CZECH TECHNICAL UNIVERSITY IN PRAGUE

FACULTY OF TRANSPORTATION SCIENCES

Bc. Petr Jandík

Prediction of Dynamic Traffic Conditions on Urban Roads Using GPS Traces

Master's thesis

2015

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
Fakulta dopravní
děkan
Konviktská 20, 110 00 Praha 1

K612...Ústav dopravních systémů

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení studenta (včetně titulů):

Bc. Petr Jandík

Kód studijního programu a studijní obor studenta:

N 3710 – DS – Dopravní systémy a technika

Název tématu (česky):	**Predikce dopravních charakteristik v intravilánu za pomoci dat z GPS jednotek**
Název tématu (anglicky):	Prediction of Dynamic Traffic Conditions on Urban Roads Using GPS Traces

Zásady pro vypracování

Při zpracování diplomové práce se řiďte osnovou uvedenou v následujících bodech:

- popis a analýza charakteristik dopravního proudu
- možnosti získávání dat o aktuálních dopravních charakteristikách na pozemních komunikacích v intravilánu
- vývoj dopravního modelu, který bude využívat aktuální data z GPS jednotek pro zisk dopravních charakteristik
- odhad cestovních dob na základě zjištěných dat
- možnosti využití takto získaných dat a jejich přínos
- srovnání metody využívající data z mobilních GPS jednotek a metody využívající data ze statických detektorů

Rozsah grafických prací: stanoví vedoucí diplomové práce

Rozsah průvodní zprávy: minimálně 55 stran textu (včetně obrázků, grafů a tabulek, které jsou součástí průvodní zprávy)

Seznam odborné literatury: stanoví vedoucí diplomové práce

Vedoucí diplomové práce: **doc. Ing. Jiří Čarský, Ph.D.**

Datum zadání diplomové práce: **26. června 2014**
(datum prvního zadání této práce, které musí být nejpozději 10 měsíců před datem prvního předpokládaného odevzdání této práce vyplývajícího ze standardní doby studia)

Datum odevzdání diplomové práce: **31. května 2015**
a) datum prvního předpokládaného odevzdání práce vyplývající ze standardní doby studia a z doporučeného časového plánu studia
b) v případě odkladu odevzdání práce následující datum odevzdání práce vyplývající z doporučeného časového plánu studia

L. S.

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
Fakulta dopravní
3

prof. Ing. Pavel Přibyl, CSc.
vedoucí
Ústavu dopravních systémů

prof. Dr. Ing. Miroslav Svítek
děkan fakulty

Potvrzuji převzetí zadání diplomové práce.

Bc. Petr Jandík
jméno a podpis studenta

V Praze dne .... 26. června 2014

CZECH TECHNICAL UNIVERSITY IN PRAGUE
Faculty of Transportation Sciences
Dean's office
Konviktská 20, 110 00 Prague 1, Czech Republic

K612 .... Department of Transportation Systems

MASTER'S THESIS ASSIGNMENT

(PROJECT, WORK OF ART)

Student's name and surname (including degrees):
Bc. Petr Jandík

Code of study programme code and study field of the student:
N 3710 – DS – Transportation Systems and Technology

Theme title (in Czech): **Predikce dopravních charakteristik v intravilánu za pomoci dat z GPS jednotek**

Theme title (in English): Prediction of Dynamic Traffic Conditions on Urban Roads Using GPS Traces

Guides for elaboration

During the elaboration of the master's thesis follow the outline below:

- description and analysis of traffic conditions
- possibilities of getting data of live traffic conditions on urban roads
- development of a traffic model using live data from GPS traces for a prediction of traffic conditions
- prediction of travel time on base of the predicted characteristics
- possibilities of usage of these data and their benefit
- comparision of approaches using live data from GPS traces and using data from embedded sensors

Graphical work range: assessed by the master's thesis supervisor

Accompanying report length: at the minimum of 55 pages of text
(including figures, graphs, and tables, which are a part
of the accompanying report
Bibliography: assessed by the master's thesis supervisor

Master's thesis supervisor: **Doc. Ing. Jiří Čarský, Ph.D.**

Date of master's thesis assignment: **June 26, 2014**
(date of the first assignment of this work, that has be minimum of 10 months before the deadline of the theses submission based on the standard duration of the study)

Date of master's thesis submission: **May 31, 2015**

a) date of first anticipated submission of the thesis based on the standard study duration and the recommended study time schedule
b) in case of postponing the submission of the thesis, next submission date results from the recommended time schedule

L. S.

prof. Ing. Pavel Přibyl, CSc.
head of the Department
of Transportation Systems

prof. Dr. Ing. Miroslav Svítek
dean of the faculty

I confirm assumption of master's thesis assignment.

Bc. Petr Jandík
Student's name and signature

Prague .... June 26, 2014

Prohlášení

Předkládám tímto k posouzení a obhajobě diplomovou práci, zpracovanou na závěr studia na ČVUT v Praze Fakultě dopravní.

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o etické přípravě vysokoškolských závěrečných prací.

Nemám závažný důvod proti užití tohoto školního díla ve smyslu § 60 Zákona č.121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Montrealu dne 30. května 2015 ....

podpis

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

Fakulta dopravní

Predikce dopravních charakteristik v intravilánu za pomoci dat z GPS jednotek

diplomová práce

květen 2015

Bc. Petr Jandík

ABSTRAKT

Současné rozšíření GPS technologií do každodenního života může být využito mnoha způsoby. Jedním z nich je i monitorování aktuálních dopravních charakteristik. Vozidla s GPS jednotkami na palubě totiž mohou být chápána jako měřicí plovoucí vozidla. Tato práce se zabývá analýzou takových dat z případové studie na sběrných komunikacích v Montrealu. Dále je vypracován dopravní model oblasti a je uvedena technika jeho zkalibrování pomocí změřených GPS dat.

KLÍČOVÁ SLOVA

Dopravní modelování, mikrosimulace, SUMO, GPS, simulované žíhání, doba jízdy.

CZECH TECHNICAL UNIVERSITY IN PRAGUE

Faculty of Transportation Sciences

Prediction of Dynamic Traffic Conditions on Urban Roads Using GPS Traces

Master's thesis

May 2015

Bc. Petr Jandík

ABSTRACT

The contemporary penetration of GPS technology in everyday life might be used in many different ways. One of them is a monitoring of live traffic conditions. Vehicles equipped with GPS onboard units can be considered as float vehicles. This thesis analyzes those data from a case study in Montreal. The traffic model of the area is created and a method how to calibrate the model by GPS data is provided.

KEY WORDS

Traffic modelling, microsimulation, SUMO, GPS, simulated annealing, travel time.

CONTENTS

Electronical enclosures:

Enclosure 1: Data from Traffic Analysis

1.1: GPS Traces in morning peak hour

1.2: Aggregated statistics in morning peak hour

1.3: Counted Data

1.4: Plots of Speed during Trajectories

Enclosure 2: SUMO Model

2.1: SUMO Model

2.2: Program for Model Calibration (developed by Anas Balboul, Polytechnique Montreal)

1.Used Abbreviations

DUE	dynamic user equilibrium
FC	floating car
FCD	floating car data
FHWA	Federal Highway Administration
HCM2000	Highway Capacity Manual 2000
HCM2010	Highway Capacity Manual 2010
k	density
k_j	jam density
NB	northbound
OSM	OpenStreetMap
PHF	peak hour factor
q	traffic flow
SB	southbound
TT	travel time
u_f	free flow speed
u_i	spot speed of i-th vehicle
u_s	space-mean speed
u_t	time-mean speed
vph	vehicles per hour

2. Introduction

There has been a phenomenal increase in congestion on urban roads of basically all big cities. Thus having the ability to predict short-term changes in link-by-link traffic conditions (i.e. speed and flow) has various key applications, from demand management to dynamic route guidance for avoiding congested roads. Conventionally, embedded sensors like loop detectors or camera technology have been used to an idea of the traffic conditions. The high cost and maintenance requirements however hindered their widespread usage. Moreover, the end users (drivers) do not have access to such data streams. The penetration of GPS technology in day-to-day routing now gives us a new way to assess traffic conditions. They are like probes that have the ability to give us the system pulse.

The advantage of similar data is the fact that they are very easily accessible. They already exist here in a big volume and they can be collected every day. For example taxi fleets are omnipresent on roads in dense urban and are most likely equipped with GPS. Another source of data might be any company monitoring its fleet cars by GPS. They all thus form the perfect probes that can give us a measurement on the pulse of the traffic.

The main idea of this approach is to develop a simulation environment which use a historical counted data as the prior of the phenomena (i.e. traffic flow) on one hand, while it is updated and calibrated by live GPS traces on the other hand. This simulation environment, which would act as a parallel reality, would be able to give us instant updates on the link-by-link speeds, flow, and densities.

The resulting online short-term prediction framework would have various applications with objective to decrease travel times in general. This vital information can then be fed to dynamic congestion management strategies or back to users for improved route guidance. It can be used by the traffic department of a city in pre-emptive demand management (e.g. adjusting signal timing and dynamic display signs) at both system level and neighbourhood level. While the detailed long-term patterns of road conditions from this simulation environment could be greatly useful for sustainability analysis of more long-term policies like infrastructure investments.

2.1. Traffic characteristics

If we want to talk about a description and prediction of traffic characteristics, it is suitable to provide at least their brief listing with definitions and mutual relations. The three basic characteristics in traffic flow theory are flow, speed, and density.

2.1.1. Traffic flow

Traffic flow describes a traffic demand on a particular road.

Traffic flow can be also found in some literature named as *flow*, *flow rate* or *volume*. Nevertheless, all these terms can be used interchangeably.

It is defined as a number of vehicles passing a point in a given period of time. It is usually expressed in units of vehicles per hour (vph). Other possible units are for instance vehicles per hour per lane (vphpl), passenger car units per hour (pcu/hr), or passenger car units per hour per lane (pcphpl or in easier readable form pc/h/ln).

The equation for counting traffic flow is simple:

$$q = \frac{n}{t}$$

where q is traffic flow (vph), n is number of vehicles passing a spot on the road in a given interval t.

The special value of traffic flow is capacity c which defines the maximum hourly rate under prevailing roadway conditions.

Since it is convenient to measure traffic flow in 15-minute intervals, a quantity called *peak hour factor* (PHF) is presented:

$$PHF = \frac{hourly\ traffic\ flow\ (vph)}{peak\ 15 - minute\ flow\ within\ the\ hour\ (vph)}$$

Hence is obvious that PHF can theoretically be in the interval 0.25≤PHF≤1.0. Peak hour factor can be understood as an indicator of flow fluctuations within the hour.

2.1.2. Speed

Based on different approaches how to calculate speed, there are two main interpretations - *time-mean speed* and *space-mean speed*.

Time-mean speed is defined as the average speed of vehicles passing a spot on a road. The standard notation is u_t.

$$u_t = \frac{\sum_{i=1}^{n} u_i}{n}$$

where:

u_t is time-mean speed in km/h,

u_i is spot speed of vehicle *i* measured at a spot along a road, usually by a radar or laser gun (km/h),

n is number of vehicles.

Space-mean speed is defined as the average travel speed of vehicles between two points at the distance *D* apart. It is computed as:

$$u = u_s = \frac{D}{\bar{t}} = \frac{D}{\frac{1}{n}\sum_{i=1}^{n}\frac{D}{u_i}} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{u_i}}$$

where:

u_s is space-mean speed,

D is the distance of two points on the road (km)

$\bar{t}$ is the average travel time (h).

Space-mean speed is more useful in the context of traffic analysis and is determined on the basis of the time necessary for a vehicle to travel some known length of a roadway. For these reasons it is also signified simply as *u*.

Space-mean speed gives more emphasis to high u_i, for this reason $u_s \leq u_t$.

The special value of speed is *free flow speed (FFS or u_f)* which defines the space-mean speed on the particular part of the road which is reached by a unrestricted traffic flow under prevailing roadway conditions. HCM2010 defines FFS as the mean speed of passenger cars operating in flow less than 1 000 pc/hr/ln. FFS is determined by road geometry, cross section, quality of road surface, and all psychological factors making impact on drivers.

2.1.3. Density

Density is defined as the number of vehicles per unit length of roadway at a time instant. The standard notation is k and its unit is vehicles per kilometre (veh/km) or vehicles per kilometre per lane (veh/km/ln). It is expressed as

$$k = \frac{n}{D}$$

where n is the number of vehicles occupying length D of roadway at some specified time.

It is rather difficult to measure density directly, unless there is an opportunity to use an aerial or satellite photography. Therefore it is more often estimated indirectly by measuring the inflow and outflow of vehicles at a road section over time, but the initial state must be known in that case.

The special case of density is so called jam density (k_j), which is the maximum possible density on the roadway when the speed of the flow is nearly zero.

2.1.4. Relationships among traffic characteristics

If there is a requirement of analysis of traffic conditions, the macroscopic approach is used.

The fundamental equation describing average conditions on a given link for a specific time period is:

$$q = u * k.$$

The equation assumes that the flow is uninterrupted and stable, i.e., all travelling at about the same speed.

The graphical interpretation is shown in Figure 1.

The fundamental diagram in the presented form assumes a linear speed-density model. That assumption is represented by Greenshield's traffic stream model. Nevertheless, it is not the only traffic stream model which can be used. There is several others - Underwood's model, northwestern model, Pipes-Munjal model, Drew's model, and others. Each of them can be used advantageously in certain situations. Even their combinations creating multi-regime models might be applied.

The benefit of using a linear representation of the speed-density relationship is that it provides a basic insight into the relationships among traffic flow, speed, and density interactions without clouding these insights by the additional complexity that a nonlinear speed-density relationship introduces. However, it is important to note that field studies have shown that the speed-density relationship $u = f(k)$ tends to be nonlinear at low densities and high densities. In fact, the overall speed-density relationship is better represented by three relationships: (1) a nonlinear relationship at low densities that has speed slowly declining from free flow value u_f, (2) a linear relationship over the large medium-density region, and (3) a nonlinear relationship near jam density k_j as the speed asymptotically approaches zero with increasing density. (Mannering, & Washburn).

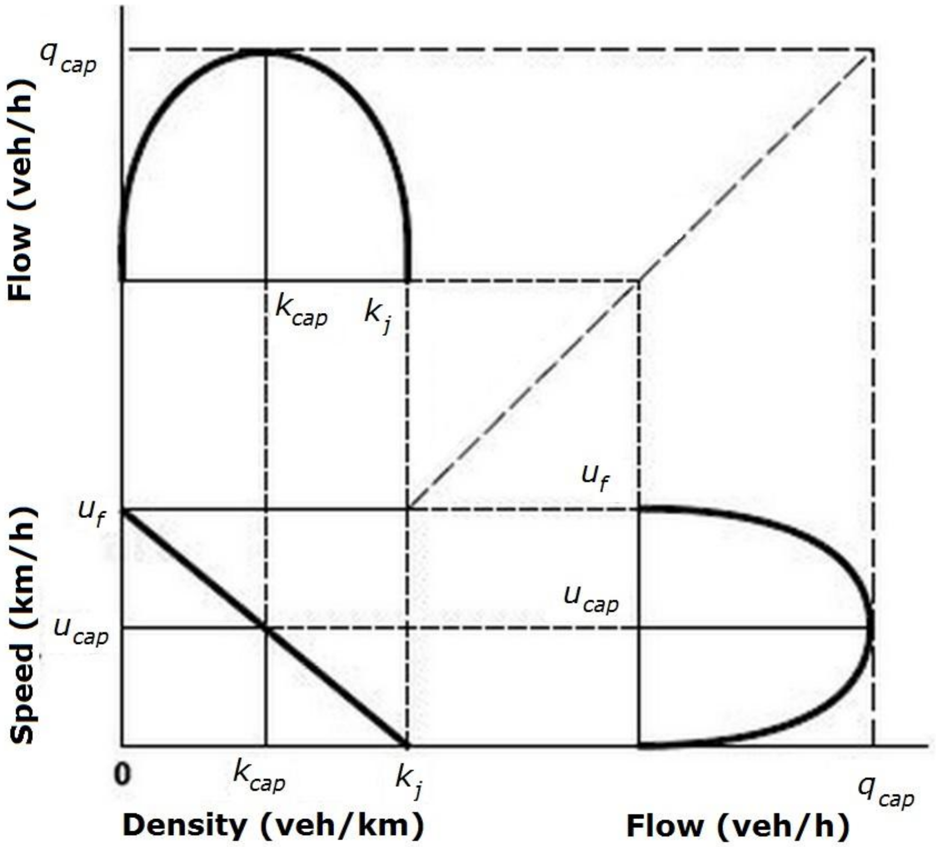


Figure 1: Flow-density, speed-density, and speed-flow relationships(*Mannering, & Washburn)*

2.2. Travel time

There are the basic traffic characteristics presented in the previous chapter 2.1. It is important to know that they are not the only ones what can be measured or computed in the traffic flow assignments.

One of the characteristic which is perceived by the end users (i.e. drivers) and affects their decisions about their trips is travel time (TT). TT is defined as a time taken by a vehicle to traverse a given section of a highway with length *D*.

$$TT = \frac{D}{average\ speed}$$

A precious estimation of travel time might decrease cost impacts by a possibility to avoid links with congestion and improve a quality of carriage companies by delivering their shipments in the required time period.

The value of TT at a particular trajectory is not affected only by a road geometry, but also by factors specific for the particular instant of the trip. These factors are for instance actual weather and state of the carriageway, and surely also actual level of service based on traffic conditions. It is said in general that travel time is affected by people, vehicles, and facilities.

There is several options how to measure TT. The first group includes ways which record passing of a car through measuring points. While methods from the second group use measuring probes moving in traffic flow which record their progress.

Examples of the first group might be license plate method in any possible variant. It finds difference in arrival time for vehicles arriving in points A and B. The assumption for getting a relevant result is that the clocks at both points are synchronized. It can be done manually by observers in its basic form. The more common way is automatic matching of some ID of cars. As an ID can be considered plates read by sensors of automatic vehicle identification systems monitoring the flow or tags used in electronic toll collection systems (onboard units). A computation of TT based on data from embedded detectors is another example belonging among these methods.

The second group using a moving observer is more interesting in the terms of this thesis. These methods are generally based on using floating cars as detectors. The variants are described in chapter 2.3.

Based on traffic flow theory and level-of-service concept, as traffic flow increases, speed decreases, and therefore travel time increases. There have been developed several link performance functions which present a mathematical relationship between route travel time and route traffic flow.

