Macroscopic Fundamental Diagrams for Stockholm Using FCD Data

A Master’s Thesis by

Gao Feng

Feb 2011

Royal Institute of Technology (KTH)
Division of Traffic and Logistics
Abstract

Macroscopic fundamental diagrams (MFD) reveal the relations among flow, speed and density in a large geographic region. After literature review on macroscopic analysis, the similar methodology is applied in this thesis. The purpose of the thesis is to find the evidence that is able to prove MFD existing in Stockholm urban region. Both floating car data (FCD) based on global positioning system (GPS) data from taxis travelling in Stockholm region and traffic data from fixed detectors data source are used to construct the fundamental diagrams. Geographically, the usage of data is extended from single link to multiple links, then to the entire study region. The temporal phase is restricted in one weekday and weekend.

The diagram of flow vs. speed based on single detector is found disordered, by contrast, the diagrams of cumulative flow, speed and density for all detectors represent orderly. MFD diagrams proposed in Yokohama case study by Geroliminis and Daganzo are reproduced with cumulative data in this thesis. Therefore, it can be proved that MFD exists when using data from multiple links. However, the cumulative data from fixed detectors only represents the traffic on links where they locate, not the entire region. To overcome it, GPS data from taxis, which covers the whole region, is analyzed with same method. Because full taxis travel in the same manner as normal vehicles, they are selected to approximate traffic in whole region. A neat curve of flow vs. speed is produced and it coincides with corresponding diagram in the reference paper. It enhances the conclusion that MFD exists in the entire study region. Moreover, based on the constant ratio between average link flow and region exit flow, a controlling density policy is discussed in aiming for maximizing trip completion.

Key words: macroscopic fundamental diagram, FCD data, taxis GPS data
Acknowledgements

There are a lot of persons the author should express gratitude to. First of all, I would like to thank my supervisor Albania Nissan and Mahmood and examiner Haris for all the helpful advices and time you devote into this thesis. Without your guidance and assist, I could not decipher the difficulties and focus on the right way. This thesis would not have the possibility to complete.

The next people to thank are Tobias Johansson and Christina Akbar who work in the city of Stockholm Traffic Administration as strategic planning. They kindly give me a hand in the data source preparation and support author with the fixed detectors traffic data. In addition, they are always patient to answer me any problems.

Besides, I desire to thank all the references writers, particularly Geroliminis and Daganzo. I would also like to thank every person who reads this paper and contributes information to this thesis.

Finally, I want to thank my parents and friends, you always stand on my side, no matter what I choose. Without your selfless support, this thesis would not have been as good as it is.

Thank you all!
Figure Lists

1-1 Flow Density and Speed Curves (Greenshields, 1934) .......................... 1
1-3 Average Number of Vehicles across the Cordon (Stockholm Trafikkontoret, 2009) ................................................................ 5
1-4 Progressive Analysis from Single to Multiple then to Entire Region ..... 6
2-1 Greenshields Measurements at a Point near Highway (Greenshields, 1934) ................................................................. 9
2-2 Speed and Density Linear Relation (Greenshields in 1934) .................. 10
2-3 Speed and Traffic Flow Relation (Greenshields in 1934) ...................... 10
2-4 Simulation Results in a Closed CBD-Type Street Network (Williams in 1987) ................................................................. 12
3-1 GPS Data Transfer Flow .................................................................. 17
3-2 Study Region same as the Congestion Charge System
(http://www.transportstyrelsen.se) ......................................................... 20
3-3 Fixed Detectors Data Process Flow .................................................. 22
3-4 Taxis GPS Data Process Flow .......................................................... 23
4-1 All the fixed Data Collection Detectors in Stockholm city (By Google Earth) .................................................................. 25
4-2 All the Fixed Data Collection Detectors in the Study Region (By Google Earth) ................................................................. 26
4-3 Two Single Detectors Location in the Map ......................................... 28
4-4 Flow vs. Speed of Two Single Detectors in One Day .......................... 29
4-5 Average Flow Fluctuates along Time of One Weekday and Weekend Day ..... 29
4-6 Aggregate Time Mean Speed and Space Mean Speed in One Day .......... 30
4-7 Unweighted Flow vs. Speed Aggregated in One Day ......................... 31
4-8 Unweighted Flow vs. Speed of Yokohama by Geroliminis in 2007 ....... 31
4-9 Unweighted Flow vs. Density in Yokohama by Geroliminis in 2007 ...... 31
4-10 Unweighted Flow vs. Density Aggregated in One Day ..................... 31
4-11 Weighted Flow vs. Density in Yokohama by Geroliminis in 2007 ...... 32
4-12 Weighted Flow vs. Density Aggregated in One Day ....................... 32
4-13 Weighted Flow vs. Space Mean Speed of One Day ......................... 32
4-14 Detector 513 Flow vs. Speed for Five Weekdays ............................. 33
4-15 Exit Flow and Weighted Flow Fluctuate along Time in One Day ....... 35
4-16 Ratio between Exist Flow and Weighted Flow in One Day ............... 35
4-17 Individual Taxi Trace in One Weekday (By Google Earth) ................. 37
4-18 Taxis Flow Estimation Process ...................................................... 38
4-19 Taxis Move in and Out Ratio to All Vehicles in One Day ................... 39
4-20 All Vehicle Flow Fluctuate in One Day .......................................... 42
4-22 All Region Vehicles Number of Vehicles vs. Speed in One Day Estimated from Taxis GPS Data ................................................... 43
4-21 All Region Number of Vehicles vs. Speed in Yokohama by Geroliminis in 2007 ................................................................. 43
Table Lists

