.

The method developed in this study defines the state of the entire network based on user-centric streaming data. The method defines the recommended state of the network link by link among possible states. These states are defined based on the availability of travel time data. Therefore, by collecting travel time of links using enough floating cars we can access travel time series of various patterns of traffic behaviour in the urban road network as possible scenarios. We need to know the prior probability of each possible state. Also, we determine the probability of measurements that is related to available streaming data. We developed the conditional model by considering available streaming data and simulation outputs.

The proposed Bayesian approach helps us choose among possible states of traffic conditions. We can develop all ideal possible states using ideal dataset. Therefore, we can consider different patterns of traffic behaviour of the network users which help us in considering different possible scenarios and improving stochastic prediction of traffic conditions in the network.

8.2 Limitations

We mentioned in section 6.1 that feeding more data to the proposed Bayesian framework leads us to predict traffic conditions in urban road network more stochastically. With enough coverage of probe vehicles we can predict traffic condition more precisely. But there are some limitations for collecting enough data for appropriate covering the network by probe vehicles. GPS devices installed in Bus and Taxi fleets enable us to access a source of data. However, the data recorded by the GPS devices in these fleets are not good representative of measurements about the network state. We can think about the other sources of streaming data, for instance, the data produced by the cars that use GPS devices for routing in the network.

The ability of Poisson distribution in modelling the state of the crowded urban roads is challengeable. When density increases, the probability cannot be determined based on the mean value using Poisson formula. For instance, if the mean value is 200, for the density value of more than 133, the value of probability is not calculable by implementing the Poisson formula using usual processors. This limitation in using Poisson distribution is indicated in more detail in Table 8.1.

The proposed method of this thesis has a limitation where it is not possible to predict the

Table 8.1 Limitation in implementing Poisson distribution

λ	Density	Used value
200-207	133 or more	$1.01 * 10^{-7} - 8.97 * 10^{-9}$
193-199	134 or more	$1.04 * 10^{-6} - 2.10 * 10^{-7}$
185-192	135 or more	$1.97 * 10^{-5} - 2.07 * 10^{-6}$
178-184	136 or more	$1.55 * 10^{-4} - 3.48 * 10^{-5}$

short-term future of the urban road network. It is only possible to stochastically define the most probable state of the network at each time-step. It can be mentioned that when we define the most probable state at each time-step, the short-term future of the state can be considered for short-term predicting the future. But we know that there are considerable fluctuations in link's parameters between the minimum and maximum values (that can be defined from the most and the least probable states). Therefore, the prediction of short-term future of urban road network based on streaming data received at each current time-step is not possible by this method and can lead us to unsuitable results. We can define which state at the time of receiving streaming data is the most probable state.

CHAPTER 9 FUTURE STUDIES AND DIRECTIONS

We mentioned in section 8.2 that the proposed method of this study is suitable for predicting traffic conditions and estimating the most probable state using the received streaming data. Regarding this limitation, we can develop the proposed approach by considering the effect of different recommended states of the network by the proposed Bayesian approach at all preceding times from past (for instance, a fixed time) until the current time-step for forecasting the most probable state of the network at short-term future. As an idea, we can consider a combination of each possible state by applying its probability.

SUMO was used in this study to simulate traffic conditions. There are other softwares that could be used. We can use them and compare their results in the simulation part of the proposed approach. We recommend a separate study to review their abilities in microsimulating traffic conditions by focusing on the differences in applied assumptions and the produced outputs. We think the recommended study would be able to help us in understanding the better usage of each available software in different cases. We selected SUMO in this study for some reasons: it uses car-following framework; and it is able to produce outputs at microscopic level (for each vehicle) and at macroscopic level (link by link). But we think that testing ability of the other available softwares helps us to decide about the best software more precisely.

We assumed that when the value of reported speed by probe vehicles becomes zero, link densities fluctuations can be modeled by implementing Poisson distribution. While the Poisson formula is not able to describe the links by high value of density and we know that Poisson distribution is accurate when the values of mean and variance equal, we recommend to evaluate usage of other distributions. For instance, negative binomial distribution can be a choice. This distribution in a specific case can be transfered to Poisson distribution. Negative binomial distribution that considers the impact of difference between the two parameters (mean and variance) of distribution can be examined in the other study.

Also, the reported speed of probe vehicle (when the speed value is none-zero) and average speed of links considered normally distributed. We think that this assumption can be tested in other study. We can analyze recorded data of speed of the probe vehicles and detect the accuracy of Normal distribution for this indicator.

We need to have access to the dataset that highly covers the speed and travel time information in most of the links in the network (e.g. DataMobile http://www.datamobileapp.ca/). Most of drivers use GPS devices for routing in the network. Therefore, accessing to this source of streaming data can help us to achieve more complete data.

To extend this study to a complete network, we need to simulate traffic conditions by using 22 ideal TT datasets mentioned in section 4.2.3. We can collect the related TT data from GPS devices used by drivers in the network and create a historical dataset of experienced traffic conditions (TT and speed) in the network. Using the historical dataset we can define that each day is categorized in which group of 22 groups of days. Also, we can define the probability of each simulated state between 22 states. This dataset needs to be updated over time every year.

We need some output files from simulated states. For a complete network as real case study, we need to know that at each time of a day how many vehicles are located after each probe vehicle at each point of each link as it was described in more detail in section 3.5.2. The output information from the related simulations can be saved in a different database for using in real-time predicting the most probable state of the network at each time of a day.

At each time-step, we must calculate probability of measurements, and measure conditional model. Therefore, the supporting database that we need from all simulated possible states must be ready at each time-step. At each time-step, we only need to use processors to process received streaming data. All the data that we need for calculations that can be independent from received streaming data must manage to be prepared before the main part of calculations that uses streaming data.

BIBLIOGRAPHY

Alecsandru, C. D., 2006. A stochastic mesoscopic cell-transmission model for operational analysis of large-scale transportation networks. Ph.D. thesis, Louisiana State University.

Benekohal, R., Treiterer, J., 1988. CARSIM: Car-following model for simulation of traffic in normal and stop-and-go conditions. Transportation research record (1194).

Brackstone, M., McDonald, M., 1999. Car-following: a historical review. Transportation Research Part F: Traffic Psychology and Behaviour 2 (4), 181–196.
URL http://www.sciencedirect.com/science/article/pii/S136984780000005X

Canaud, M., Mihaylova, L., EL FAOUZI, N. E., Billot, R., Sau, J., Jan. 2013. A Bayesian Approach to Real-time Traffic State Estimation using a Particle PHD Filter with Appropriate Clutter Intensity. In : 92nd Transportation Research Board Annual Meeting.
URL https://hal.archives-ouvertes.fr/hal-00925890

Chandler, R. E., Herman, R., Montroll, E. W., 1958. Traffic dynamics: studies in car following. Operations research 6 (2), 165–184.

Charypar, D., Axhausen, K. W., Nagel, K., 2007. An event-driven parallel queue-based microsimulation for large scale traffic scenarios. ETH, Eidgenössische Technische Hochschule Zürich, IVT.

Chen, H., Rakha, H. A., 2014. Real-time travel time prediction using particle filtering with a non-explicit state-transition model. Transportation Research Part C: Emerging Technologies 43, Part 1, 112–126, special Issue on Short-term Traffic Flow Forecasting.
URL http://www.sciencedirect.com/science/article/pii/S0968090X14000400

City of Montréal, 2016. Comptage des véhicules et des piétons. Data retrieved from: http://donnees.ville.montreal.qc.ca/dataset/comptage-vehicules-pietons.

Coifman, B., 2002. Estimating travel times and vehicle trajectories on freeways using dual loop detectors. Transportation Research Part A: Policy and Practice 36 (4), 351–364.

Coifman, B., Krishnamurthy, S., 2007. Vehicle reidentification and travel time measurement across freeway junctions using the existing detector infrastructure. Transportation Research Part C: Emerging Technologies 15 (3), 135–153.
URL http://www.sciencedirect.com/science/article/pii/S0968090X07000137

Danalet, A., Farooq, B., Bierlaire, M., 2014. A Bayesian approach to detect pedestrian destination-sequences from WiFi signatures. Transportation Research Part C: Emerging Technologies 44, 146–170.
URL http://www.sciencedirect.com/science/article/pii/S0968090X14000874

Dhamaniya, A., Chandra, S., 2013. Speed Characteristics of Mixed Traffic Flow on Urban Arterials. International Journal of Civil, Architectural Science and Engineering 7 (11), 330–335.

Farooq, B., 2013. Prediction of Dynamic Traffic Conditions on Urban Roads Using GPS Traces: A Bayesian updating approach. Tech. rep., École Polytechnique de Montréal, Département des génies civil, géologique et des mines.

Furnish, P., Wignall, D., 2009. Making the Most of Models: Using Models To Develop More Effective Transport Policies And Strategies.

Gazis, D. C., Herman, R., Rothery, R. W., 1961. Nonlinear follow-the-leader models of traffic flow. Operations research 9 (4), 545–567.