FHWA uses its BPR function:

$$t = t_0 \left[1 + \propto \left(\frac{v}{c} \right)^{\beta} \right]$$

where:

t is travel time of a link,
t_0 os a free-flow travel time of a link,
v is link traffic flow (volume) of a link,
c is link capacity,
α, β are calibration constants (typically $\alpha = 0.15$, $\beta = 4$).

Second function is Davidson's formula:

$$t = t_0 \left[1 + J \left(\frac{v}{c - v} \right) \right]$$

where J is a calibration parameter of value 0.4 to 0.6 for urban arterials surveyed in my case study.

HCM2000 provides Akcelik Delay Function:

$$t = t_0 + D_0 + 0.25T \left[(X - 1) + \sqrt{(X - 1)^2 + \frac{16JXL^2}{T^2}} \right]$$

where:

D_0 is zero-flow control delay at signalized intersections (h),
T is expected duration of demand (typically 1 hour) (h),
X is link demand to capacity ratio (v/c),
J is calibration parameter (-)
L is link length (mi).

Other link performance functions are logit-based function computing total delay as sum of sum of link speed delay and intersections delay, or conical function.

2.3. Methods of collecting data

I already mentioned that there are different ways of measuring travel time in the previous chapter 2.2. As collecting relevant traffic data is absolutely crucial for any kind of traffic analysis, traffic control systems, and transportation demand management techniques, I present a summary of the most common methods how to do that.

2.3.1. Traffic detectors

Equipment used to collect data is also known as sensors. In traffic engineering industry, it is more often referred to as detectors. Each type of detector has its own measuring capability, method of operation, advantages and disadvantages.

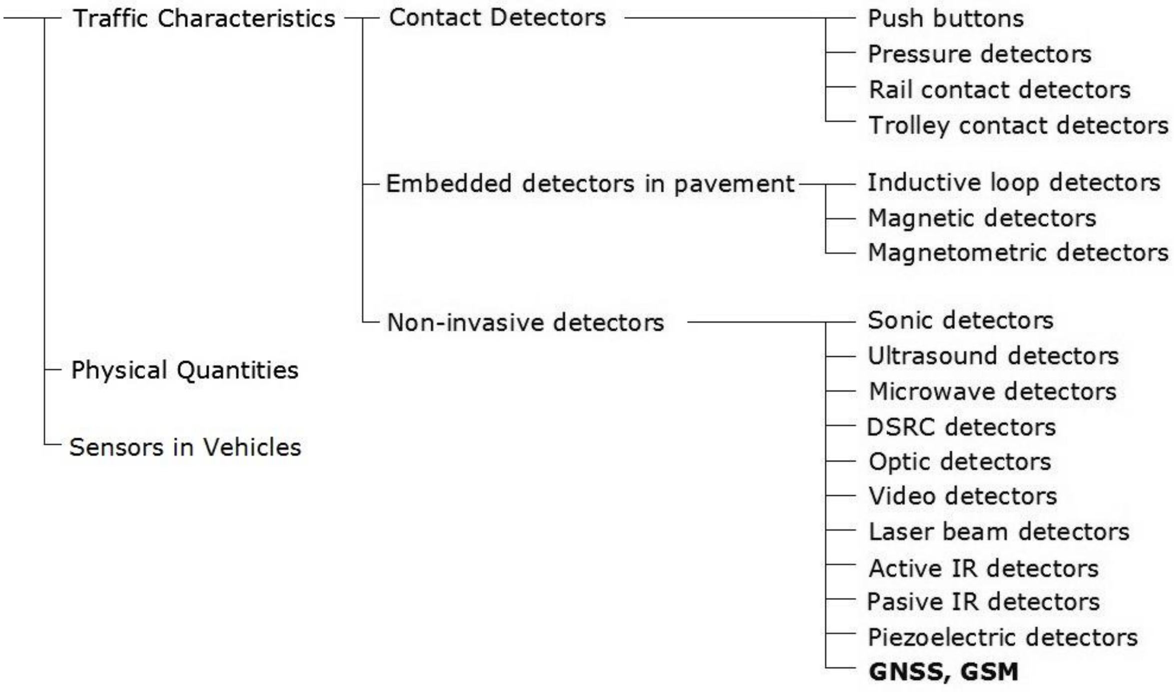


Figure 2: Overview of sensors in transportation engineering

In terms of this thesis detectors for measuring traffic characteristics are important.

Inductive loop detectors are an example of classical embedded sensors. They are broadly used. They are popular by road maintenance services for their ease in installation, capabilities in measuring most traffic parameters by suitable combination of number of loops, flexibility in terms of size and shape and relatively low maintenance costs.

Probably the second most often used detectors of traffic characteristics, which are gaining popularity worldwide, are video image processing systems. They make use of pixel values of video images to detect vehicles. They have got advantage of a possibility to create many types of detection zones, the fact that many processing algorithm cam be implemented, including vehicle tracking and queue measurement. Another benefits of them are that the user can see video images and that digging out the pavement surface is not needed. On the other hand their drawback is they are rather expensive and they might be affected by weather, shadows, occlusion and view angle.

There is "GNSS, GSM" highlighted in Figure 2. This is the group of detectors what I use for my case study in this thesis.

There are various traffic applications of using a knowledge of a position of the car. It can be used for monitoring and controlling of vehicles, for a dedicated traffic like public transportation vehicles or vehicles hauling toxic freight, for safety and warning systems like eCall, for navigation systems and last but not least for collecting data by floating cars. Floating cars might be represented by special measuring vehicles, by drivers - volunteers, or by contracting fleet of cars. Company fleet cars are in most cases equipped with GPS unit for its monitoring and keeping records, so it is only a question of an agreement for using these data for use of traffic engineering applications.

2.3.2. Methods of measuring travel time

The calibration of my case study is based on computation of travel times (TT) on links, therefore I describe ways of determining TT.

The very simple method of measuring TT can be a test vehicle method when drivers are asked to record their TT and path at a specific duration of study, as they drive in the real traffic stream. Sampling of vehicles and drivers must be chosen carefully.

A slightly more sophisticated version of it is a moving-observer method when another person in addition to the driver is needed to make vehicle counts (number of vehicles overtaking the test vehicle and number of vehicles passed by the test vehicle). It is usually done in both directions (AB and BA) and outbound traffic flow can be counted:

$$q_{AB} = \frac{M_{BA}+O_{AB}-P_{AB}}{t_{AB}+t_{BA}},$$

where M_{BA} is number of vehicles seen travelling in the AB direction observed while the test car is returning to A, O_{AB} is number of vehicles overtaking the test vehicle, and P_{AB} is number of vehicles overtaken by the test vehicle. A sum of TT in both directions is in the denominator.

A probe vehicle method is more advanced version of test vehicle method. GPS unit is installed onboard and it continuously record down the position and speed of the test vehicle. This is the method what was used in the case study from chapter 5 of this thesis.

The latest modification of probe vehicle method uses cell phones of drivers instead of GPS units onboard. It is justifiable as the penetration of traffic stream by probes could be nearly 100%, if the access to those data were permitted by telecommunication companies. The drawback is that the location is not as precious as from GPS units. It is usable for predicting travel times for long trips, but the location is not sufficient in dense urban areas and there are difficulties with deviations.

3. Methodology

3.1. Methodology in general

The principal idea of this project is to get knowledge about the actual traffic condition in an area of a city. In the case study I monitor an area of ca. 3.2 km x 0.5 km, it gives ca. 250 ha (600 acres).

Floating cars equipped with GPS unit drive around this area and record their precious position and actual speed every second. Data from morning peak hour are used in the case study.

Then the area is simulated in a microsimulation software - SUMO is used there. Traffic characteristics on each links of simulated floating cars are recorded as the output of the simulation.

After the simulation finishes, there are two matrixes - the first one describes real date measured on the road, while the second one contains the same parameters but for the simulated vehicles. The goal is to minimize differences between these two matrixes. For this purpose the number of vehicles is changed during an iteration process while there is an acceptably small difference. Then the number of vehicles on inflows to the area are taken as the "correct" traffic flows and traffic characteristics can be measured wherever throughout the network for any time.

3.2. Princip of microsimulations

Traffic simulation is the mathematical modelling of transportation systems (system of three arterial routes in the case study of this thesis) through the application of computer software to better help plan, design and operate transportation systems. Simulation of transportation systems started over forty years ago, and is an important area of discipline in traffic engineering and transportation planning today

Simulation in transportation is important because it can study models too complicated for analytical or numerical treatment, can be used for experimental studies, can study detailed relations that might be lost in analytical or numerical treatment and can produce attractive visual demonstrations of present and future scenarios. („Traffic Simulation”, 2001)

Traffic microsimulation is used in this thesis. That model captures the interactions of real world road traffic through a series of complex algorithms describing car following, lane changing, gap acceptance, and spatial collision detection. Microscopic traffic-simulation tools are applied to deal with dynamic and operational problems and to evaluate a range of new proposals and scenarios. The output of microsimulation might be behaviour of particular vehicles as well as the state of traffic condition on the network. It provides a control on all levels.

Simulation tools works with some stochastic probabilities. By running the simulation we can get just one result of one observation which can be interpreted as a statistical experiment. It corresponds to the real situation when the traffic flow never behaves deterministically.

In general, the strengths of simulations are that allow analysis of systems too complicated for standard mathematical methods. They allow a model of unusual situations, to study the system in real time, to run experiments without high costs of investments, to simulate situations when their experiment would not be safe, and the simulation might help to understand of hidden processes by detailed analysis. On the other hand the weaknesses are that there might be the simpler techniques for solving a particular problem, they might be time demanding and expensive, they request a lot of input data, a suitable methods of calibration and validation should be used, and last but not least the author is the creator of the simulated reality and can adjust it to whatever is required. Therefore it is sometimes said that simulations are useful but dangerous.

The principle of creating simulation models is demonstrated in Figure 3.

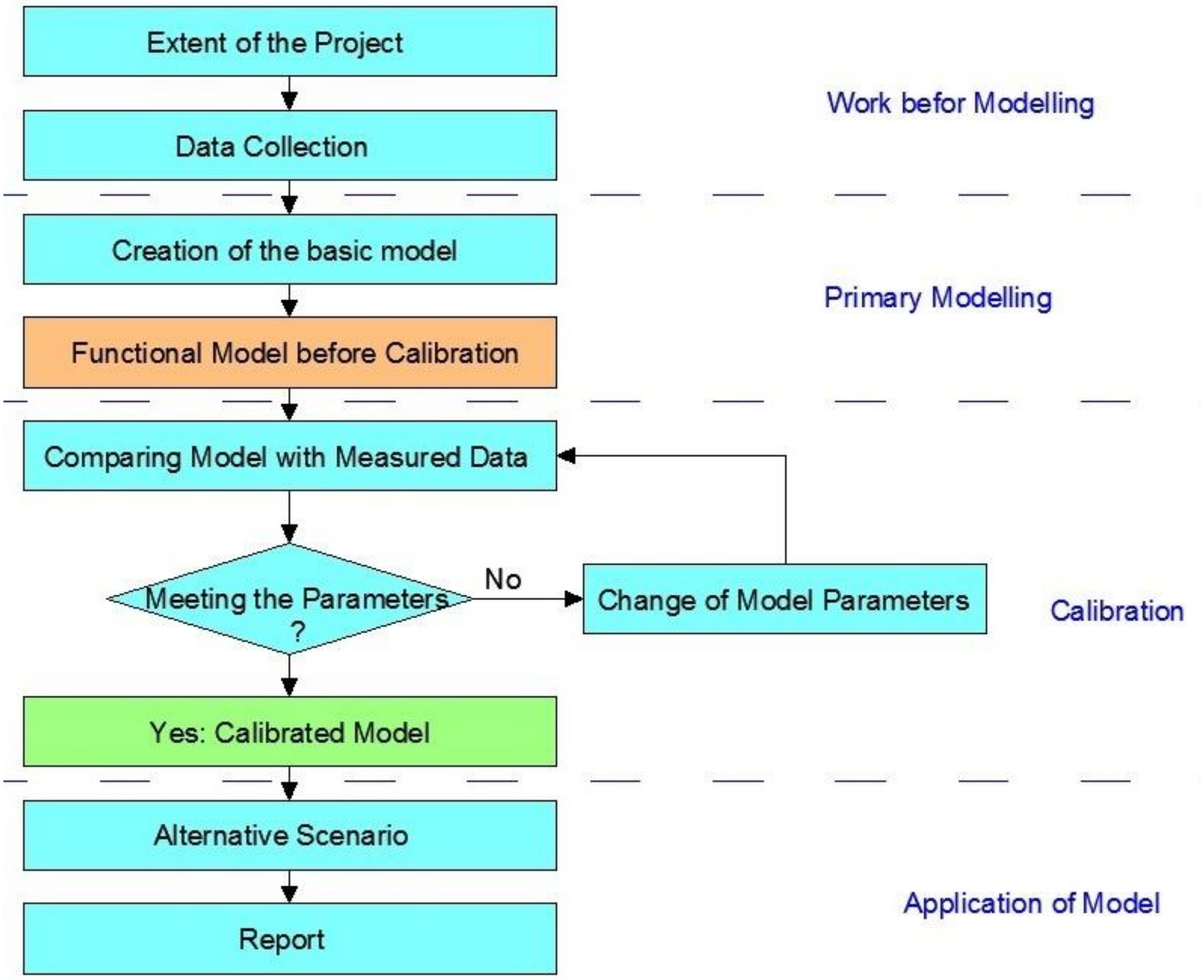


Figure 3: Process of creating simulated model

The core of traffic simulation based on discrete time and continuous space is car-following model which describes driver's behaviour following another vehicle and system of interaction among vehicles. In SUMO, which is used in this project, modified Krauß- model is set as a default setting. Since there has not appeared any reason why to change it, it is kept as a car-following model for the case study. (Krauss, 1997).

3.3. Dynamic Traffic Assignment

During the process of modelling there occurs two types of dynamics.

The first one is time dynamic when traffic conditions vary in time (dynamic route choice) and travel demand very in time and depends on the current level of congestion (dynamic departure time choice).

The second kind of dynamic is link dynamic. In static models the flow/delay-function is used to describe the relationship between speed and flow, but it does not give realistic results under oversaturated traffic condition. A more realistic modelling of traffic when demand exceeds capacity can be done by macroscopic flow simulation, queue modelling, or spillback. Dynamic models are used for short-term traffic infrastructure planning (offline) or mainly for operative use (online) when it can provide traffic information, suggest and evaluate rerouting, or some other usages like estimation of air quality, etc.

Static models are incapable to model peak demand as is compared in Figure 4. In static models all demand is assumed to be served, whereas in a dynamic models route choice will depend on the current level of congestion as the arrival time is its function.

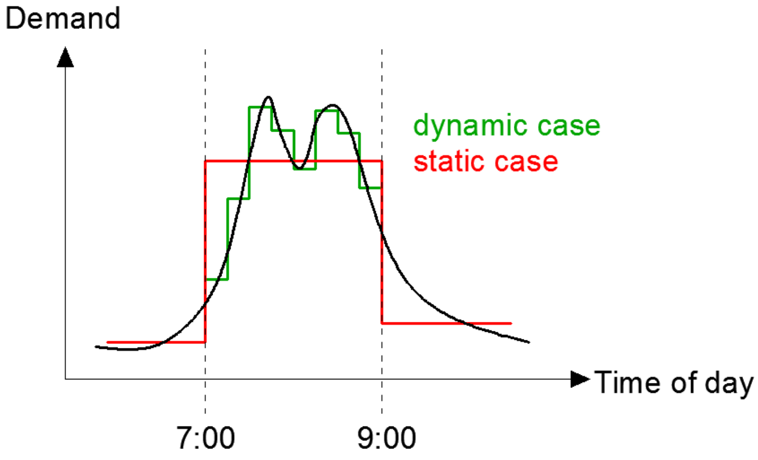


Figure 4: Static vs Dynamic Models

Dynamic traffic assignment (DTA) models have important applications to the rapidly developed advanced traveller information systems and advanced traffic management systems. In general, these DTA models can be classified into two categories: the reactive assignment model and the predictive assignment model. The reactive assignment model assumes that each traveller chooses the shortest route to his destination according to present instantaneous traffic condition. As a result of the time-varying traffic condition, travellers between the same origin-destination (OD) pair departing at the same time may arrive at the destination differently if different routes have been chosen. In contrast, the predictive assignment model considers the impact of future traffic condition on route choice behaviour, i.e. the shortest route is determined based on the actually experienced travel time or cost by a traveller leaving from a particular location at certain time.

Most of the predictive assignment models aim to satisfy the dynamic user equilibrium (DUE) condition which requires that equilibrium, the TT between same OD pair departing at the same time is minimal and equal on all routes.

However, the predictive assignment problem is more difficult than that due to the high computational requirements and vague property of the actual route TT in general networks. These model often adopt longer time intervals, for instance 15 minutes, so they may be more applicable for long-term transportation planning, rather than for instantaneous traffic analysis.

3.4. SUMO software

SUMO (Simulation of Urban MObility) is a free and open traffic simulation suite which is available since 2001. It consists of a microscopic simulator for multimodal road traffic and a host of applications for preparing simulation input data (network import and modification, traffic import, routing) and for working with simualtion outputs. SUMO allows modelling of intermodal traffic systems including road vehicles, public transport and pedestrians. Included with SUMO is a wealth of supporting tools which handle tasks such as route finding, visualization, network import and emission calculation.

SUMO is implemented in C++ and uses only portable libraries. It has been under continuous development of Institute of Transportation Systems, Berlin, Germany. SUMO version 0.22.0 is used here to perform the microsimulation.

It was chosen for this simulation because of its relatively small requirements on memory usage in CPU, the fact that it is fully editable and the user has full control over the model.

3.4.1. Building network in SUMO

It is possible to import network into SUMO from various data sources (Open StreetMap, Visum, Vissim, OpenDrive, MATsim, ArcView, Elmar's GDF, Robocup Simulation League.

In the primary phases of the project I worked with a network model imported from OpenStreetMap. However, I did not find it very convenient for further work. There is a need of making a lot of small changes to get it into a shape according to the reality. There are some commands helping with that, but it is rather difficult to keep it in a state when an inexperienced user can be sure that nothing has been omitted and it is really according to his idea. Another drawback is that since all IDs are taken from larger areas, they have ordinarily no sense in the smaller section of it. It leads to too complicated IDs which are not user friendly and can subsequently lead to some errors.