2-1 SIMILARITY BETWEEN GEROLIMINIS AND DAGANZO (2007) YOKOHAMA CASE AND THIS THESIS STOCKHOLM CASE ................................................................. 13
3-1 STOCKHOLM CONGESTION CHARGE SYSTEM TOLLS DATA SAMPLE .................. 17
3-2 STOCKHOLM FIXED DETECTORS DATA SAMPL E ........................................ 17
3-3 TAXIS GPS DATA SAMPLE ............................................................................... 18
3-4 CONGESTION CHARGE SYSTEM 18 TOLLS DETAILS ................................... 21
4-1 STATISTICAL TEST FOR IN RATIO WITH TIME T ........................................... 40
1 Introduction

1.1 Motivation and Overview

The urban transport plays a significant role in our daily life. Each day, people participate in the traffic system and spend some time in personal vehicles or public transport. However, the increasing of population and vehicles aggravate the traffic system burden and prolong the private travel time. With the consideration of improving efficiency, developing optimal road network and relieving congestion, the traffic flow theory is comprehensively and thoroughly investigated in past few decades. According to different observation scales, the traffic theory study could be divided into macroscopic, mesoscopic and microscopic. This thesis attempts to make some researches in macroscopic fundamental diagrams (MFD). Macroscopic traffic flow theory aggregates the individual behavior and analogizes the traffic to a stream. It is usually used into network links and region research. The fundamental diagrams refer to the traditional relation diagrams among traffic flow, space mean speed and density as shown in figure 1-1.

![Flow Density and Speed Curves (Greenshields, 1934)](image)
There are some successful experiences in how to make the macroscopic analysis. For instance, Greenshields (1934) collected traffic data at the side of highway by a camera and built the first fundamental diagram with conjecture of the linear relation between speed and density. Thomson (1967) found linear decreasing relation between average speed and flow with collection many years’ data in central London. Wardrop (1968) inducted the street width and intersection spacing in the flow speed relation. Nevertheless, the earlier researches generally detected the monotonic change relation that only suit for the non congestion state. Afterwards, Herman and Prigogine (1979) proposed two-fluid model that defined the cars as stop parts and moving parts. This new model captured more characteristics in heavy burden traffic condition.

While the global positioning system (GPS) is used as the new traffic data collection technology, real-time data in large scale is available for dynamical rush hour description. The operational principle of GPS is simply stated as there is a GPS receiver in the vehicles that calculates current time position by receiving GPS satellites signals. Geroliminis and Daganzo (2007) reveal that MFD exists in Yokohama (Japan) urban region with the help of fixed detectors data and GPS data.

Based on the Geroliminis and Daganzo’s (2007) research, this thesis would like to make similar analysis of MFD in Stockholm urban region. The research objective variables are derived from two kinds of data sources. One is through the fixed detectors. It supplies us abundant raw data such as vehicle numbers and average fleet speed. These records are measured by inductive loops and microwave beams at short road sections distributing around the whole Stockholm urban region. The other data source is GPS data source, what is new, comes from the taxis equipped with GPS equipments, which trace taxis movements in the comparable region of fixed detectors part. In addition, it envelops wider region and more details, for example, the individual taxis’ id, taxis status, and location. Due to GPS data is collected in the more realistic way, comparing to the former data source, this estimation result is closer to actual condition.
1.2 Traffic Condition Background

1.2.1 Other Countries’ traffic condition

While the urban population grows fast and number of automobiles increases quickly, residents in most of cities, particularly the big cities, are confronted with more serious congestion problem than before. For instance, U.S 2010 urban mobility report (David, et al. 2010) describes the scope of the mobility problem. It calculates the congestion performance for 439 U.S urban roadway section with the help of data sources from Highway Performance Monitoring System (HPMS) that provides traffic volume and road network data. In addition, the real time speed data is offered by INRIX, which is a traffic services company providing historical and real-time traffic information in United States and Europe. According to this report, the congestion cost was 24 billion dollar for 439 U.S urban regions in 1982. This congestion cost increased to 85 billion in 2000. Even worse, this congestion cost increased to 115 billion in 2009. This 115 billion congestion cost in 2009 is composed of 3.9 billion gallons of wasted fuel and 4.8 billion hours of extra time. In addition, this report points out that congestion is worse in regions of every size, but the larger cities endure longer hours of delay.