Gipps, P., 1981. A behavioural car-following model for computer simulation. Transportation Research Part B: Methodological 15 (2), 105–111.
URL `http://www.sciencedirect.com/science/article/pii/0191261581900370`

Gkiotsalitis, K., Stathopoulos, A., 2015. A utility-maximization model for retrieving users' willingness to travel for participating in activities from big-data. Transportation Research Part C: Emerging Technologies 58, Part B, 265–277.

Hanken, A., Rockwell, T., 1967. A MODEL OF CAR FOLLOWING DERIVED EMPIRICALLY BY PIECE-WISE REGRESSION ANALYSIS. IN VEHICULAR TRAFFIC SCIENCE. In : Proceedings of the Third International Symposium on the Theory of Traffic Flow.

Helly, W., 1959. Simulation of bottlenecks in single-lane traffic flow. In : Proceedings of the Symposium on Theory of Traffic Flow, Research Laboratories, General Motors. New York: Elsevier, pp. 207–238.

Herman, R., Montroll, E. W., Potts, R. B., Rothery, R. W., 1959. Traffic dynamics : analysis of stability in car following. Operations research 7 (1), 86–106.

Herring, R. J., 2010. Real-time traffic modeling and estimation with streaming probe data using machine learning. Ph.D. thesis, University of California, Berkeley.

Hofleitner, A., Herring, R., Abbeel, P., Bayen, A., 2012. Learning the dynamics of arterial traffic from probe data using a dynamic Bayesian network. IEEE Transactions on Intelligent Transportation Systems 13 (4), 1679–1693.

Jandik, P., 2015. Prediction of Dynamic Traffic Conditions on Urban Roads Using GPS Traces. Master's thesis, Czech Technical Universiy in Prague, Faculty of Transportation Sciences.

Jenelius, E., Koutsopoulos, H. N., 2013. Travel time estimation for urban road networks using low frequency probe vehicle data. Transportation Research Part B: Methodological 53, 64 – 81.
URL http://www.sciencedirect.com/science/article/pii/S0191261513000489

Kikuchi, S., Chakroborty, P., 1992. Car-following model based on fuzzy inference system. Transportation Research Record, 82–82.

Kometani, E., Sasaki, T., 1959. Dynamic behaviour of traffic with a non-linear spacing-speed relationship. In : Proceedings of the Symposium on Theory of Traffic Flow, Research Laboratories, General Motors. New York: Elsevier, pp. 105–119.

Krajzewicz, D., Erdmann, J., Behrisch, M., Bieker, L., December 2012. Recent Development and Applications of SUMO - Simulation of Urban MObility. International Journal on Advances in Systems and Mearsurements 5 (3&4), 128–138.

Li, L., Su, X., Wang, Y., Lin, Y., Li, Z., Li, Y., 2015. Robust causal dependence mining in big data network and its application to traffic flow predictions. Transportation Research Part C: Emerging Technologies 58, Part B, 292–307.

Maerivoet, S., De Moor, B., 2005. Cellular automata models of road traffic. Physics Reports 419 (1), 1–64.

Martin, P. T., Feng, Y., Wang, X., et al., 2003. Detector technology evaluation. Tech. rep., Citeseer.

Michaels, R., 1963. Perceptual Factors in Car Following. In : Proceedings of the 2nd International Symposium on the Theory of Road Traffic Flow, OECD, Paris. pp. 44–59.

Nantes, A., Billot, R., Miska, M. P., Chung, E., 2013. Bayesian Inference of Traffic Volumes Based on Bluetooth Data. In : Transportation Research Board 92nd Annual Meeting. No. 13-4838.

Park, T., Lee, S., 2004. A Bayesian approach for estimating link travel time on urban arterial road network. In : Computational Science and Its Applications–ICCSA 2004. Springer, pp. 1017–1025.

Pascale, A., Gomes, G., Nicoli, M., 2013. Estimation of highway traffic from sparse sensors: Stochastic modeling and particle filtering. In : 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp. 6158–6162.

Polson, N., Sokolov, V., 2014. Bayesian Analysis of Traffic Flow on Interstate I-55: The LWR Model. arXiv preprint arXiv:1409.6034.

Shi, Q., Abdel-Aty, M., 2015. Big Data applications in real-time traffic operation and safety monitoring and improvement on urban expressways. Transportation Research Part C: Emerging Technologies 58, Part B, 380–394.

Staley, S., 2012. Practical Strategies for Reducing Congestion and Increasing Mobility for Chicago. Reason Foundation.

Staňková, K., De Schutter, B., Sep. 2010. On freeway traffic density estimation for a jump Markov linear model based on Daganzo's cell transmission model. In : Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems (ITSC 2010). Madeira Island, Portugal, pp. 13–18.

Stone, J. V., 2013. Bayes' Rule: A Tutorial Introduction to Bayesian Analysis. Sebtel Press.

Vij, A., Shankari, K., 2015. When is big data big enough? Implications of using GPS-based surveys for travel demand analysis. Transportation Research Part C: Emerging Technologies 56, 446–462.

Vlahogianni, E. I., Karlaftis, M. G., Golias, J. C., 2014. Short-term traffic forecasting: Where we are and where we're going. Transportation Research Part C: Emerging Technologies 43, 3–19.

Xing, J., 1995. A Parameter Identification of a Car-Following Model. In : Steps Forward. Intelligent Transport Systems World Congress. No. Volume 4.

APPENDIX A - Detail of link characteristics in the chosen arterials (source: Jandik (2015))

Table A.1 Characteristic of the links in Acadie-Parc -SB direction

Node No.	Node name	Length (m)	No. of Lanes	Parking facility	FFS [I]
1	Acadie/Liege	-			
2	Acadie/Jarry	521	3	-	50
3	Acadie/St. Roch	539	3	-	54
4	Acadie/Jean-Talon	469	3	-	49
5	Acadie/Beaumont	228	1 of 2 [II]	-	43
6	Beaumont/Outremont	262	1 of 2	-	43
7	Beaumont/Ourbes	272	1 of 2	-	47
8	Beaumont/Parc	208	1 of 2	-	42
9	Parc/Beaubien	495	2	-	53
10	Parc/Van Horn	238	2(-1) [III]	P	71
11	Parc/Bernard	435	2(-1)	P	60
12	Parc/(Bernard-St.Viateur)	158	2(-1)	P	58
13	Parc/St.Viateur	139	2(-1)	P	49
14	Parc/(St.Viateur-Fairmount)	242	2(-1)	P	59
15	Parc/Fairmount	175	2(-1)	P	48
16	Parc/Laurier	163	2(-1)	P	33
17	Parc/St. Joseph	127	2(-1)	P	52
18	Parc/Villeneuve	191	2(-1)	P	54
19	Parc/Mt. Royal	294	2(-1)	P	50
20	Parc/Ch. Ste. Catherine	87	2	-	54
21	Parc/(Ch. Ste. Catherine-Duluth)	300	3	P	60

I. Free flow speed that was defined by the maximum average speed on the link measured by a test car.
II. 1 of 2 lines of the related two-way street can be used for this direction.
III. There are two lines available for this direction most of the time. But in the evening rush hours one line is used for the other direction.

Table A.2 Characteristic of the links in Acadie-Parc -NB direction

Node No.	Node name	Length (m)	No. of Lanes	Parking facility	FFS
1	Parc/(Ch. Ste. Catherine-Duluth)	-			
2	Parc/Mt. Royal	394	2+2 TL [I]	-	62
3	Parc/Villeneuve	294	1(+1) [II]	P	51
4	Parc/St. Joseph	173	1(+1)	P	46
5	Parc/Laurier	134	1(+1)	P	51
6	Parc/Fairmount	167	1(+1)	P	45
7	Parc/(St.Viateur-Fairmount)	211	1(+1)	P	56
8	Parc/St.Viateur	200	1(+1)	P	48
9	Parc/(Bernard-St.Viateur)	153	1(+1)	P	47
10	Parc/Bernard	143	1(+1)	P	50
11	Parc/Van Horne	434	1(+1)	P	50
12	Prac/Beaubien	256	2	-	56
13	Parc/Beaumont	464	2 of 4	-	60
14	Parc/Jean-Talon	354	2 of 4	(P) [III]	49
15	Parc/Hutchisont	74	2 of 4	-	38
16	Querbes	132	2 of 4	2*P [IV]	38
17	Parc/Bloomfield	152	2 of 4	2*P	51
18	Parc/Wiseman	176	2 of 4	2*P	32
19	Parc/Acadie	172	2 of 4	2*P	34
20	Acadie/St. Roch	487	3	(P)	45
21	Acadie/Jarry	529	3	(P)	51
22	Acadie/ Liege	491	3	(P)	58

I. There are two lines available for this direction. Also, two additional lines is available for left turns.

II. There are one line available for this direction most of the time. But in the evening rush hours one additional line is used for this direction.