For these reasons I chose the basic way of building the network. It is defined by own XML-descriptions created by hand. I would highly recommend it for areas of my size (81 nodes and 195 edges in this case). It brings much better control over the model. However, during the process of defining the network I also used the model imported from OSM with an advantage. All nodes in my model has the same coordinates as in the OSM model. By building the network by hand some details are lost in comparison with an imported network. Nevertheless, all crucial information are kept. The level of the preserved details must be set in the beginning. As I focus in my simulation on three north-south arterials roads, minor parallel streets are left out as they do not cause any delay on the observed routes, carry only very small local traffic volumes, which have got their origin or destination on that street in most cases, so they do not serve as an alternative for Av du Parc, St. Urbain and St. Laurent which connect the uptown and downtown.

A SUMO network file describes the traffic-related part of a map, the roads and intersections the simulated vehicles run along or across. At a coarse scale, a SUMO network is a directed graph. Nodes, usually named "junctions" in SUMO-context, represent intersections, and "edges" roads or streets. All edges are unidirectional.

The SUMO network file is named as "my.net.xml" in this work. Although being readable (XML) by human beings, a it is not meant to be edited by hand. SUMO XML description files together with NETCONVERT should be always used.

The XML description files for creating the network are the followings :

- **my.nod.xml**

File containing all definition of nodes, which represent either intersections (named 111-319) or they have a function of source or sink.

Every node is described in a single line which looks like this:

```
<node id="<STRING>" x="<FLOAT>" y="<FLOAT>" type="<TYPE>"/>
```

Each of these attributes has a certain meaning and value range (only the used attributes are listed):

Table 1: Attributes of my.nods.xml

Attribute Name	Value Type	Description
id	id (string)	The name of the node.
x	float	The x-position of the node on the plane in meters.
y	float	The y-position of the node on the plane in meters.
type	"priority", "traffic_light"	Way of control of the junction.

- **my.edg.xml**

Within the edges file, each description of a single edge looks like this:

```
<edge id="<STRING>" from="<NODE_ID>" to="<NODE_ID>" type="<STRING>"
speed="<FLOAT>" length="<FLOAT>" shape="<2D-POSITION> />
```

The origin and the destination nodes are defined using their IDs (from="<NODE_ID>" to="<NODE_ID>"). Each edge is unidirectional and starts at the "from"-node and ends at the "to"-node. For each edge, some further attributes are supplied. Attribute speed is meant at default to be the maximum speed allowed on the edge. In order to getting closer the real situation, it is set to a value evaluated as a free flow speed from the recorded GPS data. It can be understood as a part of a calibration process.

Table 2: Attributes of my.edg.xml

Attribute Name	Value Type	Description
`id`	id (string)	The name of the edge (must be unique). The system of naming the edges is visible from the topology scheme in Enclosure X.
`from`	referenced node id	The name of a node within the nodes-file the edge shall start at.
`to`	referenced node id	The name of a node within the nodes-file the edge shall end at.
`type`	referenced type id	The name of a type within the SUMO edge type file "my.typ.xml".
`length`	float	The length of the edge in meter. The value corresponds with the length how the trajectories were divided into links during the float car measuring.
`shape`	List of positions; each position is encoded in x,y coordinates in meters.	It gives a shape to edges when necessary.

If there is a change of the number of lanes on the edge attribute split is used instead of using two or more consequent edges. It splits the edge at defined position and inserts there a node named <EDGE_ID>.<POSITION>. The parameters of attributes can be set separately for the edge before this node and after it.

The definition of a split uses the following attributes:

Table 3: Attributes of splitted edges

Attribute Name	Value Type	Description
`pos`	float	The position along the edge at which the split shall be done (in m).
`lanes`	list of lane ids (ints)	Information which lanes should exist after the split.
`speed`	float	The speed in m/s after the split position.

- **my.typ.xml**

A separate file with type description is created to ease the definition of edges. Each description of an edge includes information about the number of lanes, the maximum speed allowed on this edge (on edges being analyzed, this value is overwrite in their own definition though in order to giving more realistic results).

Each description of a class of edges looks like this:

```
<type id="<STRING>" priority="<INT>" numLanes="<INT>"
speed="<FLOAT>"/>
```

I use edge-type file for defining the following attributes:

Table 4: Attributes of my.typ.xml

Attribute Name	Value Type	Description
`priority`	int	The priority of the edge. Vehicles on a low-priority edge have to wait until vehicles on a high-priority edge have passed the junction.
`numLanes`	int	The number of lanes of the edge.
`speed`	float	The maximum speed allowed on the edge in m/s.

All these attributes of a type might be of course defined for edges themselves.

- **my.con.xml**

The file describing connections defines how a node's incoming and outgoing edges are connected. Though this might be guessed by NETCONVERT if not given, definitions of connections between edges or lanes may be set up manually, for example to prohibit left-turns at some junctions. The definition is as the following:

```
<connection from="<FROM_EDGE_ID>" to="<T0_EDGE_ID>"
fromLane="<INT_1>" toLane="<INT_2>"/>.
```

or

```
delete from="<FROM_EDGE_ID>" to="<T0_EDGE_ID>"/>
```

If it is more convenient to remove a connection. Lanes are counted from the right (outer) to the left (inner) side of the road beginning with 0. Again the parameters:

Table 5: Attributes of my.con.xml

Attribute Name	Value Type	Description
from	referenced edge id	The name of the edge the vehicles leave.
to	referenced edge id	The name of the edge the vehicles may reach when leaving "from"
fromLane	int	The lane index of the incoming lane (numbers starting with 0).
toLane	int	The lane index of the outgoing lane (numbers starting with 0).

- **my.tls.add.xml**

This file describes traffic lights and signal plans.

The following attributes/elements are used within the tlLogic element:

Table 6: Attributes of my.tls.add.xml

Attribute Name	Value Type	Description
id	id (string)	The ID of the traffic light. It is the same as the ID of the controlled intersection (node) for easier orientation.
type	"static", "actuated"	In the solved area are traffic lights only with fixed phase durations. ("static")
programID	id (string)	The ID of the traffic light program. Makes sense only if there is more plans for one traffic lights, but cannot be left out.
offset	int	The initial time offset of the program.

The following signal colours are used:

Table 7: Phases of traffic signals in my.tls.add.xml

Character	Description
r	'red light' for a signal - vehicles must stop
y	'amber (yellow) light' for a signal - vehicles will start to decelerate if far away from the junction, otherwise they pass
G	'green light' for a signal, no priority - vehicles may pass the junction if no vehicle uses a higher priorised foe stream, otherwise they decelerate for letting it pass
g	'green light' for a signal, priority - vehicles may pass the junction

The entire file "my.tls.add.xml" was written based on signal plans provided by the municipality of the city of Montreal.

Few plans might appear incorrect at the first sight, but it is important to realize that some simplifications especially about pedestrians have been done. In order to keep it still realistic pedestrians are considered, even though they are not simulated. Therefore there are for example traffic lights "117" and "119" in the middle of links on Av de Parc, but they model controlled crosswalks, so the flow is stopped even though it might be meaningless without that knowledge.

There might be also an additional part of the same file defining the relationship between a signal plan and the controlled connections. It is not used though. The definition works in system that each movement from each lane has its own definition. This is the easiest way of definition.

Even though there might be given different signal programs it was not used in this thesis, because it simulates only the morning peak hour when the programs does not change. Another options are for actuated programs which allow dynamic changes in phase length. This is not used as it is not used in the area neither. However, both of these options can be interesting for further development of the model and simulating different scenarios.

All these files above are converted into one file "**my.net.xml**" by NETCONVERT.

NETCONVERT imports digital road networks from different sources and generates road networks that can be used by other tools from the package. It is a command line application. It assumes at least one parameter. Since I use 5 different files for creating the network as an output of NETCONVERT, configuration file "**my.netc.cfg**" summarizing all of them is provided. Therefore NETCONVERCT can be called easily only with one parameter as:

```
netconvert -c netc.cfg
```

The final output file "my.net.xml" contains all these information about the SUMO network:

- every street (edge) as a collection of lanes, including the position, shape and speed limit of every lane,

- traffic light logics referenced by junctions,
- junctions, including their right of way regulation,
- connections between lanes at junctions (nodes).

Note: This chapter 4.2.1. follows the SUMO User Documentation available online („Sumo”, 2015), but the author is confident that it is necessary to be presented here at least in this reduced form since developing of the model is a principal part of this thesis and having this knowledge is essential for understanding it.

3.4.2. Definition of Traffic Demand in SUMO

There are several ways how to define vehicles and generate their routes in SUMO. Based on the available input data, the method using flow definitions and turning ratios was chosen.

The definition of all vehicles in the network is separated into five files:

- my.flows.xml,
- my.static.flows.xml,
- my.interval.flows.xml,
- my.turns.xml,
- my.testcars.rou.xml.

This sorting is needed since each file has different role in the final demand modelling. Some vehicles are generated outside the network and then they go through the area as transit traffic, while other portion of vehicles is generated inside the area since they represent people living there.

Hence:

- **my.flows.xml**

This file generates a major part of all simulated vehicles. It contains all inflows with more than 200 vph. It means there are 18 incoming flows with the biggest effect on the traffic demand in the entire network.

The numbers of generated vehicles in the basic set up before the calibration are taken from the historical database of counted data described in chapter 4.2.3.

For further calibration each 15-minute interval is defined as three shorter 5-minute intervals with same values. The reason for doing this will be explained later in chapter 3.5.

Every flow is described in a single line which looks like this:

```
 <flow id="<STRING>" from="<FROM_EDGE_ID>" number="INT"
departlane="<DEPARTLANE>" departPos="<DEPARTPOS>" departSpeed="8.33"
speedDev="0.1"/>
```

Each of these attributes has a certain meaning and value range (only the used attributes are listed):

*Table 8: Attributes of *.flows.xml*

Attribute Name	Value Type	Description
id	id (string)	The ID of the flow. It is created as "<departing_node>_<number of interval>".
from	referenced edge id	The name of the edge the vehicles of the flow starts at.
number	int	The number of vehicles that shall be inserted during the particular interval.
departlane	string I use "best" and "random"	Determines on which lane of the edge the vehicle is tried to be inserted. "best": the least occupied lane of those who allow the vehicle the longest ride without the need to lane change. "random": a random lane is chosen.
departPos	string I use "random_free" and "base"	Determines the position on the chosen departure lane at which the vehicle is tried to be inserted. "random_free": at first, a "random" position is tried, then the "free", if the first one failed. "base": the vehicle is tried to be inserted at the position which lets its back be at the beginning of the lane.
departSpeed	float(m/s)/string (≥0,"random","max") set to "8.33"	The speed with which the vehicle shall enter the network. It is set to 8.33 m/s (30km/h).
speedDev	float set to "0.1"	The attributes "speedFactor" and "speedDev" are used to sample a vehicle specific chosen SpeedFactor from a normal distribution with mean speedFactor and deviation speedDev. Using speedFactor=1 (default) and speedDev=0.1 results in a speed distribution where 95% of the vehicles drive between 80% and 120% of the legal speed limit.

Beside these attributes describing the conditions of inserting vehicles in the network, a lot of attributes describing vehicles behaviour can be set. They can define for instance the minimum accepted gap in a queue, maximum acceleration or deceleration, number of seats and others. Since there are no better available data for that, this option of customization is not used and default values are kept.

- **my.static.flows.xml**

This file generates the second (minor) part of vehicles inserting the network as inflows. It contains all other external inflows. According the condition whether the inflow belongs to this group or to the group defined in the previous file, these inflows have volumes less than 200 vph.

Since this file is not calibrated through the iteration process, it is not as detailed as "my.flows.xml" in terms of time dispersion. It is defined in 15-minute intervals as it is the most detailed level what the data are available for.

Each flow is then described in the exactly same form as in the file "my.flows.xml".

- **my.internal.flows.xml**

The previous two files insert all the vehicles arriving into the network. However, the area generates some traffic itself. It has a size of ca. 180 hectares (450 acres). So this traffic should not be completely neglected. No data are available, so there is qualified estimate used. It assumes that there is one vehicle leaving a house in the observed one hour interval. Vehicles are distributed equally throughout the given interval. The average value might be around one vehicle per 10 meters of the length of an edge. In some cases vehicles are set to be inserted rather in the downtown direction. The total number of vehicles generated inside the area is 1537 per 60 minutes.

- **my.turns.xml**

To describe the turn definitions, a further file has to be built. Within this file, the list of percentages to chose the particular movement is defined for every approach.

Example:

```
<fromEdge id="S36">
    <toEdge id="C82" probability="0.03989"/>
    <toEdge id="S34" probability="0.96011"/>
</fromEdge>
<fromEdge id="C80">
    <toEdge id="C82" probability="0.60674"/>
    <toEdge id="S34" probability="0.39326"/>
</fromEdge>
```

The values are computed from the historical database of counted data. Even though it would be possible to divide it into four 15-minute long intervals. Average values are used. The first reason for it is that that approach gives more average values and eliminates some patterns appeared only in the time of the traffic survey. The second reason is that there usually occurs no significant differences in time period between 7:30 and 8:30.

SUMO needs an input as a file(s) defining routes with extension *.rou.xml. Therefore the files above are converted into one file "**my.rou.xml**" by JTRROUTER.

JTRROUTER computes vehicle routes that may be used by SUMO using different amount definitions and junction turning percentages. It needs the following inputs: a road network, a demand definition, and junction turning ratios.

As long as there is any change in the definition of traffic modelling, JTRROUTER must be run again.

In this case one output file is created from few flow definitions, therefore a configuration file "**my.jtrc.cfg**" containing all attributes of JTRROUTER is created. Then it can be run easily as:

```
jtrrouter -c my.jtrc.cfg
```

JTRROUTER compiles all the files above into one file "my.rou.xml".

Note: Only origin edges are defined for vehicles, not destinations, which are a result of probability calculations, therefore there are warnings that routes cannot be closed. They could be closed when using DUAROUTER using inputs in form of O-D matrixes, but this is not that case. It is not an error.

- **my.testcars.rou.xml**

Up to here, basically all vehicles modelling the real situation have been defined. For the calibration process, it is essential to declare also the test vehicles which can be subsequently analyzed. They have a role here only as probes. They represent the real floating test cars equipped with GPS units onboard.

Since their trajectories are known, they are not computed based on probabilities of turning ratios, but they are defined directly in another file.

There were in total 23 test cars during the simulated 60-miute interval. They are considered to be samples of absolutely random average vehicles from that time. Therefore, it is better to simulate also the most average vehicles. For that purpose there are generated 10 vehicles within 1 minute (the same minute as the real probe starts its route. The characteristics of these 10 vehicles are recorded and averaged for comparing them with the real values. It means that there are 10 simulated vehicles associated to one real car.

For easy identification the test cars while using SUMO-GUI, they are visualized in red colour and use their one vehicle type class "probe".

The files "my.rou.xml" and "my.testcars.rou.xml" are then read at the same time by SUMO configuration file "my.sumocfg".

The 60-minute interval (7:30 to 8:30 AM) is simulated, but as in any other similar simulation program there must be a phase when the network becomes saturated. For this reason, there is half-hour interval beginning in "0" and ending in "1800" in seconds, which is never analyzed. The saturation is made by files "my.flows.xml" and "my.static.flows". The file "my.internal.flows.xml" does not contain this saturating interval, because its purpose is not principal, but rather auxiliary.

As it can be seen in Figure 5, by far the highest traffic volume is generated by definitions in "my.flows.xml". In the first step of calibration process there are 14,601 vehicles inserted in the network in the sum of all directions and trajectories in time 7:30 to 8:30 AM.

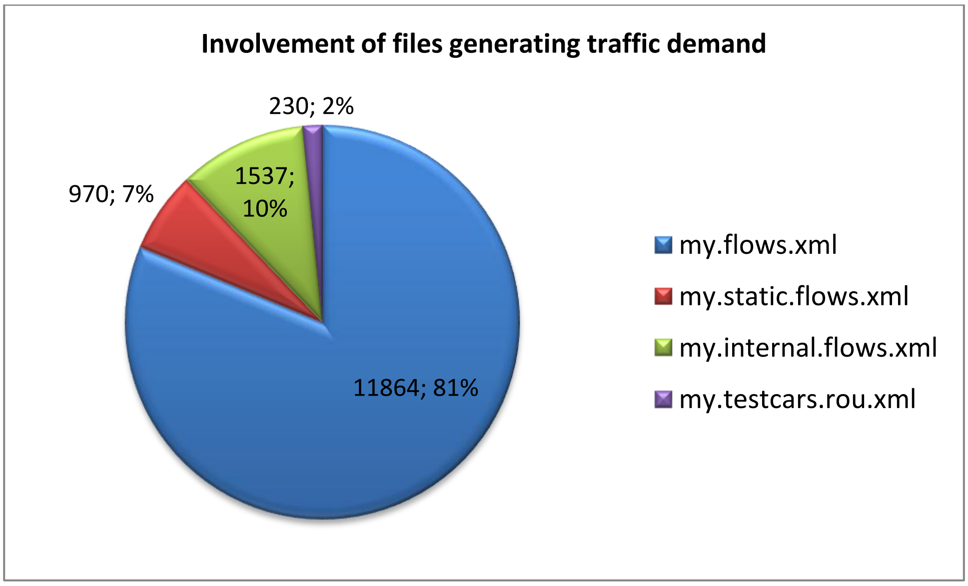


Figure 5: Generation of vehicles

3.4.3. Running the simulation

Traffic simulation in SUMO can be conducted in two ways as described below. The overview of the simulation process is given in Figure 6. All file names in the brackets are the file names used in this project.

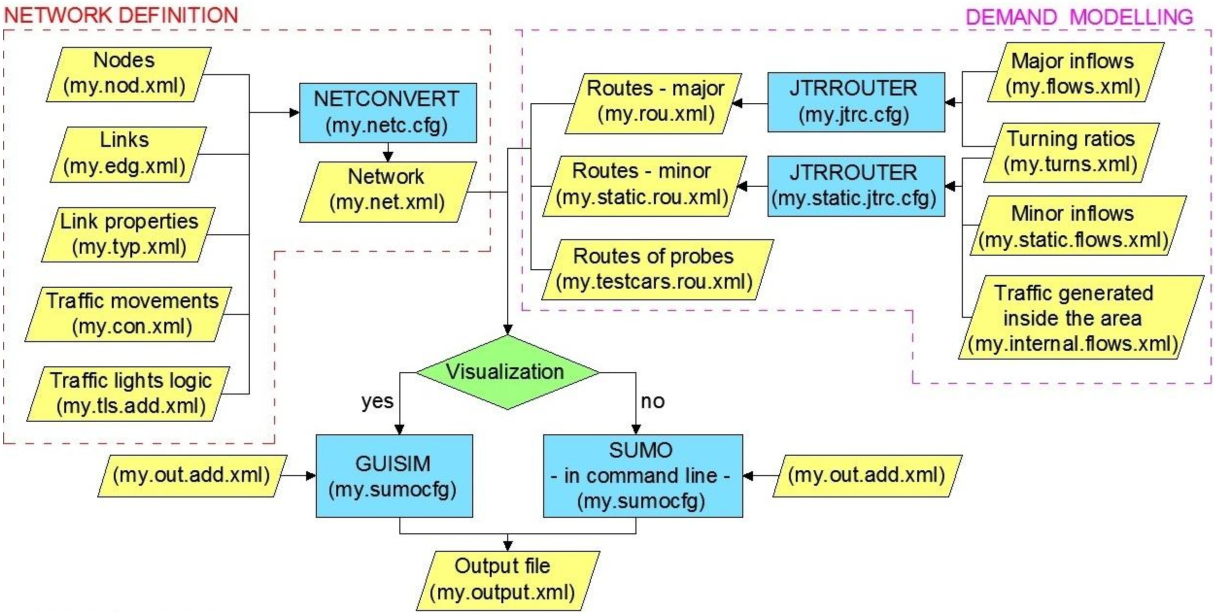
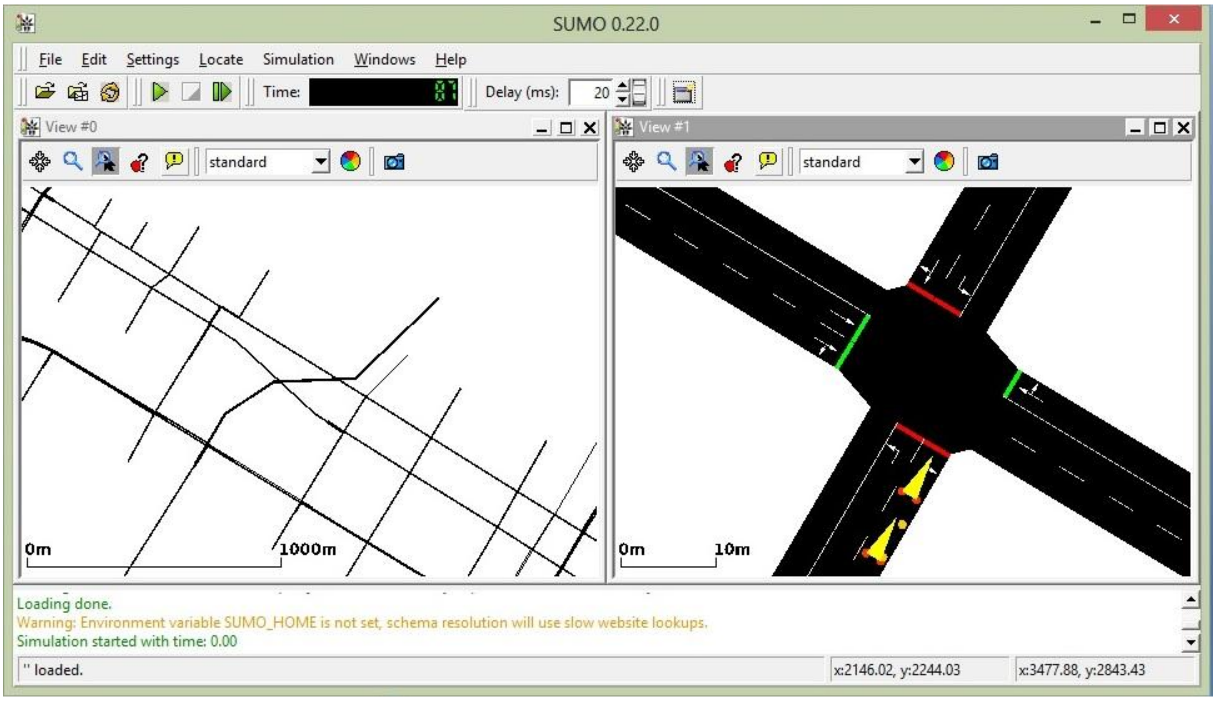


Figure 6: Overview of traffic simulation process

An efficient traffic simulation execution can be achieved with the use of command line, especially when dealing with networks or when multiple run is required. To simplify the execution process, all the required execution actions, e.g. the path and the name of the input files, the output types, the output directory and the simulation time period is specified in a configuration file "my.sumocfg". Then the simulation can be carried out with the use of the following command.

```
sumo -c my.sumocfg
```

The application of SUMO-GUI is the other way to execute the traffic simulation with SUMO. It is basically the same application as SUMO, just extended by a graphical user interface. During the execution each vehicular movement and the traffic progression can be observed and the possible bottlenecks can be visually identified.



```
Loading done.
Warning: Environment variable SUMO_HOME is not set, schema resolution will use slow website lookups.
Simulation started with time: 0.00
```

```
" loaded.    x:2146.02, y:2244.03    x:3477.88, y:2843.43
```

Figure 7: Graphical User Interface of SUMO

As long as I use GUI only for assuring myself that everything runs correctly, no advanced settings for GUI are applied.

A configuration file for all execution actions is required in SUMO-GUI. The simulation can be open by

```
sumo-gui -c my.sumocfg
```

or simply by double click or calling

```
my.sumocfg
```

The step of simulation is set to 1 second. It means that all states and parameters are recalculated every second.

The length of the simulation is 6000 seconds. First 1800 seconds are used for saturating the network by vehicles, then there is 3600 second interval of our interest, and the last 600 seconds are added for a possibility that the evaluated vehicles could have not left the network within the main time interval.

The output of the simulation is required in format of "VTypeProbe". In this order probe cars are classified as **`type="probe"`** so they can be evaluated separately beside all other vehicles. The output is specified in additional file "my.out.add.xml". This file defines that the output is saved in the following style: There is a trace of every test car every second in terms of its position (coordinates and links where it is located that second) and current speed. It means that basically the same parameters as in the real floating cars are recorded. However, this output file is transformed afterwards and aggregated by links. Travel time of each test car is saved for each link on the evaluated part of its route. This gives a comparable parameters which sum of differences is minimized by iteration process while number of cars on inputs is changed.

3.5. Model calibration

3.5.1. Technique of calibration

The first step in the entire calibration process is about setting all input values according to available data.

For this purpose counted data from historical database („Comptage des véhicules et piétons”) are used as an input for inserting traffic flows. The model is calibrated on 15-minute intervals level in this phase. However, it is still too aggregated to capture peaks within that time period. Therefore it is further refined in next steps.

The second item which is calibrated in this phase is the speed on the studied route. The speed (from our survey made by floating cars) which can be called FFS as set there as maximum speed limit. The maximum mean speed over a link is used - in most cases that speed was reached in early morning before 7 AM.

However, FFS is a parameter influenced mainly by road geometry, not by time of day, so it is valid even for the peak hour. It is set for each link separately. The static calibration is done.

Up to here, FCD are used for general setting of the model, but they do not bring information about actual "live" traffic conditions. It is achieved in the second phase of the calibration process. It works with the following idea.

By adding the input file "my.testcars.rou.xml" the floating cars from the real world are simulated. As has been already mentioned, each real car is represented by 10 simulated vehicles inserting the network in the same minute as the real car started its trip and average values are then evaluated. This way each real car is associated to one simulated vehicle at the end.

Then an output with travel times on each link is required for each model simulated vehicle. These times are given into a matrix with the following structure shown in Table 9.

Table 9: Matrix of travel times

Vehicles (23)	Links	Travel times
List of all vehicles (23) ID: "P-SB-*-avg" (6 vehicles) "P-NB-*-avg" (5 vehicles) "U-SB-*-avg" (6 vehicles) "L-NB-*-avg" (6 vehicles) * is number referenced to the order of floating car in 7:30-8:30 the on the particular route. (Starting times can be found in Table 14)	List of all links the vehicles went through: Parc - SB: 11 links Parc - NB: 10 links St. Urbain - SB: 10 links St. Laurent - NB: 12 links	TT on each link in seconds

Note: Because not all links the vehicle passed through are a part of the evaluated matrix above, only some links appearing in the route of the car are recorded into the matrix. And on contrary, when there is a change of number of lanes on the link, SUMO types them down as two links with different IDs. SUMO also makes difference between lanes of the same network. It means that IDs in the output file "my.output.xml" are slightly more complex than is required, so searching condition is brought in to record the TTs in the proper format.

There is an additional file "my.real.add.xml" with the exactly same structure, but with values of TTs measured in field by real FC.

Then it is easy to compare these two files and determine a concordance of the values.

The differences between real TTs and TTs from the simulation is summed over all vehicles and all links. Second power of differences is used to eliminate the issue that it can be sometimes positive and sometimes negative number. The objective is to minimize this sum.

Hence:

$$V = \min\left[\sum (TT_{obs} - TT_{sim})^2\right]$$

where:

V is the objective function,

TT_{obs} is the real observed travel time on a link,

TT_{sim} is the travel time from the simulation.

The sum is of 6*11 + 5*10 + 6*10 +6*12 = 248 members $(TT_{obs} - TT_{sim})^2$. (See Table 9.)

Travel time is a function of traffic demand, therefore traffic flows are being changed for the purpose to affect value of the objective function *V*. Traffic demand is being changed for an inflow defined in "my.flows.xml" in a random 5-minute interval.

The five-minute intervals were chosen, because they can still uncover some hidden fluctuation of demand which cannot be seen in the basic 15-minute interval. Of course that it would be better to work with even smaller interval, but then another aspect must be concerned. The simulation is not deterministic, but it contains a certain stochastic behaviour. If there would be a small change of inflow (e.g. +5 vehicles within interval 7:30-7:31), it could be cancelled by a change of vehicles inserted by different probabilistic computation of vehicle routes in all 60 1-minute intervals. The total value of *V* could not reflect the change made on the input.

To save some time during the computation process and to avoid unrealistic results the change of flows on input is set to be only in range <$0.5q_0$, $1.5q_0$>, where q_0 is the value of traffic flow in the original "my.flows.xml".

The iteration process is under the condition caused by SUMO topology can be encapsulated in the 3 steps:

1. Change "my.flows.xml".
2. Run JTRROUTER to compute routes with the new number of vehicles on input.
3. Run SUMO to get new values of TTs on links and compute new *V*.

If the new value of *V* is smaller than in the previous iteration, it means that the difference between simulated vehicles and the measured data decreases. In that case, the change made in "my.flows.xml" is evaluated as a step in a good way and is kept. Otherwise it is declined and the flow file is put back into the state before this step of iteration. Then another value of inflow is made and the loop runs again.

The flow chart of the iteration process is drawn in Figure 8.

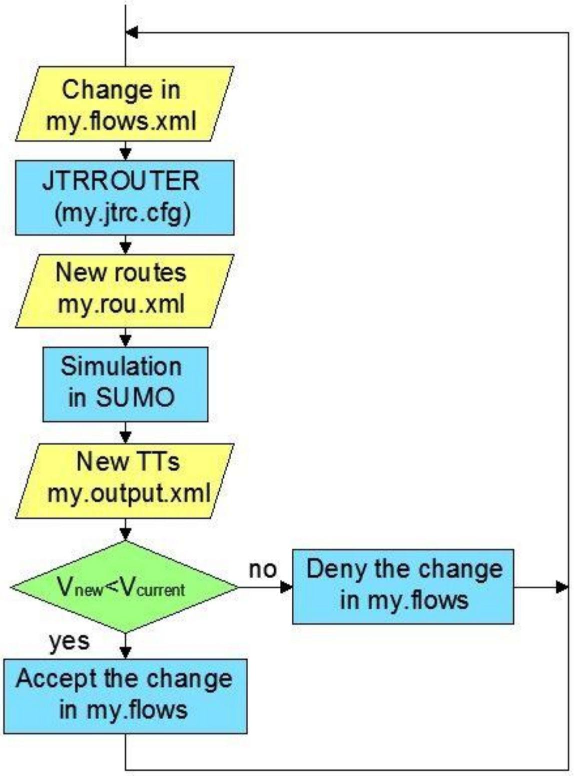


Figure 8: Iteration process of model calibration

As JTRROUTER must be run in every iteration, it is convenient to give it as small input file as possible. For that purpose, "my.flows.xml" is separated from all other files defining input flows and SUMO reads these files of routes parallelly. There is also a version in development using a separate flow xml file to every inflow and every time interval. It works with 216 small files (18 inputs * 12 5-minute intervals).

Sample of GPS traces is giving the dynamics to the simulation.

Travel time was chosen as the parameter of calibration, because it maintain some scale based on a length of links. If there were an evaluation based on average speeds on links, there would be no difference between short and long links. Although it is obvious that if there were a high difference on long links, the error would be bigger.

According to different approaches to picking the values on input to be changed, there are two methods.

3.5.2. Progressive Exhaustive Search

This method uses a brute-force search. Since there is a very high number of possible combinations of values on inputs in different ID of a flow and different time intervals, there is a proposal of bringing some logics into the process by implementing a decision tree..

1) Compute the value of V from the basic state from the iteration 0.

2. Add e.g. +10% for the number of vehicles generated in the first time interval of the strongest flow.

3) If $V_{new} < V_{current}$ add another +10% (up to +50% in total) until the sum of squares of TTT differences (V) is still smaller than in the previous step. Else -10%. Then optimum value for one in flow in one interval is found.

4) Move to the next time interval and repeat step 3) for all 12 intervals. At the end one flow is calibrated.

5) Move to the next strongest inflow and repeat steps 2) to 4).

6) All flows in all intervals are calibrated.

The drawback of this method is that it might find only local minimum of V. Therefore method of simulated annealing is preferred.

3.5.3. Simulated Annealing

Simulated annealing (SA) is a generic probabilistic metaheuristic for the global optimization problem of locating a good approximation to the global optimum of a given function in a large search space. It is often used when the search space is discrete. For certain problems, simulated annealing may be more efficient than exhaustive enumeration — provided that the goal is merely to find an acceptably good solution in a fixed amount of time, rather than the best possible solution.

The name and inspiration come from annealing in metallurgy, a technique involving heating and controlled cooling of a material to increase the size of its crystals and reduce their defects.

Table 10: Relationship between physical annealing and simulated annealing

Thermodynamic Simulation	Combinatorial Optimisation
System States	Feasible Solutions
Energy	Cost
Change of State	Neighbouring Solutions
Temperature	Control Parameter
Frozen State	Heuristic Solution

This notion of slow cooling is implemented in the Simulated Annealing algorithm as a slow decrease in the probability of accepting worse solutions as it explores the solution space. Accepting worse solutions is a fundamental property of metaheuristics because it allows for a more extensive search for the optimal solution.

$$P(accept) = \min\left(1, \frac{\exp(-\lambda V)}{\exp(-\lambda V')}\right)$$

where:

P is the probability of acceptance the change made in the particular iteration *i*,

V is the objective function $V = \sum(TT_{obs} - TT_{sim})^2$, V' is its value from the previous step,

$\lambda = c*log(1+i)$,

c is a calibration parameter defining the drop of acceptance states which are not optimum throughout the iteration process,

i is the iteration number.

The equation above says that probability that a result which is not optimum is also accepted is quite high in early phases results, but it declines with the number of iterations. If the value of the objective function *V* is smaller than in the previous iteration, then it is always accepted. If it is bigger, it is accepted only with probability equal to $\frac{\exp(-\lambda V)}{\exp(-\lambda V')}$. This method is convenient for its skill to get out of staying in local optimum.

Based on this technique using simulated annealing, there is a program written in Java being developed by Anas Balboul, student of Polytechnique Montréal. The working version of the code is in Enclosure 2.2.

This program does all the iteration procedure described in Figure X and with decision-making condition described in this chapter 3.5.3. It also records the value of *V* if the iteration is selected to be accepted. By plotting these values vs number of iteration there should be a declining trend in and its convergence theoretically to zero. Null is reached in state when there is no difference between any travel time of any simulated vehicle and associated real test car on any link.

4. Case Study

4.1. Information about the area

4.1.1. Location

The case study what I work on in this thesis is located in the city of Montreal in Canada.

Montreal is the largest city of the province Quebec and the second largest in Canada after Toronto.

The city is located in the Island of Montreal and a few smaller islands in its immediate proximity. In 2011 the city had a population of 1,649,519 at the land area of 365.13 km^2. It gives a population density of 4,517.6 persons per square kilometre. While Montreal's metropolitan area had a population of 3,824,221 and area of 4,258.31 km^2. („Montreal - Census", 2015). Montreal census metropolitan area includes dozens of smaller cities and towns out of Montreal Island. The two most important ones are Laval in the north with population of 401,553 and Longueuil in the south with population of 231,409. For this particular case study should we consider mainly Laval since its location generates trips in southbound direction during morning peak hour.

The case study area contains basically two parallel routes.

The first route from the west is Boulevard de l'Acadie (in the northern part) and Avenue du Parc (in the southern part). These streets are connected by Avenue Beaumont (in downtown direction) and Rue Jean-Talon-Ouest (in uptown direction). Boulevard de l'Acadie and Avenue du Parc are two-way roads in the solved area.

The second route consists of Boulevard St. Laurent in northbound direction and Boulevard St. Laurent - Rue Clark - Rue Saint-Urbain in southbound direction. This route operates as two-way street between motorway Autoroute Métropolitaine and Rue de Castelnau and as two one-way streets from intersection with Rue de Castelnau towards downtown.

The location of picked routes of the case study is visible in Figure 9. Routes are marked in red and downtown area is in blue.