III. Limited Parallel parking on the right side of the street.

IV. Parallel parking on both sides of the two-way street.

Table A.3 Characteristic of the links in St. Urbain (SB direction)

Node No.	Node name	Length (m)	No. of Lanes	Parking facility	FFS
1	Liege	-			
2	Guizot	253	2	P	38
3	Jarry	274	2	P	48
4	Gounod	267	2	P	48
5	Villeray	257	2	P	56
6	Faillon	243	2	P	56
7	de Castelnau	178	3	-	43
8	Jean-Talon	221	3	P(L) ^{I}	31
9	Mozart	273	2	P(L+R) ^{II}	27
10	St. Zotique	302	2	P(L+R)	52
11	Beaubien	268	2	P(L+R)	49
12	Bernard	676	2	-	37
13	St. Viateur	320	2	P(L+R),C ^{III}	50
14	Fairmont	406	2	P(L+R),C	53
15	Laurier	155	2	P(L+R),C	50
16	St. Joseph	131	2	P(L+R),C	50
17	Villeneuve	191	2	P(L+R),C	52
18	Mt. Royal	284	2	P(L+R),C	48
19	Marie-Anne	222	2	P(L+R),C	49
20	Rachel	214	2	P(L+R),C	50

I. Parallel parking on the left side of the street.

II. Parallel parking on both sides of the street.

III. A cycle lane on the right side of the carriageway.

Table A.4 Characteristic of the links in St. Laurent (NB direction)

Node No.	Node name	Length (m)	No. of Lanes	Parking facility	FFS
1	Rachel	-			
2	Marie-Anne	228	2	P(L+R)	56
3	Mt. Royal	208	2	P(L+R)	57
4	Villeneuve	266	2	P(L+R)	45
5	St. Joseph	185	2	P(L+R)	45
6	Laurier	146	2	P(L+R)	51
7	Fairmont	161	2	P(L+R)	55
8	St. Viateur	418	2	P(L+R)	59
9	Bernard	308	2	P(L+R)	60
10	Beaubien Est	566	2	P(L+R)	50
11	Beaubien Ouest	81	2	P(L+R)	50
12	St. Zotique	255	2	P(L+R)	48
13	Dante	178	2	P(L+R)	49
14	Mozart	120	2	P(L+R)	55
15	Jean-Talon	270	2	P(L+R)	57
16	de Castelnau	153	2	P	55
17	Faillon	220	2	P	55
18	Villeray	242	2	P	58
19	Gounod	269	2	P	58
20	Jarry	238	2	P	49
21	Guizot	283	2	P	58
22	Liege	250	2	P	54

APPENDIX B - Structure of file of recorded GPS traces

Table B.1 Some rows of a recorded GPS traces as an example

Time	Latitude	Longitude	Speed	HDOP	NamSats	Quality	Node	Run
previous Lines								
2005-02-18 12:08:18 PM	45308820	73350756	28.28	1.2	7	2	0	20
2005-02-18 12:08:19 PM	45308858	73350836	28.21	1.2	7	2	0	20
2005-02-18 12:08:20 PM	45308896	73350917	28.29	1.2	7	2	0	20
2005-02-18 12:08:21 PM	45308934	73351001	28.88	1.2	7	2	0	20
2005-02-18 12:08:22 PM	45308973	73351086	30.40	1.2	7	2	1	20
following lines								

APPENDIX C - Detail of start time of recording trajectories in the chosen arterials

Table C.1 Start time of recording trajectories in Acadie-Parc -NB direction

Start Time	Start Time	Start Time	Start Time
02-18 12:08:08	02-18 12:05:29	02-22 11:57:25	02-22 12:32:14
02-21 06:48:40	02-21 07:21:26	02-21 08:03:20	02-22 06:58:51
02-22 07:38:46	02-22 08:18:57	02-22 12:02:47	02-22 12:43:22
02-21 11:29:42	02-21 12:09:56	02-21 12:59:26	02-22 06:50:49
02-22 07:21:23	02-22 08:00:47	02-22 11:52:09	02-22 12:34:42
02-22 01:13:25	02-22 03:30:27	02-22 04:22:11	02-22 05:08:35
02-22 05:49:30	02-22 15:56:50	02-22 16:43:29	02-22 17:26:07
02-22 06:06:18	02-22 16:03:12	02-22 16:54:07	02-22 17:35:39
02-22 18:19:21	02-23 03:31:20	02-23 04:11:11	02-23 04:55:37
02-23 05:48:45	02-23 16:44:32	02-23 18:16:40	02-24 12:35:03
02-24 11:56:12	02-24 12:30:49	02-24 13:05:46	03-07 06:30:32
03-07 07:30:57	03-07 07:41:47	03-07 08:20:09	03-07 11:34:00
03-11 06:49:42	03-11 07:23:18	03-11 08:02:55	03-11 08:38:29
03-11 12:03:24	03-11 12:50:42	01-11 06:34:46	01-11 07:04:08
01-11 07:38:12	01-11 08:12:34	01-11 08:53:25	01-11 11:35:05
01-11 12:27:32	01-11 13:12:58	04-27 07:02:31	04-27 07:36:21
04-27 08:12:52	04-27 08:51:52		

Table C.2 Start time of recording trajectories in Acadie-Parc -SB direction

Start Time	Start Time	Start Time	Start Time
02-18 11:37:11	02-18 12:38:14	02-18 11:31:17	02-18 12:34:38
02-21 07:32:12	02-21 08:26:11	02-22 12:14:00	02-22 12:57:37
02-21 06:32:25	02-21 07:05:01	02-21 07:38:45	02-22 06:43:04
02-22 07:20:19	02-22 07:57:21	02-22 08:40:26	02-22 11:43:26
02-22 12:22:17	02-22 13:06:00	02-21 11:48:52	02-21 12:35:47
02-21 13:21:30	02-22 06:34:24	02-22 07:05:08	02-22 07:37:22
02-22 11:33:25	02-22 12:15:33	02-22 12:52:38	02-22 03:57:18
02-22 04:42:16	02-22 05:28:52	02-22 06:09:38	02-22 16:34:58
02-22 16:20:45	02-22 17:02:36	02-22 17:44:40	02-22 15:40:08
02-22 16:30:38	02-22 17:12:38	02-22 17:58:12	02-23 03:48:39
02-23 04:29:07	02-23 05:22:31	02-23 06:06:25	02-23 15:40:13
02-23 16:22:12	02-23 17:04:13	02-23 17:51:33	02-24 12:54:51
02-24 11:33:35	02-24 12:12:24	02-24 12:45:49	03-07 06:44:46
03-07 07:18:43	03-07 07:58:59	03-07 08:37:43	03-11 06:33:11
03-11 07:04:49	03-11 07:43:31	03-11 08:18:45	03-11 11:37:04
03-11 12:20:30	03-11 13:04:15	01-11 06:48:16	01-11 07:20:11
01-11 07:53:58	01-11 08:31:02	01-11 11:55:27	01-11 12:42:40
01-11 13:27:52	04-27 07:18:37	04-27 07:52:12	04-27 08:32:34

Table C.3 Start time of recording trajectories in St. Laurent-NB direction

Start Time	Start Time	Start Time	Start Time
02-18 15:35:53	02-18 16:44:00	02-18 17:41:01	02-18 03:41:24
02-18 04:50:42	02-18 05:46:34	02-23 06:30:21	02-23 07:02:16
02-23 07:43:57	02-23 08:35:49	02-23 08:42:26	02-23 11:31:56
02-23 12:19:38	02-23 13:06:34	02-24 06:30:26	02-24 07:09:05
02-24 08:03:47	02-24 06:51:45	02-24 07:36:57	02-24 08:29:20
02-24 04:46:27	02-24 05:38:44	02-25 06:3:35	02-25 07:07:46
02-25 07:50:30	02-24 15:34:35	02-24 16:20:10	02-24 17:06:40
02-24 18:03:53	02-25 06:37:03	02-25 07:10:31	02-25 07:58:36
02-25 04:01:24	02-25 05:06:34	02-25 05:59:55	02-25 11:30:39
02-25 12:19:56	02-25 11:58:39	02-25 12:43:46	02-25 15:54:15
02-25 16:49:21	02-25 17:45:28	03-07 12:27:54	

Table C.4 Start time of recording trajectories in St. Urbain-SB direction

Start Time	Start Time	Start Time	Start Time
02-18 16:15:55	02-18 17:14:47	02-18 04:21:16	02-18 05:20:01
02-23-06:46:40	02-23 07:21:58	02-23 08:10:52	02-23 07:25:31
02-23 08:16:18	02-23 11:42:40	02-23 12:31:26	02-23 13:21:53
02-23 12:01:22	02-23 12:46:43	02-24 06:50:01	02-24 07:29:25
02-24 08:26:43	02-24 06:35:21	02-24 07:17:17	02-24 08:01:13
02-24 03:32:07	02-24 04:24:44	02-24 05:15:19	02-25 06:50:03
02-25 07:29:12	02-25 08:16:14	02-24 16:00:44	02-24 16:46:14
02-24 17:35:18	02-25 06:53:23	02-25 07:35:20	02-25 08:21:10
02-25 03:37:04	02-25 04:38:05	02-25 05:38:19	02-25 11:57:50
02-25 12:48:19	02-25 11:32:20	02-25 12:22:52	02-25 15:31:59
02-25 16:26:08	02-25 17:21:23	03-07 12:06:07	03-07 12:56:37

APPENDIX D - Counted traffic data at intersections from 2003 dataset of traffic survey