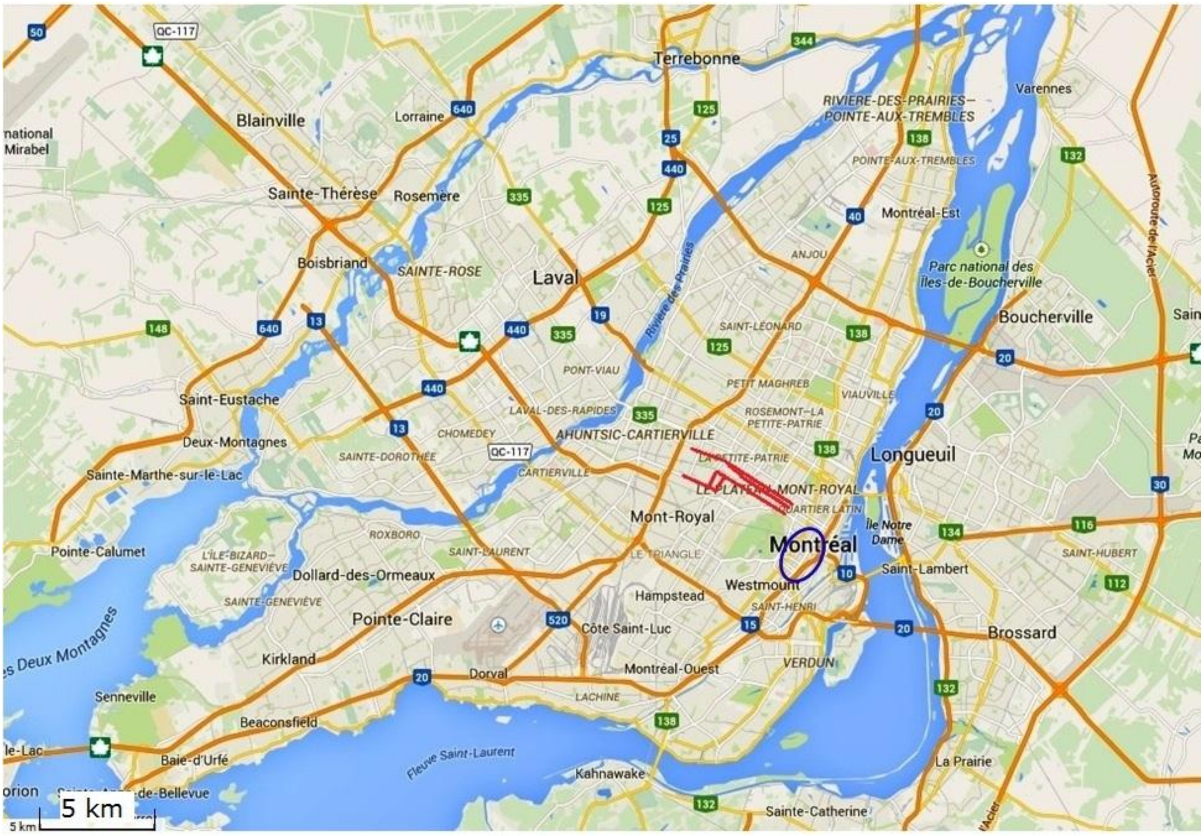


Figure 9: City of Montreal

4.1.2. Description of the streets

All three streets are arterial roads. They can be classified as roads of type "B" according to Czech standard ČSN 73 6110. Their main function is to connect highway 40 which is going through the city in northeast-southwest direction (this highway is also known as "Route Transcanadienne") to downtown area in the south part of the Island of Montreal. In view of the fact that there is a hill Mont Royal in the centre of the city which is a park with no through traffic, Avenue du Parc is the last road on the east part of this hill. The next parallel road is Chemin de la Côte-des-Neiges on the west side of Mont Royal which is in distance of 3.2 km from Av du Parc.

On the other hand, Av du Parc, St. Urbain, and St. Laurent are really close one another and they should be therefore consider together in traffic analysis. Their distance in the south part of the discussed area is 290 m, 155 m respectively.

Each of these routes is divided into several links for further analysis. The nodes between represents intersections which are controlled by traffic lights. The detailed description of roads of the case study is provided in Tables 11 and 12.

Table 11: Boulevard de l'Acadie + Avenue du Parc

Boulevard de l'Acadie + Avenue du Parc											
Southbound direction						Northbound direction					
No. of Node	Node Name	Length [m]	No. of lanes	Facilities	FFS	No. of Node	Node Name	Length [m]	No. of lanes	Facilities	FFS
1	Liege	-				22	Liege	491	3	(P)	58
2	Jarry	521	3	-	50	21	Jarry	529	3	(P)	51
3	St. Roch	529	3	-	54	20	St. Roch	487	3	(P)	45
4	Jean-Talon	469	3	-	49	19	Acadie	172	2 of 4	2xP	34
5	Beaumont	228	1 of 2	-	43	18	Wiseman	176	2 of 4	2xP	32
6	d'Outremont	262	1 of 2	-	43	17	Bloomfield	152	2 of 4	2xP	51
7	Querbes	272	1 of 2	-	47	16	Querbes	132	2 of 4	2xP	38
8	Parc	208	1 of 2	-	42	15	Hutchison	74	2 of 4	-	38
						14	Jean-Talon	351	2 of 4	(P)	49
						13	Beaumont	464	2 of 4	-	60
9	Beaubien	495	2	-	53	12	Beaubien	256	2	-	56
10	Van Horne	238	2 (-1)	P	71	11	Van Horne	434	1(+1)	P	50
11	Bernard	435	2 (-1)	P	60	10	Bernard	143	1(+1)	P	50
12	Pieton (crossing)	158	2 (-1)	P	58	9	Pieton (crossing)	153	1(+1)	P	47
13	St. Viateur	139	2 (-1)	P	49	8	St. Viateur	200	1(+1)	P	48
14	Pieton (crossing)	242	2 (-1)	P	59	7	Pieton (crossing)	211	1(+1)	P	56
15	Fairmount	175	2 (-1)	P	48	6	Fairmount	167	1(+1)	P	45
16	Laurier	163	2 (-1)	P	33	5	Laurier	134	1(+1)	P	51
17	St. Joseph	127	2 (-1)	P	52	4	St. Joseph	173	1(+1)	P	46
18	Villeneuve	191	2 (-1)	P	54	3	Villeneuve	294	1(+1)	P	51
19	Mt. Royal	294	2 (-1)	P	50	2	Mt. Royal	394	2+2 turning L	-	62
20	Ch. Ste. Catherine	87	2	-	54						
21	Pieton (crossing)	300	3	P	60	1	Pieton (crossing)	-			
total		5533						5587			

Table 12: St. Urbain + St. Laurent

St. Urbain + St. Laurent											
Southbound direction (St. Urbain)						Northbound direction (St. Laurent)					
No. of Node	Node Name	Length [m]	No. of lanes	Facilities	FFS	No. of Node	Node Name	Length [m]	No. of lanes	Facilities	FFS
1	Liege	-				22	Liege	250	2	P	54
2	Guizot	253	2	P	38	21	Guizot	283	2	P	58
3	Jarry	274	2	P	48	20	Jarry	238	2	P	49
4	Gounod	267	2	P	48	19	Gounod	269	2	P	58
5	Villeray	257	2	P	56	18	Villeray	242	2	P	58
6	Faillon	243	2	P	56	17	Faillon	220	2	P	55
7	de Castelnau	178	3	-	43	16	de Castelnau	153	2	P	55
8	Jean-Talon	221	3	P(L)	31	15	Jean-Talon	270	2	P(L+R)	57
9	Mozart	273	2	P(L+R)	27	14	Mozart	120	2	P(L+R)	55
						13	Dante	178	2	P(L+R)	49
10	St. Zotique	302	2	P(L+R)	52	12	St. Zotique	255	2	P(L+R)	48
11	Beaubien	268	2	P(L+R)	49	11	Beaubien Ouest	81	2	P(L+R)	50
						10	Beaubien Est	566	2	P(L+R)	50
12	Bernard	676	2	-	37	9	Bernard	308	2	P(L+R)	60
13	St. Viateur	320	2	P(L+R), C	50	8	St. Viateur	418	2	P(L+R)	59
14	Fairmont	406	2	P(L+R), C	53	7	Fairmont	161	2	P(L+R)	55
15	Laurier	155	2	P(L+R), C	50	6	Laurier	146	2	P(L+R)	51
16	St. Joseph	131	2	P(L+R), C	50	5	St. Joseph	185	2	P(L+R)	45
17	Villeneuve	191	2	P(L+R), C	52	4	Villeneuve	266	2	P(L+R)	45
18	Mt. Royal	284	2	P(L+R), C	48	3	Mt. Royal	208	2	P(L+R)	57
19	Marie-Anne	222	2	P(L+R), C	49	2	Marie-Anne	228	2	P(L+R)	56
20	Rachel	214	2	P(L+R), C	50	1	Rachel	-			
total		5135						5045			

Notes:

Length: length of links is measured from the axis of the crossing street to the next axis.

No. of lanes: Avenue du Parc has layout including 3 lanes - one per each direction and one extra lane in the middle of the street which has a traffic light signal. It's open for the direction with currently higher demand. It means in this case that there are two lanes in southbound direction and one lane in northbound direction in the morning, but two lanes northbound and only one lane southbound in the afternoon. This is visible also in Figure 10.

Hence: 2(-1) = 2 lanes are open in that direction, but one might be closed in other daytime

1(+1) = 1 lane is open in that direction, but another one might be opened in other daytime

In case the route was measured one-way but on two-way road, I use "1 of 2" in a meaning that there is one traffic lane for each direction on the road, or "2 of 4" if there are 2 traffic lanes per direction respectively.

Figure 10: Extra traffic lane with an option of dynamic change of its direction

Facilities: P = there is parallel parking on the link on the right side of the street

If there is a note (L) or (L+R) it says that there is a parallel parking on the left side of the street or on both sides.

C = there is a cycle lane on the right side of the carriageway.

FFS: This column shows an empirical free flow speed on the link. It is the maximum average speed on the link measured by a test car

The trajectories are marked in Figure 11. Even though it may seem there are one-way streets between Acadie and Parc, Beaumont and Jean-Talon are both two-way streets. The route is marked through them, because it was leaded in the same way during a measurement by a floating car.

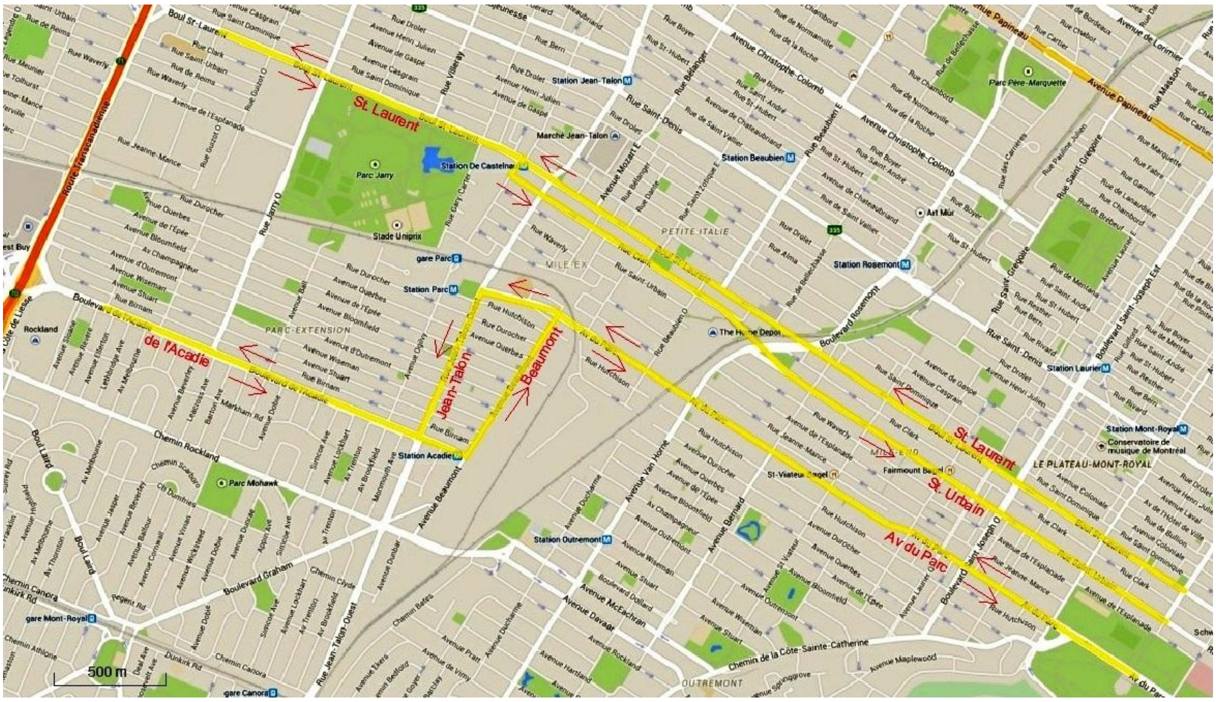


Figure 11: Map of solved area

Mainly the southern part of the area is full of miscellaneous businesses. There are dozens of stores, markets, restaurants, bars, and others. These should be considered especially during analysis on Friday afternoon in northbound direction or at weekends. As long as I am modelling a morning peak hour there is no relevant reason to assume that those facilities should be busy at that time when the demand on the roads is strongly focused on transit traffic aiming to downtown area.

Almost all intersections are controlled by traffic lights. Their schedules are fixed and time dependant. So they are switching during the day, but they cannot adapt on changing traffic demand.

There are lines of public transport buses on all three routes. The system of bus stops is different from the general European system. Distances between stops are very short - there is usually a stop nearby every intersection. Buses stop always only on a demand of a traveller. Nevertheless, they are not considered in my model. It is a sort of simplification, as there would be a requirement for a detailed traffic survey describing where and how often bus stops, and for how long time. Buses have no preferences at signal controlled intersections.

For similar reasons parking, pedestrians, and cyclists are also not considered in the model, even though they might occasionally affect a drive of a floating car.

4.2. Datasets

There are two main data sources what I use in my project. The first one is a set of GPS data from a floating test cars. The second source is a database of counted data in the city of Montreal.

4.2.1. GPS dataset description

The mainstay of this thesis is a use of GPS dataset. The data were recorded by Montreal city's test car equipped with a high-quality GPS unit.

The data were collected during several different days in 2005. Different days were chosen to have a sample for various days of a week and time of a day. The sample represents a random average day. My simulations are designed for morning peak hour during an average working day.

Table 13: List of trajectories with GPS records

Acadie + Parc SB	Acadie + Parc NB	St. Urbain SB	St. Laurent NB
Mo 7 Mar 2005 6.32	Mo 7 Mar 2005 6.30	We 23 Feb 2005 6.50	We 23 Feb 2005 6.31
Tu 22 Feb 2005 6.34	Mo 21 Feb 2005 6.48	Th 24 Feb 2005 6.40	Th 24 Feb 2005 6.31
Tu 22 Feb 2005 6.43	Tu 22 Feb 2005 6.50	Th 24 Feb 2005 6.54	Fr 25 Feb 2005 6.32
Mo 7 Mar 2005 6.44	Tu 22 Feb 2005 6.58	Fr 25 Feb 2005 6.54	Fr 25 Feb 2005 6.37
Mo 21 Feb 2005 7.05	We 27 Apr 2005 7.02	Th 24 Feb 2005 6.58	Th 24 Feb 2005 6.52
Tu 22 Feb 2005 7.05	Mo 7 Mar 2005 7.03	Th 24 Feb 2005 7.23	We 23 Feb 2005 6.57
Mo 7 Mar 2005 7.18	Mo 21 Feb 2005 7.21	We 23 Feb 2005 7.28	We 23 Feb 2005 7.03
We 27 Apr 2005 7.18	Tu 22 Feb 2005 7.21	We 23 Feb 2005 7.31	Fr 25 Feb 2005 7.08
Tu 22 Feb 2005 7.20	We 27 Apr 2005 7.36	Th 24 Feb 2005 7.35	Th 24 Feb 2005 7.10
Mo 21 Feb 2005 7.32	Tu 22 Feb 2005 7.38	Fr 25 Feb 2005 7.36	Th 24 Feb 2005 7.11
Tu 22 Feb 2005 7.37	Mo 7 Mar 2005 7.41	Fr 25 Feb 2005 7.40	Th 24 Feb 2005 7.37
Mo 21 Feb 2005 7.38	Tu 22 Feb 2005 8.00	Th 24 Feb 2005 8.07	We 23 Feb 2005 7.44
We 27 Apr 2005 7.52	We 27 Apr 2005 8.12	We 23 Feb 2005 8.17	We 23 Feb 2005 7.48
Tu 22 Feb 2005 7.57	Tu 22 Feb 2005 8.18	We 23 Feb 2005 8.23	Fr 25 Feb 2005 7.51
Mo 7 Mar 2005 7.58	Mo 7 Mar 2005 8.20	Fr 25 Feb 2005 8.23	Fr 25 Feb 2005 7.59
Mo 21 Feb 2005 8.26	We 27 Apr 2005 8.51	Fr 25 Feb 2005 8.27	Th 24 Feb 2005 8.04
We 27 Apr 2005 8.32		Th 24 Feb 2005 8.32	Th 24 Feb 2005 8.30
Mo 7 Mar 2005 8.37			We 23 Feb 2005 8.36
			We 23 Feb 2005 8.43

Note: All these trajectories have been analyzed. The highlighted trajectories have been used to calibrate the model simulation. They are representatives of trajectories of one-hour long time period between 7.30 and 8.30.

Beside these data points there are available data for times around midday and for the afternoon when the flows in northbound direction dominate. There are also individual records from other days available, but they were not used for not finding them a good examples of average days with carrying some certain homogeneity during recording them or just too incomplete. For these reason days when more trajectories had been measured were preferred.

The author realizes that the situation might have changed significantly since that time. However, it is not a real problem. The simulation is rather pseudo-live. It means that it is only a question of measuring new data in field to have the model up to date according to live conditions. The basic principles would still be valid.

The test car was recording its position, speed and additional parameters every second. It provides a very detailed source for reverse modelling of the drive of the car. The example of original recorded data, how the city's employees provided them to this project, is shown in Table 15. The data were provided in format of .xls sheets.

Table 14: Example of recorded GPS traces

Time	Latitude	Longitude	Speed	HDOP	NumSats	Quality	Node	Run
- previous lines -								
23.2.2005 7:56:18	45319668	73368476	23.46232	1.6	8	2	15	12
23.2.2005 7:56:19	45319700	73368542	24.03913	1.6	8	2	15	12
23.2.2005 7:56:20	45319733	73368612	24.59397	1.6	7	2	16	12
23.2.2005 7:56:21	45319765	73368683	24.50665	1.6	8	2	16	12
- following lines -								

The column 1 is the time of the specific line.

The columns 2 and 3 represents GPS position in DDdddddd° form. For subsequent work they are converted into DD°MM.mmm' form which is a value behind the original table in Enclosure 1.1

The column 4 is current speed in mph. It is converted into kph units in another column of the table.

The column 5 represents horizontal dilution of precision in meters. A measure of the geometric quality of a GPS satellite configuration in the sky. HDOP is a factor in determining the relative accuracy of a horizontal position. The smaller the HDOP number, the better the geometry.

The columns 6 and 7 shows a strength and quality of the GPS signal. The information about the number of satellites what GPS unit was connected to is provided.