Table D.1 Counted traffic data at intersections

ID	Intersection	Date	Heure	ND	NTD	NG	ED	ETD	EG	SD	STD	SG	OD	OTD	OG
101	Acadie/Jean-Talon	2003-09-15	07:30	187	213	101	31	246	2	2	173	0	8	99	4
101	Acadie/Jean-Talon	2003-09-15	07:45	168	231	106	36	249	3	1	140	0	5	112	0
101	Acadie/Jean-Talon	2003-09-15	08:00	151	218	97	30	251	1	4	165	0	5	141	2
101	Acadie/Jean-Talon	2003-09-15	08:15	154	209	97	33	248	1	2	154	0	4	135	2
102	Acadie/Beaumont	2003-09-15	07:30	73	4	103	32	76	3	1	6	12	0	24	58

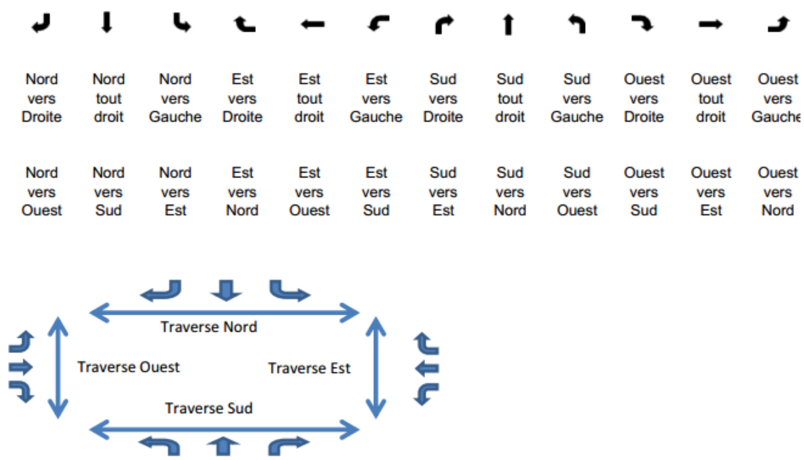


Figure D.1 Description of possible directions in counting traffic volumes at each intersection

APPENDIX E - Analysis of average speed (source: Jandik (2015))

Figure E.1 Average experienced speed within morning rush hour on 2005-02-23

APPENDIX F - Analysis of speed and travel time distribution within three time spans of day

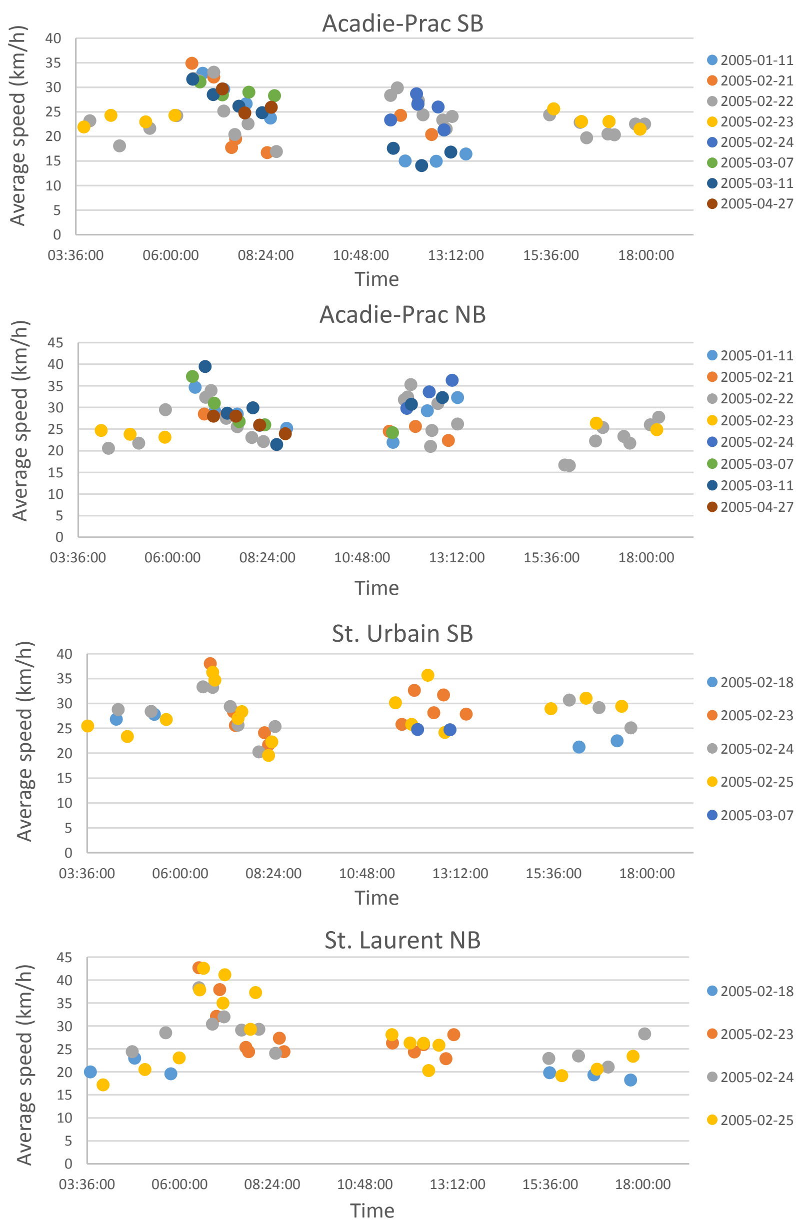


Figure F.1 Average experienced speed in different time spans of the days

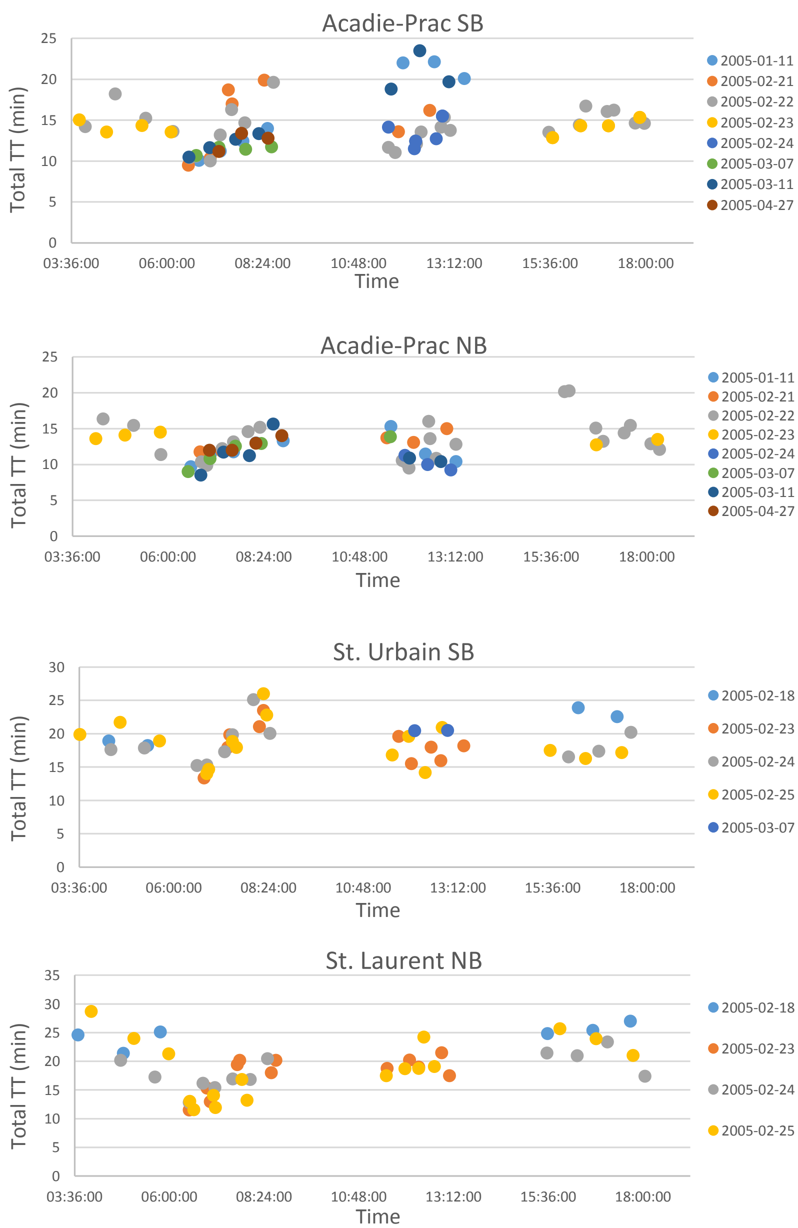


Figure F.2 Average experienced total travel time in different time spans of the days

APPENDIX G - Turning probabilities defined in order to create input files required to simulate traffic conditions