The column 8 is the number of link on the trajectory. However, they do not correspond with the number in the first column of Table 11 and 12.

The column 9 distinguishes individual drives of the test car.

4.2.2. Processing of data

It is obvious that these data are too detailed for traffic analyses on the flow level. On that account they must have been processed.

In the first step they have been aggregated according to columns "Node" and "Run". That is a way how tables containing information about travel time, average speed, time when the car was not moving, and number of stops were created. All of them are separated in terms of links and in terms of runs. All these tables are provided in Enclosure 1.2.

An analysis made in GIS software QGIS was chosen for further, more detailed, analysis. The points where the test vehicle stopped were analyzed. Their precious position was displayed in the map of the area and the distance from the stop line was measured. If we accept the assumption that each test car represents an absolutely random car of the traffic flow, we can evaluate it in terms of average length of queue and average delay for each intersection. (The average length for one vehicle in a queue is considered to be 6 meters.)

4.2.3. The historical database (counted data)

The second data source is a database of counted data in Montreal area. These data were obtained by municipality of the city. Data were collected manually in 15 minute intervals over a day. Data were collected for all main intersections around the city. The traffic survey in the solved area took place in September, October and November 2003. The database is freely available online („Comptage des véhicules et piétons”).

4.3. Data analysis

4.3.1. Analysis of average speeds and travel times over the entire trajectories

This chapter provides an overview of analyzed characteristics of traffic in the solved area. Data from GPS traces recorded in-situ were used for the following waveforms of average speed Figure 12 and travel times Figure 13 over the entire trajectories.

Plots displaying travel times over the trajectories are also provided as they might show better the differences between peak hour traffic and times with moderate traffic. However, they still have the same information value as the previous plots displaying speed. Their disadvantage in comparison with plots of average speeds is that they cannot be compared one another, because each trajectory has different length.

All these plots show that there is a big difference between morning and afternoon. The traffic is slower in the afternoon, but it evinces bigger dispersion at that time.

Slightly unexpected are the values of the midday times. There is no significant drop of travel times when the morning traffic disappears. This appears especially in the northbound direction, which does not carry the major portion of home-work trips.

In terms of different weekdays there are no significant differences with one exception. There is a very strong traffic demand on Friday afternoon in uptown direction. This is explained by travellers stopping and parking on Avenue du Parc and St. Laurent for the purpose of doing shopping in the local small businesses on their way home. These drivers interrupt the flow by that behaviour.

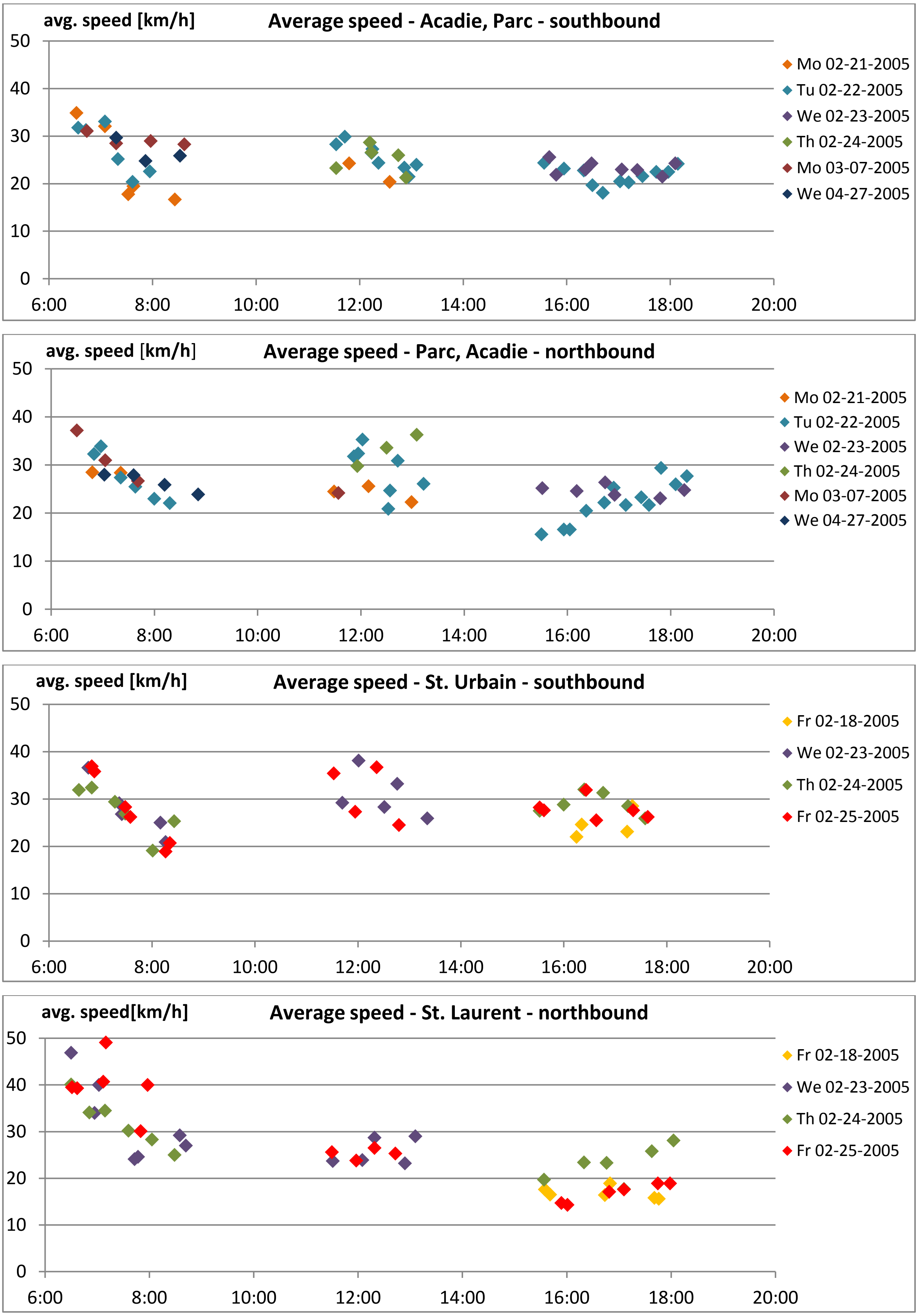


Figure 12: Plots of Average Travel Speed vs. Time-of-Day

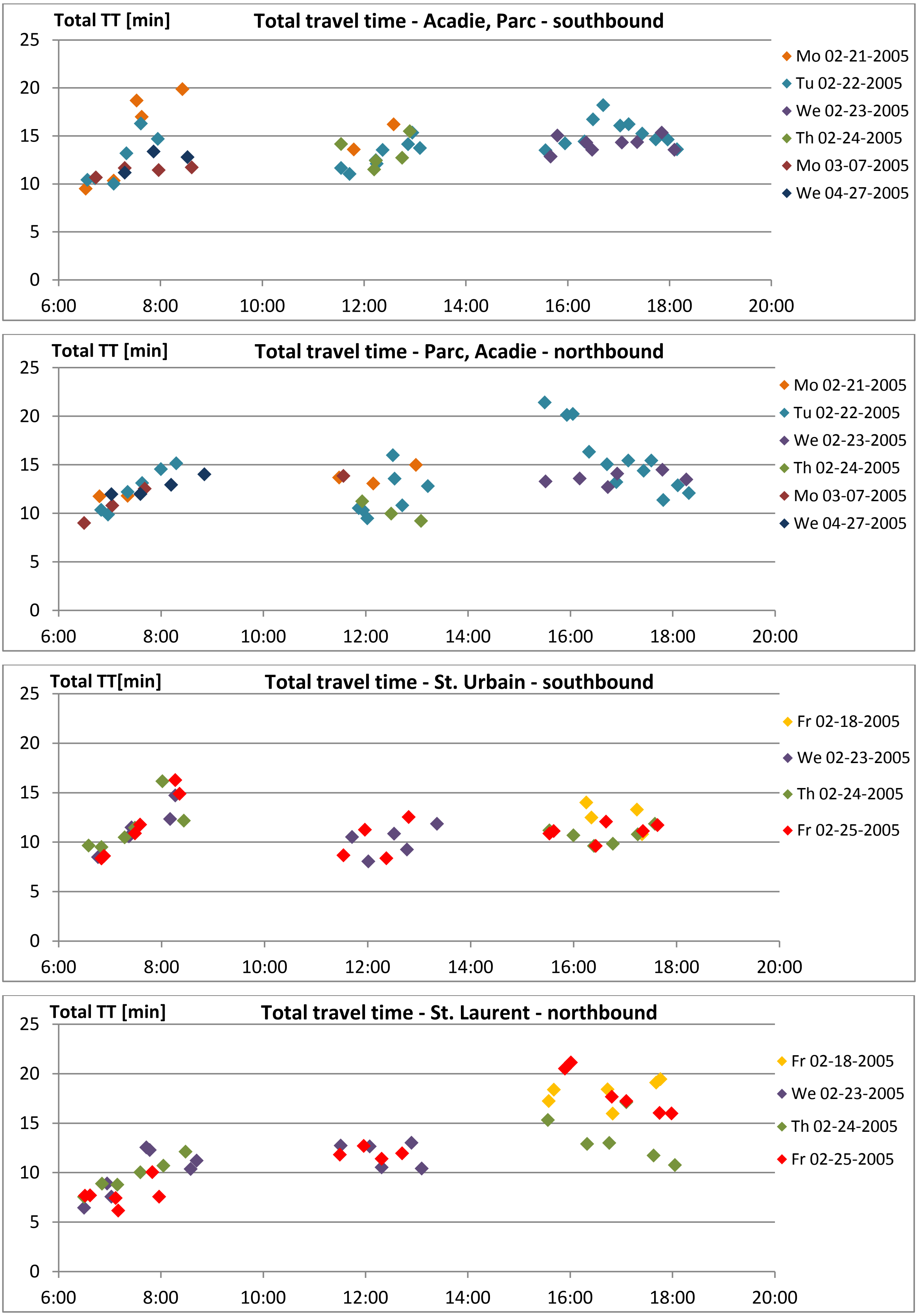


Figure 13: Plots of Total Travel Time vs Time-of-Day

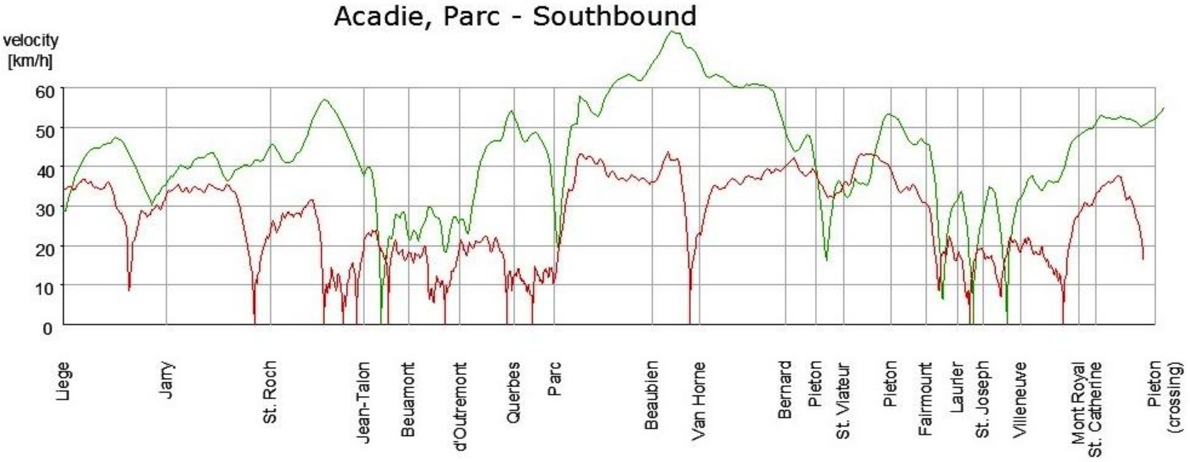
In general, the measured data are more consistent in the morning period. They denote bigger dependence on growing traffic flow how people start travelling to downtown area to work.

These reasons led to idea to focus in the simulation for time between 7.30 and 8.30 AM. Trajectories in that time have already significant delays caused by big volumes of traffic to be interesting for deeper studying.

4.3.2. Analysis of speed over trajectories

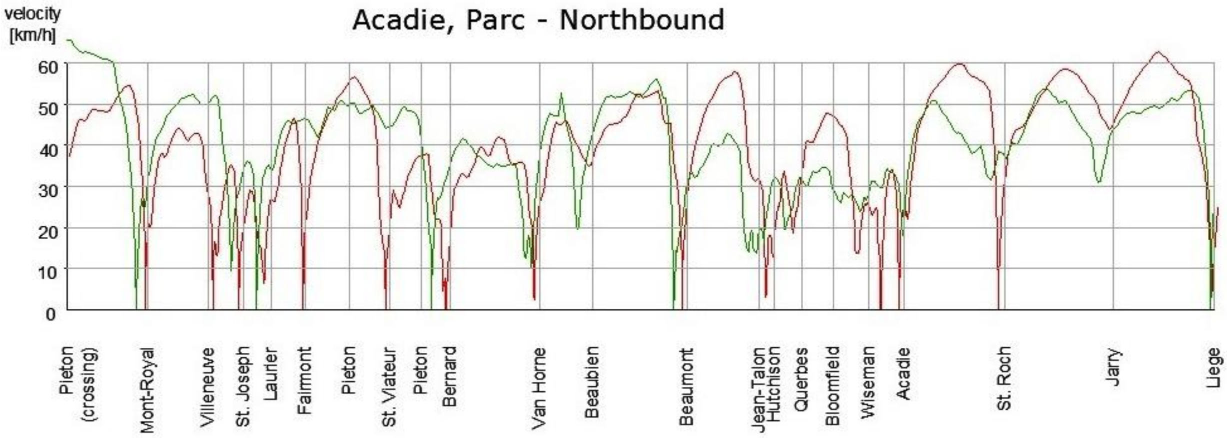
The previous analysis aggregate data from all links together. It is useful to get the general overview of reached average speed, but it is not possible to identify where delays appeared. Therefore I also analyzed each trajectory on very detailed level and I plotted spot speeds recorded every second.

There are examples of differences of the courses of speed measured in early morning before the peak hour started and during the peak hour shown in Figure 14. The complete plots separating the samples into four intervals (before 7.00, 7.00 - 7.30, 7.30 - 8.00, after 8.00) are in Enclosure 1.4. Enclosure 1.4 involves also additional information of the analysis - the distance of the test vehicle from stop line and the time in seconds for how long time the vehicle stopped there are included.

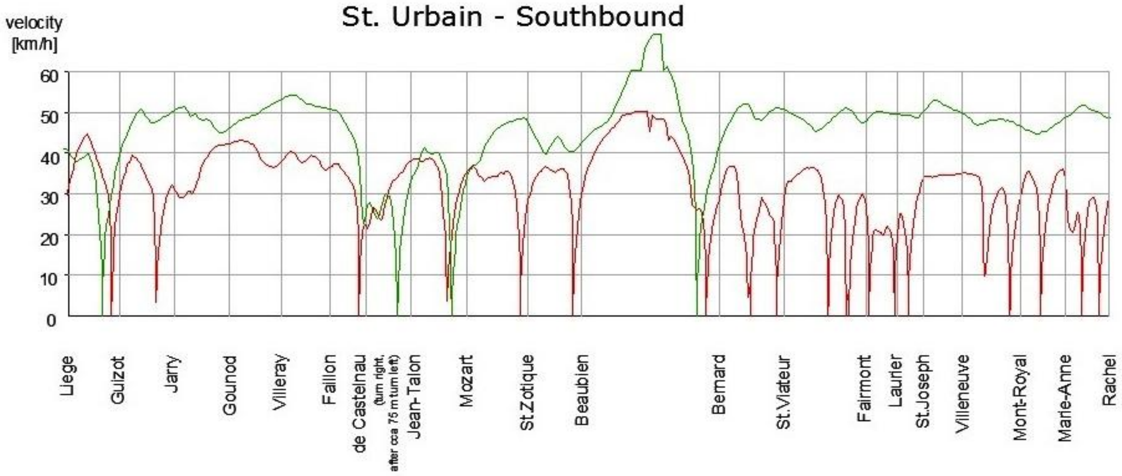


	total TT	time of 0 km/h	average speed
— Mo 02-21-2005 6.32 - 6.41	9 min 31 s	44 s (8% of TT)	34.9 km/h
— Mo 02-21-2005 8.26 - 8.46	19 min 53 s	333 s (27% of TT)	16.7 km/h

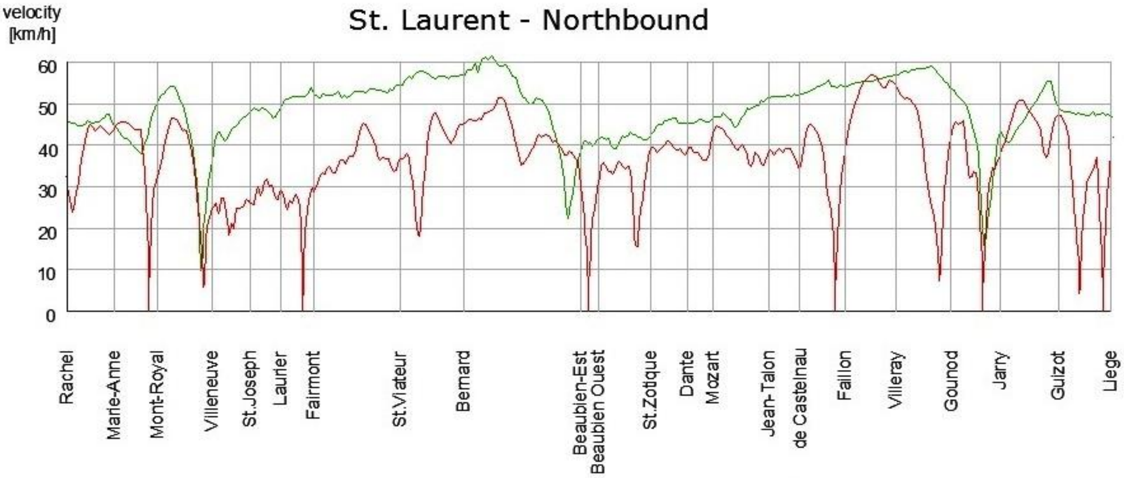
Figure 14: Plots of speed over trajectories (to be continued)



	total TT	time of 0 km/h	average speed
🟩 Tu 02-22-2005 6.50 - 7.01	10 min 22 s	57 s (9% of TT)	23.0 km/h
🟥 Tu 02-22-2005 8.00 - 8.15	14 min 34 s	260 s (30% of TT)	32.3 km/h



	total TT	time of 0 km/h	average speed
🟩 Fr 02-25-2005 6.54 - 7.03	8 min 21 s	55 s (11% of TT)	36.9 km/h
🟥 Fr 02-25-2005 8.23 - 8.40	16 min 16 s	296 s (30% of TT)	18.9 km/h



	total TT	time of 0 km/h	average speed
🟩 We 02-23-2005 6.31 - 6.37	6 min 27 s	0 s (0% of TT)	46.9 km/h
🟥 We 02-23-2005 7.48 - 8.01	12 min 17 s	178 s (24% of TT)	24.6 km/h

Figure 14: Plots of spot speed over trajectories

We can see from the plots above that the fundamental diagram shown in Figure 1 works. The higher the traffic demand is, the slower the flow is. It is a valid statement for links. Furthermore, we can observe the increase of stops during the trips. If we assume that traffic lights are in coordination, it can be explained by too high demand that not allow the flow to reach the speed what the signal timings were counted to. It leads to the state when the vehicle is not able to drive through the arterial road in its green wave and is forced to stop at almost every intersection. This happens mostly in the south part of the solved area where the demand is highest.