Table G.1 Traffic pattern in turning at intersections using counted data

Intersection	ND	NTD	NG	ED	ETD	EG	SD	STD	SG	OD	OTD	OG
Jean-Talon/Wiseman	-	-	-	0.036	0.964	0.000	0.258	0.441	0.301	0.000	0.888	0.112
Jean-Talon/Parc	-	-	-	0.000	0.808	0.192	0.202	0.000	0.798	0.475	0.525	0.000
Mozart/St-Laurent inter. S	-	-	-	1.000	0.000	0.000	0.014	0.932	0.054	0.000	0.304	0.696
Dante/Saint-Laurent	-	-	-	1.000	0.000	0.000	0.024	0.976	0.000	-	-	-
St-Laurent/St-Zotique inter. S	-	-	-	0.358	0.642	0.000	0.060	0.923	0.017	0.000	0.000	1.000
Beaubien/St-Laurent inter. N	-	-	-	-	-	-	0.000	0.656	0.344	0.000	0.000	1.000
Beaubien/St-Laurent inter. S	-	-	-	1.000	0.000	0.000	0.067	0.933	0.000	-	-	-
Bernard / Saint-Laurent	-	-	-	-	-	-	0.008	0.956	0.036	0.000	0.645	0.355
Bernard / Saint-Laurent	-	-	-	-	-	-	0.008	0.956	0.036	0.000	0.645	0.355
Saint-Laurent / Saint-Viateur	-	-	-	0.587	0.413	0.000	0.053	0.881	0.066	0.015	0.443	0.542
Fairmount / Saint-Laurent	-	-	-	0.347	0.653	0.000	0.016	0.935	0.049	0.000	0.322	0.678
Saint-Joseph / Saint-Laurent	-	-	-	0.135	0.865	0.000	0.073	0.902	0.025	0.002	0.859	0.139
Saint-Laurent / Villeneuve	-	-	-	0.290	0.703	0.007	0.000	0.977	0.023	-	-	-
Mont-Royal / Saint-Laurent	-	-	-	0.149	0.851	0.000	0.045	0.862	0.093	0.000	0.523	0.477
Marie-Anne / Saint-Laurent	-	-	-	0.941	0.000	0.059	0.040	0.960	0.000	0.000	0.607	0.393
Rachel / Saint-Laurent	-	-	-	0.403	0.597	0.000	0.038	0.908	0.054	0.000	0.874	0.126
Jean-Talon / Saint-Laurent	1.000	0.000	0.000	0.088	0.912	0.000	0.038	0.776	0.186	0.000	0.998	0.002
Castelnau / Saint-Laurent	0.953	0.017	0.030	0.370	0.630	0.000	0.024	0.967	0.009	-	-	-
Acadie / Beaumont	0.425	0.019	0.557	0.274	0.711	0.015	0.227	0.318	0.455	0.024	0.305	0.671
Beaumont / Outremont	0.406	0.000	0.594	0.004	0.996	0.000	-	-	-	0.000	1.000	0.000
Acadie / Jean-Talon	0.342	0.451	0.208	0.115	0.879	0.006	0.014	0.986	0.000	0.043	0.942	0.015
Clark / Jean-Talon	0.159	0.822	0.020	0.003	0.856	0.141	-	-	-	0.248	0.752	0.000
Beaumont / Querbes	0.154	0.007	0.838	0.165	0.818	0.018	0.333	0.556	0.111	0.005	0.954	0.041
Bernard / Saint-Urbain	0.116	0.879	0.004	0.000	0.595	0.405	-	-	-	0.918	0.082	0.000
Beaubien / Clark	0.109	0.881	0.010	0.000	0.822	0.178	-	-	-	0.690	0.310	0.000
Parc / Van Horne	0.101	0.899	0.000	0.137	0.863	0.000	0.082	0.915	0.003	0.057	0.941	0.001
Laurier / Saint-Laurent	0.100	0.900	0.000	0.125	0.875	0.000	0.034	0.901	0.064	0.003	0.743	0.253
Bernard / Parc	0.098	0.902	0.000	0.079	0.668	0.253	0.078	0.922	0.000	0.255	0.382	0.363
Bloomfield / Jean-Talon	0.081	0.337	0.582	0.000	0.943	0.057	-	-	-	0.023	0.977	0.000
Jarry / Saint-Laurent	0.078	0.922	0.000	0.138	0.686	0.176	0.033	0.893	0.075	0.339	0.578	0.083
Mont-Royal / Saint-Urbain	0.076	0.876	0.048	0.000	0.804	0.196	-	-	-	0.386	0.614	0.000
Mont-Royal / Parc	0.063	0.936	0.001	0.106	0.595	0.298	0.300	0.697	0.004	0.027	0.722	0.251
Jean-Talon / Querbes	0.059	0.598	0.343	0.091	0.909	0.000	0.083	0.833	0.083	0.055	0.914	0.031
Fairmount / Saint-Urbain	0.057	0.913	0.030	0.000	0.730	0.270	-	-	-	0.455	0.545	0.000
Saint-Urbain / Saint-Viateur	0.050	0.843	0.107	0.006	0.733	0.262	-	-	-	0.299	0.692	0.009
Faillon / Saint-Laurent	0.048	0.906	0.046	0.325	0.163	0.512	0.021	0.970	0.009	0.120	0.326	0.554
Parc / Saint-Viateur	0.045	0.954	0.001	0.202	0.562	0.236	0.083	0.915	0.002	0.139	0.644	0.218
Fairmount / Parc	0.043	0.957	0.000	0.182	0.491	0.327	0.032	0.966	0.002	0.287	0.228	0.485
Clark / Saint-Zotique	0.041	0.935	0.024	0.000	0.543	0.457	-	-	-	0.213	0.787	0.000
Saint-Urbain / Villeneuve	0.040	0.959	0.001	0.019	0.497	0.484	-	-	-	1.000	0.000	0.000
Hutchison / Jean-Talon	0.037	0.059	0.903	0.182	0.792	0.026	-	-	-	0.034	0.951	0.015
Laurier / Parc	0.034	0.966	0.000	0.062	0.689	0.249	0.119	0.879	0.001	0.185	0.733	0.082
Beaubien / Parc	0.034	0.800	0.165	0.423	0.103	0.474	0.105	0.846	0.048	0.286	0.367	0.347
Laurier / Saint-Urbain	0.032	0.948	0.020	0.000	0.776	0.224	-	-	-	0.316	0.684	0.000
Clark / Mozart	0.030	0.943	0.026	0.000	0.510	0.490	0.000	0.000	1.000	0.250	0.750	0.000
Saint-Joseph / Saint-Urbain	0.027	0.873	0.100	0.006	0.713	0.280	-	-	-	0.090	0.910	0.000
Parc / Villeneuve	0.024	0.976	0.000	0.208	0.464	0.327	0.034	0.964	0.001	0.189	0.566	0.245
Rachel / Saint-Urbain	0.020	0.876	0.104	0.000	0.293	0.707	-	-	-	-	-	-
Parc / Saint-Joseph	0.006	0.993	0.001	0.034	0.611	0.355	0.217	0.783	0.000	0.158	0.789	0.053
Beaumont / Parc	0.003	0.997	0.000	0.143	0.143	0.714	0.014	0.528	0.458	0.932	0.027	0.041
Marie-Anne / Saint-Urbain	0.000	0.960	0.040	-	-	-	-	-	-	0.706	0.294	0.000
Acadie / Jarry	0.000	0.846	0.154	0.148	0.000	0.852	0.067	0.933	0.000	-	-	-
Saint-Laurent / Villeray	0.000	0.967	0.033	0.384	0.004	0.612	0.029	0.971	0.000	-	-	-

APPENDIX H - Introducing characteristics of the nodes and the links used in simulating states by SUMO

Node No.	Intersection name
111	Hatchson/Jean-Talon
112	Jean-Talon/Parc
113	Beaumont/Parc
114	Beaubien Ouest/Parc
115	Van-Horn / Parc
116	Bernard / Parc
117	Bernard & Saint Viateur
118	Saint Viateur/Parc
119	Saint Viateur & Fairmont
120	Fairmont/Parc
121	Laurier/Parc
122	Saint Joseph/Parc
123	Villeneuve/Parc
124	Mont-Royal/Parc
125	Jeane-Mance/Mont-Royal

Node No.	Intersection name
204	Jean-Talon/Clark
205	Mozart Ouest/Clark
206	Saint Zotiqe/Clark
207	Beaubien Ouest/Clark
208	Bernard /Urbain
209	Saint Viateur/Urbain
210	Fairmont/Urbain
211	Laurier/Urbain
212	Saint Joseph/Urbain
213	Villeneuve/Urbain
214	Mont-Royal/Urbain
215	Marie-Anne/Urbain
216	Rachel /Urbain

Node No.	Intersection name
304	Jean-Talon/StLaurent
305	Mozart Est/StLaurent
306	Mozart Ouest/StLaurent
307	Dante/StLaurent
308	Saint Zotiqe/StLaurent
309	Baubien Ouest/StLaurent
310	Beaubien Est/StLaurent
311	Bernard /StLaurent
312	Saint Vlateur/StLaurent
313	Fairmont/StLaurent
314	Laurier/StLaurent
315	Saint Joseph/StLaurent
316	Villeneuve/StLaurent
317	Mont-Royal/StLaurent
318	Marie-Anne/StLaurent
319	Rachel /StLaurent

Figure H.1 Node numbers and the related intersections

Link Name		Start Node	End Node	From	To
P43	43,43.370	113	114	Beaumont/Parc	Beaubien Ouest/Parc
P45		114	115	Beaubien Ouest/Parc	Van-Horn / Parc
P47		115	116	Van-Horn / Parc	Bernard / Parc
P49		116	117	Bernard / Parc	Bernard & Saint Viateur
P51		117	118	Bernard & Saint Viateur	Saint Viateur/Parc
P53		118	119	Saint Viateur/Parc	Saint Viateur & Fairmont
P55		119	120	Saint Viateur & Fairmont	Fairmont/Parc
P57		120	121	Fairmont/Parc	Laurier/Parc
P59		121	122	Laurier/Parc	Saint Joseph/Parc
P61		122	123	Saint Joseph/Parc	Villeneuve/Parc
P63		123	124	Villeneuve/Parc	Mont-Royal/Parc

Link Name		Start Node	End Node	From	To
P64	64,64.50	124	123	Mont-Royal/Parc	Villeneuve/Parc
P62		123	122	Villeneuve/Parc	Saint Joseph/Parc
P60		122	121	Saint Joseph/Parc	Laurier/Parc
P58		121	120	Laurier/Parc	Fairmont/Parc
P56		120	119	Fairmont/Parc	Saint Viateur & Fairmont
P54		119	118	Saint Viateur & Fairmont	Saint Viateur/Parc
P52		118	117	Saint Viateur/Parc	Bernard & Saint Viateur
P50	50,50.106	117	116	Bernard & Saint Viateur	Bernard / Parc
P48	48,48.306	116	115	Bernard / Parc	Van-Horn / Parc
P46	46,46.190	115	114	Van-Horn / Parc	Beaubien Ouest/Parc
P44		114	113	Beaubien Ouest/Parc	Beaumont/Parc
P42	42,42.190	113	112	Beaumont/Parc	Jean-Talon/Parc
P39		112	111	Jean-Talon/Parc	Hatchson/Jean-Talon

Figure H.2 Link numbers and the related intersections in Acadie-Parc route

Link Name		Start Node	End Node	From	To
S7		204	205	Jean-Talon/Clark	Mozart Ouest/Clark
S9		205	206	Mozart Ouest/Clark	Saint Zotiqe/Clark
S11		206	207	Saint Zotiqe/Clark	Beaubien Ouest/Clark
S13	13, 13.605	207	208	Beaubien Ouest/Clark	Bernard /StUrbain
S15		208	209	Bernard/StUrbain	Saint Viateur/StUrbain
S17		209	210	Saint Viateur/StUrbain	Fairmont/StUrbain
S19		210	211	Fairmont/StUrbain	Laurier/StUrbain
S21		211	212	Laurier/StUrbain	Saint Joseph/StUrbain
S23		212	213	Saint Joseph/StUrbain	Villeneuve/StUrbain
S25		213	214	Villeneuve/StUrbain	Mont-Royal/StUrbain

Link Name		Start Node	End Node	From	To
S32		317	316	Mont-Royal/StLaurent	Villeneuve/StLaurent
S30		316	315	Villeneuve/StLaurent	Saint Joseph/StLaurent
S28		315	314	Saint Joseph/StLaurent	Laurier/StLaurent
S26		314	313	Laurier/StLaurent	Fairmont/StLaurent
S24		313	312	Fairmont/StLaurent	Saint Vlateur/StLaurent
S22		312	311	Saint Viateur/StLaurent	Bernard /StLaurent
S20		311	310	Bernard /StLaurent	Beaubien Est/StLaurent
S18		310	309	Beaubien Est/StLaurent	Baubien Ouest/StLaurent
S16		309	308	Baubien Ouest/StLaurent	Saint Zotiqe/StLaurent
S14		308	307	Saint Zotiqe/StLaurent	Dante/StLaurent
S12		307	306	Dante/StLaurent	Mozart Ouest/StLaurent
S10	10,8,8.19	306	305	Mozart Ouest/StLaurent	Mozart Est/StLaurent
S8		305	304	Mozart Est/StLaurent	Jean-Talon/StLaurent

Figure H.3 Link numbers and the related intersections in St. Laurent - St. Urbain route

APPENDIX I - The codes used in running SUMO to define the probe vehicles and travel times of the links

Introducing the probe vehicles in state 2 as an instance

```
<routes>
<vType id="probe"/>
<route id="Parc-SB" edges="P36 P36.695 P43 P43.370 P45 P47 P49 P51 P53 P55 P57 P59
P61 P63 P65"/>
<route id="Parc-NB" edges="P66 P64 P64.50 P62 P60 P58 P56 P54 P52 P50 P50.106 P48
P48.306 P46 P46.190 P44 P42 P42.190 P39 P33"/>
<route id="Urbain-SB" edges="S5 S5.780 S7 S9 S11 S13 S13.605 S15 S17 S19 S21 S23 S25
S27 S29 S31"/>
<route id="Laurent-NB" edges="S38 S36 S34 S32 S30 S28 S26 S24 S22 S20 S18 S16 S14 S12
S10 S8 S8.190 S6"/>
<flows>
< !– Parc - SB –>
<interval begin="1770" end="1830">
<flow id="U-SB-1" type="probe" from="S5" number="10" departlane="best"
departPos="random-free" departSpeed="8.33" speedDev="0.1" color="1,0,0"
route="Urbain-SB"/>
</interval>
<interval begin="2130" end="2190">
<flow id="L-NB-1" type="probe" from="S38" number="10" departlane="best"
departPos="random-free" departSpeed="8.33" speedDev="0.1" color="1,0,0"
route="Laurent-NB"/>
</interval>
<interval begin="2190" end="2250">
<flow id="P-SB-1" type="probe" from="P36" number="10" departlane="best"
departPos="random-free" departSpeed="8.33" speedDev="0.1" color="1,0,0"
route="Parc-SB"/>
</interval>
<interval begin="2250" end="2310">
<flow id="P-NB-1" type="probe" from="P66" number="10" departlane="best"
departPos="random-free" departSpeed="8.33" speedDev="0.1" color="1,0,0"
```

route="Parc-NB"/>
</interval>
<interval begin="3390" end="3450">
<flow id="P-SB-2" type="probe" from="P36" number="10" departlane="best"
departPos="random-free" departSpeed="8.33" speedDev="0.1" color="1,0,0"
route="Parc-SB"/>
</interval>
<interval begin="3570" end="3630">
<flow id="P-NB-2" type="probe" from="P66" number="10" departlane="best"
departPos="random-free" departSpeed="8.33" speedDev="0.1" color="1,0,0"
route="Parc-NB"/>
</interval>
<interval begin="3630" end="3690">
<flow id="U-SB-2" type="probe" from="S5" number="10" departlane="best"
departPos="random-free" departSpeed="8.33" speedDev="0.1" color="1,0,0"
route="Urbain-SB"/>
</interval>
<interval begin="3750" end="3810">
<flow id="L-NB-2" type="probe" from="S38" number="10" departlane="best"
departPos="random-free" departSpeed="8.33" speedDev="0.1" color="1,0,0"
route="Laurent-NB"/>
</interval>
<interval begin="4650" end="5710">
<flow id="P-NB-3" type="probe" from="P66" number="10" departlane="best"
departPos="random-free" departSpeed="8.33" speedDev="0.1" color="1,0,0"
route="Parc-NB"/>
</interval>
<interval begin="5130" end="5190">
<flow id="U-SB-3" type="probe" from="S5" number="10" departlane="best"
departPos="random-free" departSpeed="8.33" speedDev="0.1" color="1,0,0"
route="Urbain-SB"/>
</interval>
<interval begin="5310" end="5370">
<flow id="L-NB-3" type="probe" from="S38" number="10" departlane="best"
departPos="random-free" departSpeed="8.33" speedDev="0.1" color="1,0,0"
route="Laurent-NB"/>

</interval>
</flows>
</routes>

Introducing travel time of the links in state 2 as an example

<additional>
<!– The real values what we want to convergate to. –>
<!– Parc - SB –>
<vehicle id="P-SB-1-avg">
<lane id="P43" TT="243"/>
<!– Sum of TT on lane id="P43" and lane id="P43.370" –>
<lane id="P45" TT="22"/>
<lane id="P47" TT="31"/>
<lane id="P49" TT="12"/>
<lane id="P51" TT="34"/>
<lane id="P53" TT="19"/>
<lane id="P55" TT="17"/>
<lane id="P57" TT="17"/>
<lane id="P59" TT="61"/>
<lane id="P61" TT="38"/>
<lane id="P63" TT="93"/>
</vehicle>
<vehicle id="P-SB-2-avg">
<lane id="P43" TT="133"/>
<!– Sum of TT on lane id="P43" and lane id="P43.370" –>
<lane id="P45" TT="54"/>
<lane id="P47" TT="44"/>
<lane id="P49" TT="15"/>
<lane id="P51" TT="23"/>
<lane id="P53" TT="17"/>
<lane id="P55" TT="21"/>
<lane id="P57" TT="34"/>
<lane id="P59" TT="60"/>
<lane id="P61" TT="34"/>
<lane id="P63" TT="88"/>