4.3.3. Analysis of average speed on links in peak hour

The graphic representation of that phenomena is in Figure 15. The biggest delay is measured around St. Joseph which is an important arterial road crossing the solved area in the east-west direction. Van Horne and Jean-Talon also carry a big volumes, but the effect of the queuing is divided among longer adjacent links. The average delay at intersections are elaborated in Appendix A.

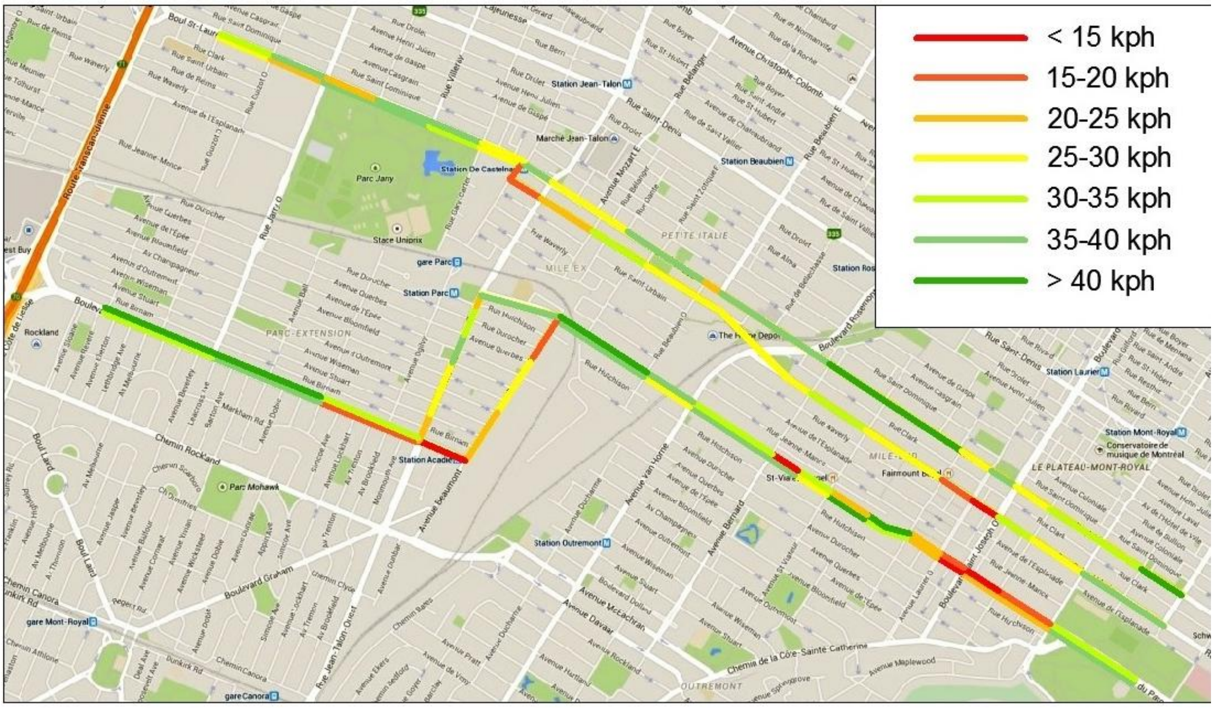


Figure 15: Average speed on links in time 7:30 - 8:30 AM

The Appendix A contains the following information:

Example:

Parc X Mt. Royal

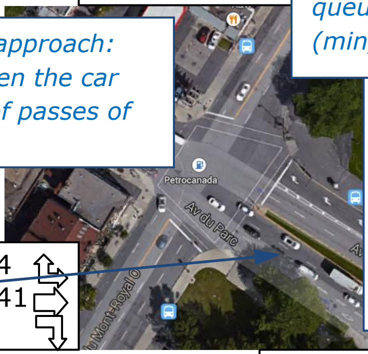
Streets of the intersection.

	# stops		pos. in queue	queuing time
Car stopped	6/9	min	1	24 s
		max	4	45 s
		avg.	2	38 ± 7 s

Analysis of delay from this approach: Position of the test vehicle in the queue and waiting time. (min, max and average values)

Analysis of delay from this approach: ratio number of passing when the car stopped and total number of passes of the test vehicles.

102-1510-1
↲ ⇩



Hourly traffic flows between 7.30 and 8.30 according to the historical database (counted flows).

North and South are data for the trajectory of the test vehicle.

Layout of the intersection.

84 ↰
241 ⇨
9 ↳

Layout of traffic lanes.

3-574-247
⇧ ⇧ ↱

	# stops		pos. in queue	queuing time
Car stopped	5/8	min	1	13 s
		max	10	41 s
		avg.	5	26 ± 10 s

			northbound	southbound
Link to Pieton (crossing)	length		394 m	387 m
	# lanes		2	3
	empirical FFS (kph)		61,9	60,1
	avg. speed (st.d.)		32,8	36,1 ± 17,9
	stops on link		only at ligths	only at ligths
	travel time	min	23 s	18 s
		max	78 s	99 s
		avg.	55 ± 22 s	39 ± 28 s

Analysis of data from the adjacent link:
Empirical FFS: measured as the max recorded average speed on the link (usually between 6.30 and 7:00)
Stops on link: if there was recorded a stop somewhere else but the queue at the intersection, then distance from the stop line ahead and time of that stop is given.
Travel time: includes the queing time at the analyzed intersection. (min, max, average values)

4.4. Application of the model

After the calibration process is completed, we have the accurate information about all incoming traffic flows in 5-minute intervals. It provides us a possibility to evaluate any query about traffic conditions in any time within the simulated one hour period. The traffic characteristics can be evaluated as aggregated values of travel times, mean speed, density, flow on any link in the network. There is also a option to evaluate inquiries of dynamic traffic assignment, such as when is the most convenient time for a vehicle to depart and what values of its indicator is going to be reached.

Prediction can be done rather easily based on the current measured traffic characteristics. In its short-term form it would be used for adaptive traffic signal timings, prediction of occupancy rate of car parks, garages, etc. Long-term predictions are used for planning of investments in road network. These predictions but cannot be done based on few measurements, but require a lot of data to create a repeating patterns which can be extrapolated by mathematical tools.

5. Conclusion

5.1. Summary

The thesis focuses on a description of traffic characteristics and on ways how to collect data for their determination in its theoretical part.

The practical part includes a detailed traffic analysis of an area of three arterial roads in the city of Montreal. A traffic microsimulation model of that area is created in the open source software tool SUMO. Subsequently the model is calibrated based on pseudo-live GPS traces collected by floating cars.

The output of this calibrated model might have plenty of further applications, from demand management to dynamic route guidance for avoiding congested roads. It is true that similar model can be created by using conventionally embedded sensors like loop detectors or camera technology, but those require high cost investments and certain maintenance, while my software solution can be based on using GPS data from vehicles that actually already generate these FCD.

5.2. Future research directions

This topic is certainly not closed with a dead end, but on the contrary there is a huge potential for further development.

The model itself can be particularized by adding diversity of traffic flow including buses of public transportation. Another thing which has been left out in the model is the parallel parking on the side of the arterial roads. Even though vehicles making the parking manoeuvre impact the following vehicles.

This thesis works with pseudo live GPS traces. For a commercial usage there must be developed a script for automatic online collection of FCD.

Beside calibration of demand using GPS data there should be created traffic scenario management analysis. These scenarios might study a change of traffic conditions under different boundary conditions. Speed limit might be changed, different signal timings can be implemented or changed from static to actuated or both together can be studied and evaluated. Indicators on which base the scenarios would be analyzed might still be travel time or queue length can be measured as well.

In summary, the primary theoretical framework for monitoring traffic conditions based on GPS traces has been developed. Its further usage depends on the final user, but I dare to say that there are might arise various possible short-term and long-term applications.

Appendix A - Overview of Analyzed Intersections

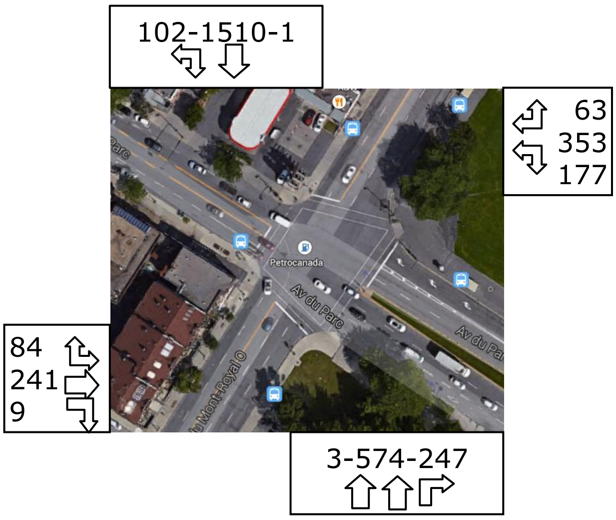
The author has the same analyses also for the following intersections:

- Acadie X Jarry
- Acadie X Beaumont
- Acadie X Jean-Talon
- St. Laurent X Jarry
- St. Laurent X Villeray
- St. Laurent X de Castelnau

These intersections are not given, because they are not in SUMO model. However, they have been a part of the network which was analyzed in a detailed way as is described in Chapter 4.

Parc X Mt. Royal

	# stops		pos. in queue	queuing time
Car stopped	6/9	min	1	24 s
		max	4	45 s
		avg.	2	38 ± 7 s

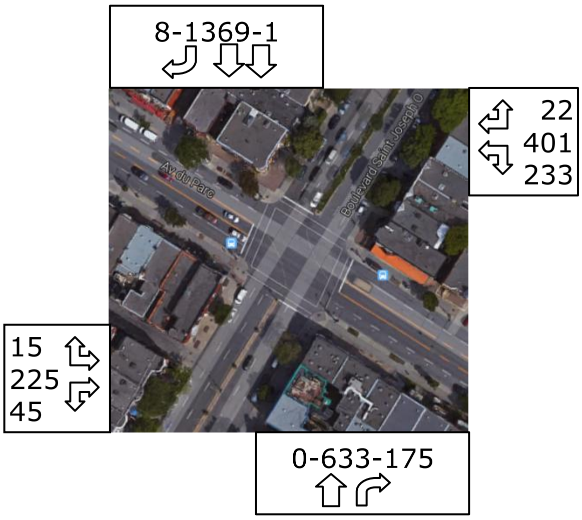


	# stops		pos. in queue	queuing time
Car stopped	5/8	min	1	13 s
		max	10	41 s
		avg.	5	26 ± 10 s

			northbound	southbound
Link to Pieton (crossing)	length		394 m	387 m
	# lanes		2	3
	empirical FFS (kph)		61,9	60,1
	avg. speed (st.d.)		32,8	36,1 ± 17,9
	stops on link		only at ligths	only at ligths
	travel time	min	23 s	18 s
		max	78 s	99 s
		avg.	55 ± 22 s	39 ± 28 s

Parc X St. Joseph

	# stops		pos. in queue	queuing time
Car stopped	8/9	min	1	27 s
		max	13	43 s
		avg.	8	34 ± 5 s



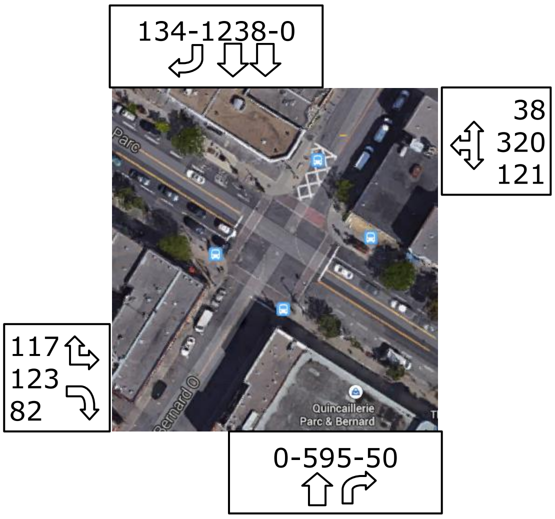
	# stops		pos. in queue	queuing time
Car stopped	8/8	min	1	29 s
		max	15	58 s
		avg.	4	42 ± 9 s

			northbound	southbound
Link to Villeneuve	length		173 m	191 m
	# lanes		1+	1+
	empirical FFS (kph)		45,9	54,0
	avg. speed (st.d.)		8,5 ± 3,4	20,1 ± 2,4
	stops on link		5/8 avg. 126 m for 28 s	only at lights
	travel time	min	44 s	27 s
		max	125 s	43 s
		avg.	87 ± 33 s	35 ± 4 s

Parc X Bernard

			northbound	southbound
Link to Van Horne	length		434 m	435 m
	# lanes		1+	1+
	empirical FFS		50,4	59,9
	avg. speed (st.d.)		30,8 ± 2,5	39,4 ± 10,2
	stops on link		only at lights	only at lights
	travel time	min	44 s	31 s
		max	57 s	93 s
		avg.	51 ± 4 s	44 ± 18 s

	# stops		pos. in queue	queuing time
Car stopped	1/9	min	3	22 s
		max	3	22 s
		avg.	3	22 ± 0 s



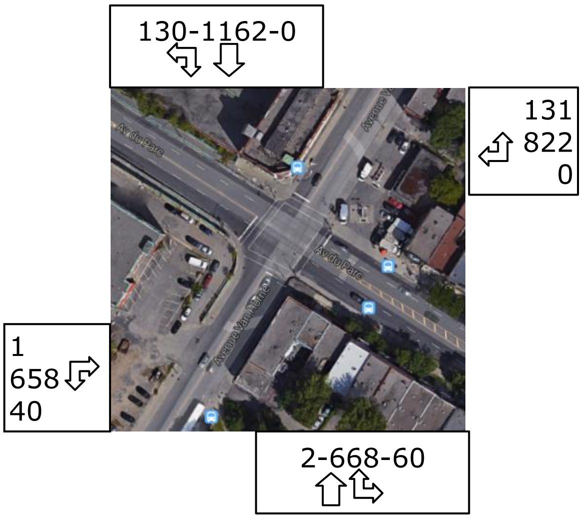
	# stops		pos. in queue	queuing time
Car stopped	6/8	min	1	14 s
		max	17	26 s
		avg.	6	21 ± 5 s

			northbound	southbound
Link to Pieton (crossing)	length		143 m	158 m
	# lanes		1+	1+
	empirical FFS (kph)		49,7	57,8
	avg. speed (st.d.)		13,2 ± 4,1	39,3 ± 7,6
	stops on link		1/8 avg. 114 m for 3 s	only at lights
	travel time	min	24 s	12 s
		max	58 s	23 s
		avg.	41 ± 10 s	15 ± 4 s

Parc X Van Horne

			northbound	southbound
Link to Beaubien	length		256 m	238 m
	# lanes		2	2
	empirical FFS		55,9	70,8
	avg. speed (st.d.)		37,1 ± 14,6	29,5 ± 14,4
	stops on link		only at lights	only at lights
	travel time	min	18 s	16 s
		max	75 s	64 s
		avg.	34 ± 23 s	38 ± 19 s

	# stops		pos. in queue	queuing time
Car stopped	4/9	min	1	26 s
		max	5	49 s
		avg.	2	33 ± 9 s



	# stops		pos. in queue	queuing time
Car stopped	4/8	min	3	4 s
		max	12	6 s
		avg.	6	5 ± 1 s

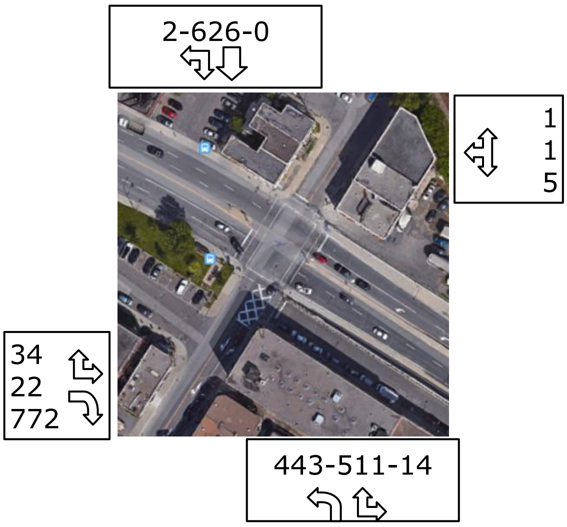
			northbound	southbound
Link to Bernard	length		434 m	191 m
	# lanes		1+	1+
	empirical FFS (kph)		50,4	54,0
	avg. speed (st.d.)		30,8 ± 2,5	20,1 ± 2,4
	stops on link		only at lights	only at lights
	travel time	min	44 s	27 s
		max	57 s	43 s
		avg.	51 ± 4 s	35 ± 4 s

Parc X Beaumont

			northbound
Link to Jean Talon	length		351 m
	# lanes		3
	empirical FFS		49,5
	avg. speed (st.d.)		38,9 ± 7,3
	stops on link		only at lights
	travel time	min	25 s
		max	43 s
		avg.	33 ± 6 s

Link to Querbes		
		southbound
length		208 m
# lanes		1
empirical FFS		42,2
avg. speed (st.d.)		19,3 ± 8,3
stops on link		1/9 avg. 129 m for 3 s
travel time	min	20 s
	max	85 s
	avg.	47 ± 20 s

Car stopped			
# stops		pos. in queue	queuing time
5/9	min	1	4 s
	max	18	22 s
	avg.	8	15 ± 6 s

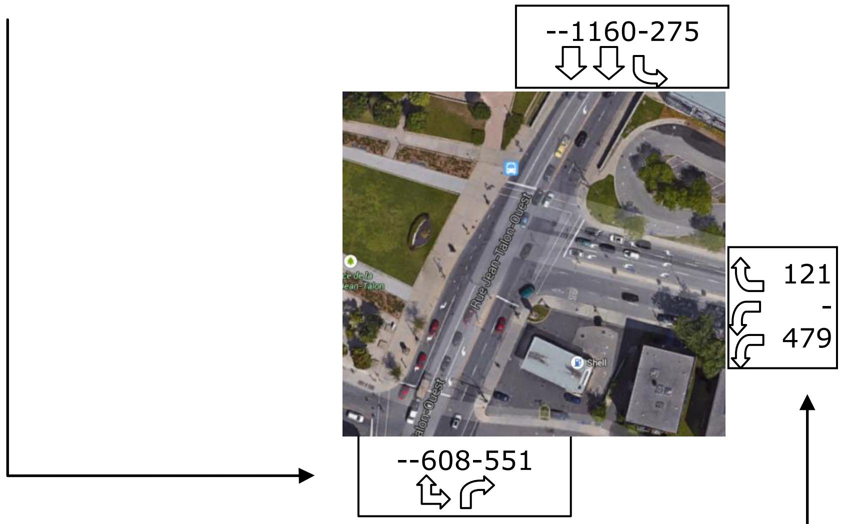


Car stopped	# stops		pos. in queue	queuing time
	2/8	min	2	18 s
		max	8	29 s
		avg.	5	24 ± 6 s

			northbound	southbound
Link to Beaubien	length		464 m	495 m
	# lanes		2	2
	empirical FFS (kph)		60,0	53,2
	avg. speed (st.d.)		44,3 ± 11,3	36,9 ± 8,6
	stops on link		1/8 avg. 204 m for 47 s	only at lights
	travel time	min	28 s	38 s
		max	73 s	93 s
		avg.	41 ± 14 s	52 ± 16 s

Parc X Jean Talon

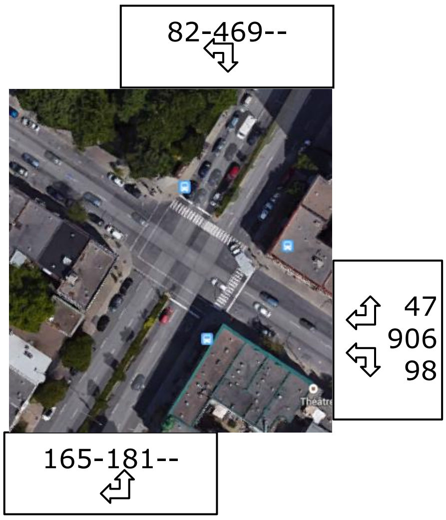
			northbound
Link to Querbes	length		206 m
	# lanes		2
	empirical FFS		35,3
	avg. speed (st.d.)		25,2 ± 7,0
	stops on link		2/8 at Hutchison lights avg. 20 m for 5 s
	travel time	min	21 s
		max	52 s
		avg.	32 ± 10 s



Car stopped	# stops		pos. in queue	queuing time
	3/8	min	1	6 s
		max	9	14 s
		avg.	5	9 ± 4 s

			northbound
Link to Beaumont	length		351 m
	# lanes		3
	empirical FFS (kph)		49,5
	avg. speed (st.d.)		38,9 ± 7,3
	stops on link		only at lights
	travel time	min	25 s
		max	43 s
		avg.	33 ± 6 s

St. Laurent X Mt. Royal

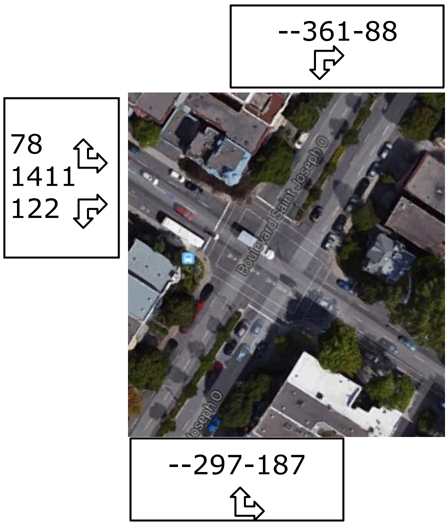


	# stops		pos. in queue	queuing time
Car stopped	2/9	min	1	35 s
		max	1	38 s
		avg.	1	37 ± 2 s

			northbound
Link to Marie-Anne	length		208 m
	# lanes		2
	empirical FFS (kph)		56,8
	avg. speed (st.d.)		31,8 ± 11,5
	stops on link		only at lights
	travel time	min	17 s
		max	62 s
		avg.	29 ± 18 s

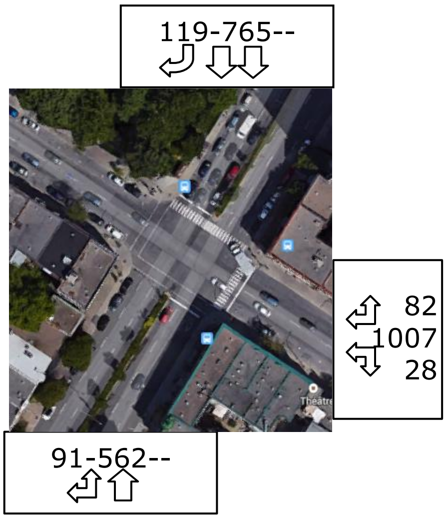
St. Urbain X Mt. Royal

	# stops		pos. in queue	queuing time
Car stopped	3/10	min	5	24 s
		max	8	40 s
		avg.	7	32 ± 7 s



			southbound
Link to Marie-Anne	length		222 m
	# lanes		2
	empirical FFS (kph)		48,6
	avg. speed (st.d.)		36,3 ± 7,5
	stops on link		1/10. 102 m for 5 s
	travel time	min	17 s
		max	35 s
		avg.	23 ± 5 s

St. Laurent X St. Joseph

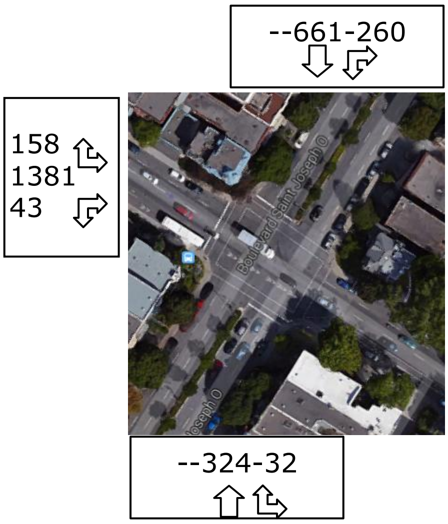


	# stops		pos. in queue	queuing time
Car stopped	2/9	min	1	40 s
		max	14	45 s
		avg.	8	43 ± 3 s

			northbound
Link to Villeneuve	length		185 m
	# lanes		2
	empirical FFS (kph)		45,4
	avg. speed (st.d.)		28,8 ± 12,6 s
	stops on link		1/9 101 m for 3 s
	travel time	min	16 s
		max	80 s
		avg.	33 ± 25 s

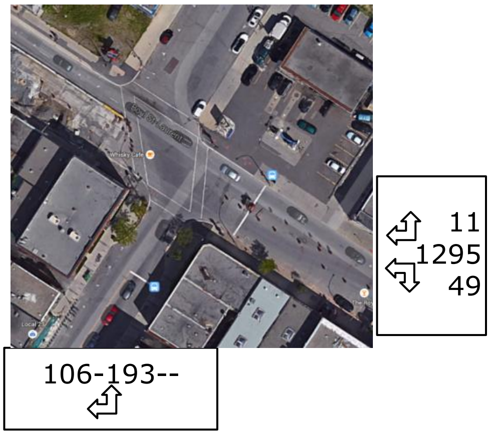
St. Urbain X St. Joseph

	# stops		pos. in queue	queuing time
Car stopped	7/10	min	2	1 s
		max	16	38 s
		avg.	8	25 ± 15 s



			southbound
Link to Villeneuve	length		191 m
	# lanes		2
	empirical FFS (kph)		51,7
	avg. speed (st.d.)		33,8 ± 4,5
	stops on link		none
	travel time	min	17 s
		max	24 s
		avg.	21 ± 2 s

St. Laurent X Bernard

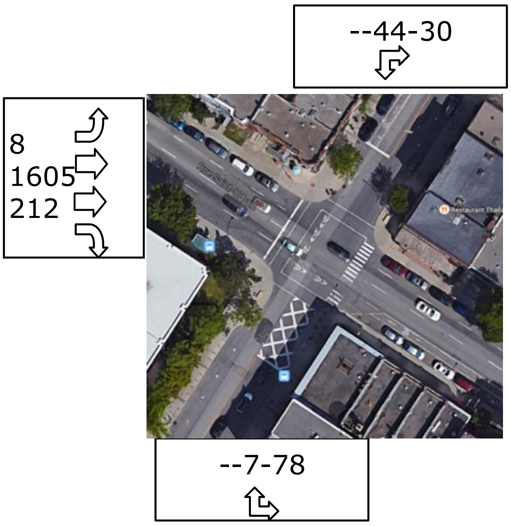


	# stops		pos. in queue	queuing time
Car stopped	0/9	min	-	-
		max	-	-
		avg.	-	-

			northbound
Link to St. Viateur	length		308 m
	# lanes		2
	empirical FFS (kph)		60,3
	avg. speed (st.d.)		46,2 ± 6,5
	stops on link		none
	travel time	min	19 s
		max	30 s
		avg.	25 ± 3 s

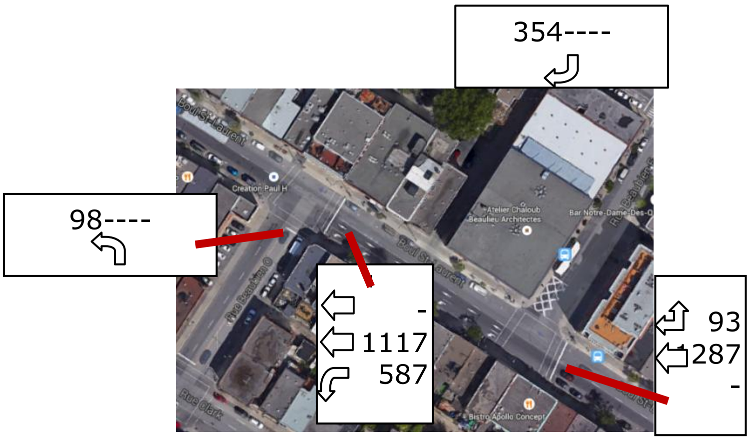
St. Urbain X Bernard

	# stops		pos. in queue	queuing time
Car stopped	10/10	min	3	5 s
		max	17	29 s
		avg.	9	17 ± 9 s



			southbound
Link to St. Viateur	length		320 m
	# lanes		2
	empirical FFS (kph)		49,7
	avg. speed (st.d.)		30,2 ± 13,0
	stops on link		none
	travel time	min	25 s
		max	94 s
		avg.	50 ± 28 s

St. Laurent X Beaubien



North: Car stopped	# stops		pos. in queue	queuing time
	4/9	min	1	31 s
		max	3	36 s
		avg.	2	33 ± 2 s

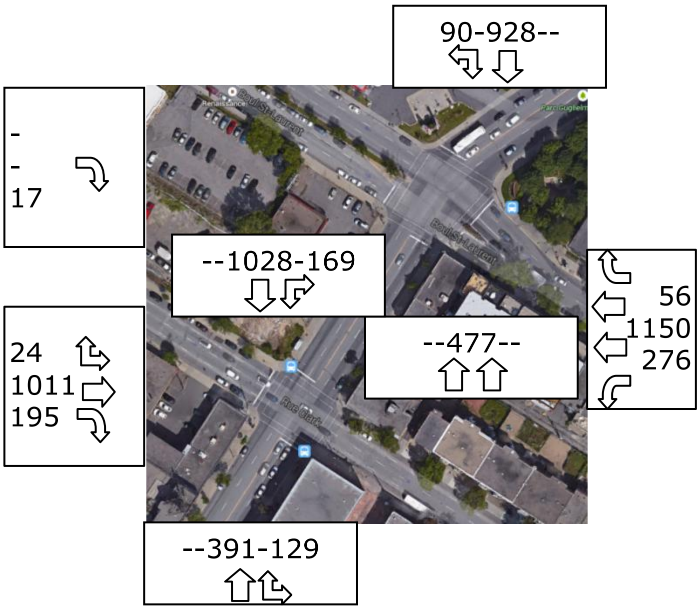
South: Car stopped	# stops		pos. in queue	queuing time
	2/9	min	5	6 s
		max	10	31 s
		avg.	8	19 ± 13 s

Link to Bernard			northbound
	length		566 m
	# lanes		2
	empirical FFS (kph)		50,4
	avg. speed (st.d.)		38,6 ± 8,6
	stops on link		1/9 129 m for 9 s
	travel time	min	41 s
		max	96 s
		avg.	56 ± 16 s

Jean Talon X St. Laurent X Clark

			northbound (St. Laurent)	southbound (Clark)
Link to de Castelnau	length		153 m	221 m
	# lanes		2	3
	empirical FFS		55,0	30,8
	avg. speed (st.d.)		39,7 ± 3,8	17,2 ± 7,8
	stops on link		only at lights	none
	travel time	min	13 s	26 s
		max	17 s	83 s
		avg.	14 ± 1 s	55 ± 19 s

	# stops		pos. in queue	queuing time
Car stopped	7/10	min	1	8 s
		max	9	43 s
		avg.	3	29 ± 10 s



	# stops		pos. in queue	queuing time
Car stopped	6/9	min	1	3 s
		max	10	16 s
		avg.	5	8 ± 5 s

			northbound (St. Laurent)	southbound (Clark)
Link to Mozart	length		270 m	273 m
	# lanes		2	2
	empirical FFS (kph)		57,2	26,9
	avg. speed (st.d.)		27,8 ± 6,1	23,9 ± 2,0
	stops on link		only at lights	only at lights
	travel time	min	25 s	37 s
		max	49 s	47 s
		avg.	37 ± 7 s	42 ± 4 s

Appendix B - Photos of the Area

Avenue du Parc

Intersection Avenue du Parc X Van Horne

Example of minor road in West-East direction (here St. Viateur)

Rue Laurier - Crossing Blvd. St. Urbain

Queue at NB approach of St. Laurent at the intersection St. Joseph

Blvd. St. Laurent in its northern part (around Av. Mozart)

Bibliography

Balakrishna, R. (2006). *Off-line Calibration of Dynamic Traffic Assignment Models* (Thesis). MIT, MIT.

Ben-Akiva, M., Bierlaire, M., Burton, D., Kousopoulos, H., & Mishalani, R. (2001). Network State Estimation and Prediction for Real-Time Traffic Management. *Networks and Spatial Economics*(1).

Brotcorne, L., De Wolf, D., Gendrau, M., & Labbé, M. (2000). A Dynmic User Equilibrium Model for Traffic Assignment in Urban Areas. *CiteSeer*(2). Retrieved from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.13.8799&rep=rep1&type=pdf

Ekström, J., Olstam, J., & Tapani, A. Linköpings universitet. (2013). *Dynamic Traffic Models*.

Cheu, R. (2014). *Traffic Engineering: handouts*. University of Texas at El Paso.

Cheu, R. (2013). *Transportation Engineering: handouts*. University of Texas at El Paso.

Comptage des véhicules et piétons: Flux de données au format KML. Retrieved from: http://donnees.ville.montreal.qc.ca/storage/f/2013-10-06T19%3A05%3A42.653Z/comptage-vehicules-pietons.kml

Dynamic Traffic Models. (2013). In: Ekström, J., Olstam, J., & Tapani, A. *Dynamic Traffic Models*. Linköpings universitet. Retrieved from: http://staffwww.itn.liu.se/~joaek/TNK062/dta_4in1.pdf

Huang, H., & H.K.Lam, W. (n.d.) *Modeling and*.

Huang, H., & Lam, W. Modeling and solving the dynamic user equilibrium route and departure time choice problem in network with queues. Retrieved from: http://semen.buaa.edu.cn/Teacher/papers/Huang-Lam-TDA.pdf

Jenelius, E., & Koutsopoulos, H. (2013). Travel time estimation for urban road networks using low frequency probe vehicle data. *Transportation Research: Part B*(53).

Koukol, M. (2013). *Traffic Microsimualtions: handouts*. Prague.

Krauss, S. (1997). *Microscopic Modeling of Traffic Flow: Investigation of Collision Free Vehicle Dynamics* (PhD Thesis). University of Cologne, Köln.

Leksono, C., & Andriyana, T. (2012). *Roundabout Microsimulation Using SUMO* (Thesis). Linköping s Universitet, Linköping.

Lu, L., Xu, Y., Antoiniou, C., & Ben-Akiva, M. (2013). An enhanced SPSA algorithm for the calibration of Dynmic Traffic Assignment models. *Transportation Research: Part C*(1).

Mannering, F., & Washburn, S. (2012). *Principles of Highway Engineering and Traffic Analysis*. (5th Edition). USA: J. Wiley.

Montreal - Census. (8 January 2015). Retrieved from: http://www12.statcan.gc.ca/census-recensement/2011/as-sa/fogs-spg/Facts-cma-eng.cfm?LANG=Eng&GK=CMA&GC=462

Přibyl, P. (2014). *Traffic Flow Theory: handouts*. CTU in Prague.

Ross, S. (c1997). *Simulation*. (2nd ed., xii, 282 p.) San Diego: Academic Press.

Sumo. (2015). Retrieved from: http://sumo.dlr.de/wiki/Main_Page

Traffic Simulation. (2001-). *Wikipedia: the free encyclopedia* [online]. Retrieved from: http://en.wikipedia.org/wiki/Traffic_simulation

Transportation Research Circular: Dynamic Traffic Assignment. (2011). (A Primer). Retrieved from: http://onlinepubs.trb.org/onlinepubs/circulars/ec153.pdf

Vaze, V. (2007). *Calibration of Dynamic Traffic Assignment Models with Point-to-Point Traffic Survaeillance* (Thesis). MIT, MIT.

Wikipedia: the free encyclopedia [online].