</vehicle>
< !– Parc - NB –>
<vehicle id="P-NB-1-avg">
<lane id="P64" TT="68"/>
< !– Sum of TT on lane id="P64" and lane id="P64.50" –>
<lane id="P62" TT="56"/>
<lane id="P60" TT="26"/>
<lane id="P58" TT="50"/>
<lane id="P56" TT="18"/>
<lane id="P54" TT="35"/>
<lane id="P52" TT="13"/>
<lane id="P50" TT="47"/>
< !– Sum of TT on lane id="P50" and lane id="P50.106" –>
<lane id="P48" TT="55"/>
< !– Sum of TT on lane id="P48" and lane id="P48.306" –>
<lane id="P46" TT="21"/>
< !– Sum of TT on lane id="P46" and lane id="P46.190" –>
<lane id="P44" TT="34"/>
<lane id="P42" TT="25"/>
< !– Sum of TT on lane id="P42" and lane id="P42.190" –>
<lane id="P39" TT="55"/>
</vehicle>
<vehicle id="P-NB-2-avg">
<lane id="P64" TT="28"/>
< !– Sum of TT on lane id="P64" and lane id="P64.50" –>
<lane id="P62" TT="109"/>
<lane id="P60" TT="26"/>
<lane id="P58" TT="49"/>
<lane id="P56" TT="17"/>
<lane id="P54" TT="41"/>
<lane id="P52" TT="17"/>
<lane id="P50" TT="46"/>
< !– Sum of TT on lane id="P50" and lane id="P50.106" –>
<lane id="P48" TT="52"/>
< !– Sum of TT on lane id="P48" and lane id="P48.306" –>
<lane id="P46" TT="23"/>

< !– Sum of TT on lane id="P46" and lane id="P46.190" –>
<lane id="P44" TT="73"/>
<lane id="P42" TT="28"/>
< !– Sum of TT on lane id="P42" and lane id="P42.190" –>
<lane id="P39" TT="52"/>
</vehicle>
<vehicle id="P-NB-3-avg">
<lane id="P64" TT="97"/>
< !– Sum of TT on lane id="P64" and lane id="P64.50" –>
<lane id="P62" TT="121"/>
<lane id="P60" TT="25"/>
<lane id="P58" TT="15"/>
<lane id="P56" TT="32"/>
<lane id="P54" TT="20"/>
<lane id="P52" TT="45"/>
<lane id="P50" TT="58"/>
< !– Sum of TT on lane id="P50" and lane id="P50.106" –>
<lane id="P48" TT="54"/>
< !– Sum of TT on lane id="P48" and lane id="P48.306" –>
<lane id="P46" TT="19"/>
< !– Sum of TT on lane id="P46" and lane id="P46.190" –>
<lane id="P44" TT="31"/>
<lane id="P42" TT="43"/>
< !– Sum of TT on lane id="P42" and lane id="P42.190" –>
<lane id="P39" TT="54"/>
</vehicle>
< !– St. Urbain - SB –>
<vehicle id="U-SB-1-avg">
<lane id="S7" TT="43"/>
<lane id="S9" TT="59"/>
<lane id="S11" TT="41"/>
<lane id="S13" TT="77"/>
< !– Sum of TT on lane id="S13" and lane id="S13.605" –>
<lane id="S15" TT="32"/>
<lane id="S17" TT="34"/>
<lane id="S19" TT="22"/>

<lane id="S21" TT="58"/>
<lane id="S23" TT="23"/>
<lane id="S25" TT="35"/>
</vehicle>
<vehicle id="U-SB-2-avg">
<lane id="S7" TT="46"/>
<lane id="S9" TT="32"/>
<lane id="S11" TT="28"/>
<lane id="S13" TT="82"/>
< !– Sum of TT on lane id="S13" and lane id="S13.605" –>
<lane id="S15" TT="91"/>
<lane id="S17" TT="43"/>
<lane id="S19" TT="68"/>
<lane id="S21" TT="28"/>
<lane id="S23" TT="24"/>
<lane id="S25" TT="73"/>
</vehicle>
<vehicle id="U-SB-3-avg">
<lane id="S7" TT="37"/>
<lane id="S9" TT="58"/>
<lane id="S11" TT="39"/>
<lane id="S13" TT="76"/>
< !– Sum of TT on lane id="S13" and lane id="S13.605" –>
<lane id="S15" TT="26"/>
<lane id="S17" TT="47"/>
<lane id="S19" TT="15"/>
<lane id="S21" TT="57"/>
<lane id="S23" TT="22"/>
<lane id="S25" TT="30"/>
</vehicle>
< !– St. Laurent - NB –>
<vehicle id="L-NB-1-avg">
<lane id="S32" TT="48"/>
<lane id="S30" TT="17"/>
<lane id="S28" TT="12"/>
<lane id="S26" TT="13"/>

<lane id="S24" TT="33"/>
<lane id="S22" TT="25"/>
<lane id="S20" TT="47"/>
<lane id="S18" TT="49"/>
<lane id="S16" TT="27"/>
<lane id="S14" TT="16"/>
<lane id="S12" TT="14"/>
<lane id="S10" TT="14"/>
<!–
Sum of TT on lane id="S10" and lane id="S8" and lane id="S8.19"
–>
</vehicle>
<vehicle id="L-NB-2-avg">
<lane id="S32" TT="48"/>
<lane id="S30" TT="19"/>
<lane id="S28" TT="11"/>
<lane id="S26" TT="12"/>
<lane id="S24" TT="31"/>
<lane id="S22" TT="19"/>
<lane id="S20" TT="41"/>
<lane id="S18" TT="6"/>
<lane id="S16" TT="24"/>
<lane id="S14" TT="18"/>
<lane id="S12" TT="26"/>
<lane id="S10" TT="26"/>
<!–
Sum of TT on lane id="S10" and lane id="S8" and lane id="S8.19"
–>
</vehicle>
<vehicle id="L-NB-3-avg">
<lane id="S32" TT="48"/>
<lane id="S30" TT="80"/>
<lane id="S28" TT="19"/>
<lane id="S26" TT="23"/>
<lane id="S24" TT="40"/>
<lane id="S22" TT="23"/>

<lane id="S20" TT="96"/>
<lane id="S18" TT="10"/>
<lane id="S16" TT="28"/>
<lane id="S14" TT="61"/>
<lane id="S12" TT="16"/>
<lane id="S10" TT="16"/>
< !–
Sum of TT on lane id="S10" and lane id="S8" and lane id="S8.19"
–>
</vehicle>
</additional>

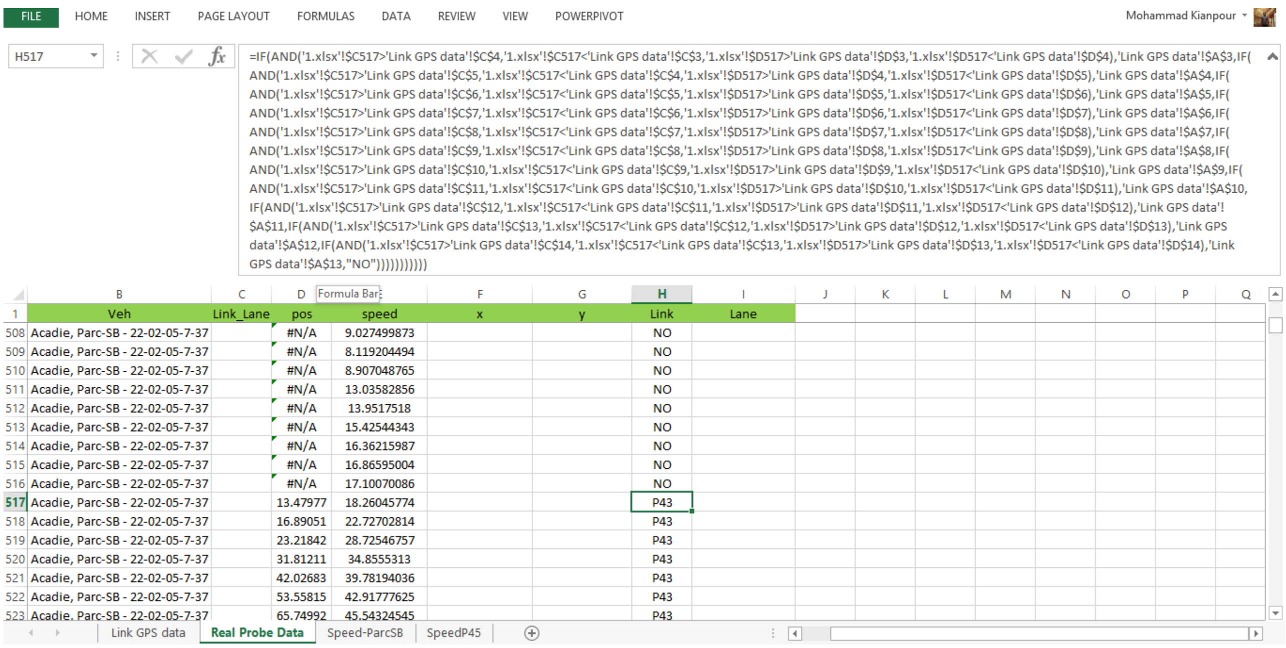
APPENDIX J - Latitude and longitude of the links

	Acadie-Parc SB				
Link	Intersection	Lat	Long	Start	End
P43	Beaumont/Parc	45.52921	-73.6179	113	114
P45	Beaubien Ouest/Parc	45.52697	-73.6129	114	115
P47	Van-Horn / Parc	45.52577	-73.6101	115	116
P49	Bernard / Parc	45.52367	-73.6056	116	117
P51	Bernard & Saint Viateur	45.52297	-73.6039	117	118
P53	Saint Viateur/Parc	45.52219	-73.6022	118	119
P55	Saint Viateur & Fairmont	45.52128	-73.6	119	120
P57	Fairmont/Parc	45.52047	-73.5976	120	121
P59	Laurier/Parc	45.51964	-73.5958	121	122
P61	Saint Joseph/Parc	45.519	-73.5944	122	123
P63	Villeneuve/Parc	45.51815	-73.5923	123	124
	Mont-Royal/Parc	45.51684	-73.5894		
	Acadie-Parc NB				
P64	Mont-Royal/Parc	45.51684	-73.5894	124	123
P62	Villeneuve/Parc	45.51815	-73.5923	123	122
P60	Saint Joseph/Parc	45.519	-73.5944	122	121
P58	Laurier/Parc	45.51964	-73.5958	121	120
P56	Fairmont/Parc	45.52047	-73.5976	120	119
P54	Saint Viateur & Fairmont	45.52128	-73.6	119	118
P52	Saint Viateur/Parc	45.52219	-73.6022	118	117
P50	Bernard & Saint Viateur	45.52297	-73.6039	117	116
P48	Bernard / Parc	45.52367	-73.6056	116	115
P46	Van-Horn / Parc	45.52577	-73.6101	115	114
P44	Beaubien Ouest/Parc	45.52697	-73.6129	114	113
P42	Beaumont/Parc	45.52921	-73.6179	113	112
P39	Jean-Talon/Parc	45.53019	-73.6223	112	111
	Hatchson/Jean-Talon	45.52965	-73.6227		

	St. Urbain SB				
S7	Jean-Talon/Clark	45.5341	-73.6189	204	205
S9	Mozart Ouest/Clark	45.53275	-73.6159	205	206
S11	Saint Zotiqe/Clark	45.5314	-73.6129	206	207
S13	Beaubien Ouest/Clark	45.5301	-73.6099	207	208
S15	Bernard /Urbain	45.52618	-73.6033	208	209
S17	Saint Viateur/Urbain	45.52467	-73.6	209	210
S19	Fairmont/Urbain	45.52271	-73.5956	210	211
S21	Laurier/Urbain	45.5219	-73.5939	211	212
S23	Saint Joseph/Urbain	45.52119	-73.5924	212	213
S25	Villeneuve/Urbain	45.52029	-73.5904	213	214
	Mont-Royal/Urbain	45.51894	-73.5875		
	St. Laurent NB				
S32	Mont-Royal/StLaurent	45.52013	-73.5864	317	316
S30	Villeneuve/StLaurent	45.52146	-73.5894	316	315
S28	Saint Joseph/StLaurent	45.52235	-73.5913	315	314
S26	Laurier/StLaurent	45.52305	-73.5929	314	313
S24	Fairmont/StLaurent	45.5239	-73.5947	313	312
S22	Saint Viateur/StLaurent	45.52581	-73.5989	312	311
S20	Bernard /StLaurent	45.52733	-73.6022	311	310
S18	Beaubien Est/StLaurent	45.53021	-73.6085	310	309
S16	Beaubien Ouest/StLaurent	45.53063	-73.6094	309	308
S14	Saint Zotiqe/StLaurent	45.5319	-73.6123	308	307
S12	Dante/StLaurent	45.5327	-73.614	307	306
S10	Mozart Ouest/StLaurent	45.53337	-73.6154	306	305
S8	Mozart Est/StLaurent	45.5335	-73.6157	305	304
	Jean-Talon/StLaurent	45.53472	-73.6183		

Figure J.1 Latitude and longitude of the links in the study area defined for using in validation of the results of the proposed method

APPENDIX K - Implemented formulas using Excel for calculations




```
=IF(AND('1.xlsx'!$C517>'Link GPS data'!$C$4,'1.xlsx'!$C517<'Link GPS data'!$C$3,'1.xlsx'!$D517>'Link GPS data'!$D$3,'1.xlsx'!$D517<'Link GPS data'!$D$4),'Link GPS data'!$A$3,IF(
AND('1.xlsx'!$C517>'Link GPS data'!$C$5,'1.xlsx'!$C517<'Link GPS data'!$C$4,'1.xlsx'!$D517>'Link GPS data'!$D$4,'1.xlsx'!$D517<'Link GPS data'!$D$5),'Link GPS data'!$A$4,IF(
AND('1.xlsx'!$C517>'Link GPS data'!$C$6,'1.xlsx'!$C517<'Link GPS data'!$C$5,'1.xlsx'!$D517>'Link GPS data'!$D$5,'1.xlsx'!$D517<'Link GPS data'!$D$6),'Link GPS data'!$A$5,IF(
AND('1.xlsx'!$C517>'Link GPS data'!$C$7,'1.xlsx'!$C517<'Link GPS data'!$C$6,'1.xlsx'!$D517>'Link GPS data'!$D$6,'1.xlsx'!$D517<'Link GPS data'!$D$7),'Link GPS data'!$A$6,IF(
AND('1.xlsx'!$C517>'Link GPS data'!$C$8,'1.xlsx'!$C517<'Link GPS data'!$C$7,'1.xlsx'!$D517>'Link GPS data'!$D$7,'1.xlsx'!$D517<'Link GPS data'!$D$8),'Link GPS data'!$A$7,IF(
AND('1.xlsx'!$C517>'Link GPS data'!$C$9,'1.xlsx'!$C517<'Link GPS data'!$C$8,'1.xlsx'!$D517>'Link GPS data'!$D$8,'1.xlsx'!$D517<'Link GPS data'!$D$9),'Link GPS data'!$A$8,IF(
AND('1.xlsx'!$C517>'Link GPS data'!$C$10,'1.xlsx'!$C517<'Link GPS data'!$C$9,'1.xlsx'!$D517>'Link GPS data'!$D$9,'1.xlsx'!$D517<'Link GPS data'!$D$10),'Link GPS data'!$A$9,IF(
AND('1.xlsx'!$C517>'Link GPS data'!$C$11,'1.xlsx'!$C517<'Link GPS data'!$C$10,'1.xlsx'!$D517>'Link GPS data'!$D$10,'1.xlsx'!$D517<'Link GPS data'!$D$11),'Link GPS data'!$A$10,
IF(AND('1.xlsx'!$C517>'Link GPS data'!$C$12,'1.xlsx'!$C517<'Link GPS data'!$C$11,'1.xlsx'!$D517>'Link GPS data'!$D$11,'1.xlsx'!$D517<'Link GPS data'!$D$12),'Link GPS data'!
$A$11,IF(AND('1.xlsx'!$C517>'Link GPS data'!$C$13,'1.xlsx'!$C517<'Link GPS data'!$C$12,'1.xlsx'!$D517>'Link GPS data'!$D$12,'1.xlsx'!$D517<'Link GPS data'!$D$13),'Link GPS
data'!$A$12,IF(AND('1.xlsx'!$C517>'Link GPS data'!$C$14,'1.xlsx'!$C517<'Link GPS data'!$C$13,'1.xlsx'!$D517>'Link GPS data'!$D$13,'1.xlsx'!$D517<'Link GPS data'!$D$14),'Link
GPS data'!$A$13,"NO"))))))))))
```

	B	C	D	E	F	G	H	I
1	Veh	Link_Lane	pos	speed	x	y	Link	Lane
508	Acadie, Parc-SB - 22-02-05-7-37		#N/A	9.027499873			NO	
509	Acadie, Parc-SB - 22-02-05-7-37		#N/A	8.119204494			NO	
510	Acadie, Parc-SB - 22-02-05-7-37		#N/A	8.907048765			NO	
511	Acadie, Parc-SB - 22-02-05-7-37		#N/A	13.03582856			NO	
512	Acadie, Parc-SB - 22-02-05-7-37		#N/A	13.9517518			NO	
513	Acadie, Parc-SB - 22-02-05-7-37		#N/A	15.42544343			NO	
514	Acadie, Parc-SB - 22-02-05-7-37		#N/A	16.36215987			NO	
515	Acadie, Parc-SB - 22-02-05-7-37		#N/A	16.86595004			NO	
516	Acadie, Parc-SB - 22-02-05-7-37		#N/A	17.10070086			NO	
517	Acadie, Parc-SB - 22-02-05-7-37		13.47977	18.26045774			P43	
518	Acadie, Parc-SB - 22-02-05-7-37		16.89051	22.72702814			P43	
519	Acadie, Parc-SB - 22-02-05-7-37		23.21842	28.72546757			P43	
520	Acadie, Parc-SB - 22-02-05-7-37		31.81211	34.8555313			P43	
521	Acadie, Parc-SB - 22-02-05-7-37		42.02683	39.78194036			P43	
522	Acadie, Parc-SB - 22-02-05-7-37		53.55815	42.91777625			P43	
523	Acadie, Parc-SB - 22-02-05-7-37		65.74992	45.54324545			P43	

Figure K.1 Defining the related link by the reported location of real probe vehicle