U.S 439 Cities' Congestion Cost

![U.S 439 Cities' Congestion Cost](chart.png)

*1-2 U.S Congestion Cost Change from 1982 to 2009(Urban Mobility Report 2010)*
Parallel situation also occurs in Europe. Since 2008, INRIX started to collect traffic data and built evaluation report for some European countries. INRIX required billions of GPS probe vehicles and mobile devices travelling Europe reporting the traffic information, such as speed, location and time with particular anonymous head. They developed efficient models for interpreting probe vehicle reports and established estimation of travel patterns in major cities. For instance, in 2010, INRIX made scorecard results for United Kingdom, Belgium, Germany, Luxembourg, France and Netherland. From this report, within these countries, the most congestion city usually is the capital, which is the biggest city in the country as well, except the Ruhrgebiet, which is traditional industrial district in Germany. Among these six countries, the most congestion city is Paris and drivers in Paris average waste 70 hours per year in traffic. The most congestion bottleneck on French roads takes about 52% longer than the average travel time.

### 1.2.2 Stockholm Traffic condition and policies

Stockholm, as the capital and the largest city of Sweden, is facing serious congestion pressure as well. There are 2,034,480 inhabitants in the Stockholm region about 6,488 square km and 837,031 in the city of Stockholm city (2010). There are 560,000 vehicles crossing the inner cordon per working day. Given that the car ownership increases 2.5% per year, the urban traffic system is overwhelmed by the increasing travel demand. As estimation, there was 600-800 million € congestion cost per year in 2005. What is worse, there are about 10-100 cases of cancer caused by atmospheric pollution and 50,000 inhabitants are exposed to over 65 DBA noise (Gunilla, 2008).

Due to this serious congestion dilemma, the Stockholm municipal has tried a number of successful methods to progress the urban traffic service level. One of the most famous policies is the congestion charge system. The congestion charge system trail was firstly implemented in 2006 for six months with charge of 10 or 20 Kronor at different time phase between 06:30 and 18:30 in and out of the boundary. Then, the permanent congestion tax became effective from august 2007. The trial in 2006 reduced 22% vehicle numbers matched up to 2005. In 2008, the
drop of the traffic flow across the cordon was 18% compared to 2005 (Stockholm Trafikkontoret, 2009) as shown in figure 1-3. The decline of the traffic flow not only decreased the vehicle numbers. It also improved the air quality by reducing the emissions of nitrogen oxides around 8 percent and fossil fuel carbon dioxide about 4 percent between 2006 and 2008.

Besides the congestion tax, Stockholm municipal also implemented the Motorway Control System (MCS). The north part of E4 about 8 km was implemented MCS in 1996 and south part about 12 km in 2004. MCS suggests speed dynamically by automatic detecting incidents. The evaluation indicates that the MCS system enhances the traffic safety by improving travel speed homogeneity and reducing the lane changes frequency between the middle lane and left lane (Nissan et al. 2006)

Additionally, with the consideration of safe transport services, Stockholm city council planned Intelligent Speed Adaptation (ISA) system since 2003 (ISA in Stockholm, 2005). ISA is a system to help drivers at the correct speed. It utilizes GPS equipments to decide the accurate location of the moving vehicles. There is a computer screen that provides synchronized speed limits and warms to the drivers. There was a trial project from March 2003 to September 2005. According to the
evaluation report, ISA reduces average excess over speed limit distinctly under higher speed limitation roads. After the trail, 75% of users want to continue using ISA. Based on these results, Stockholm Road Administration aims to equip all the city’s vehicles with ISA by 2010.

In brief, there are several active traffic management and control policies in Stockholm. These strategies alleviate the increasing automobile stress on the urban traffic. But most of the evaluation reports are from the inhabitants’ aspects and focus on the travel time, safety, healthy, efficiency and equality. This thesis attempts to explore the urban region from macroscopic and fundamental relations of flow, speed and density.

1.3 Thesis Purpose and Progressive Analysis

In view of fact that macroscopic fundamental diagrams (MFD) have been effectively studied and verified in some prior works, this thesis intends to find out the evidence that could proves the MFD exists in the Stockholm prescript region by making use of two kinds of new data sources. The demonstration process is relied on construction and comparison of traffic flow, speed and density relation diagrams.

![Diagram illustrating progressive analysis from single to multiple then to entire region]
The progressive analysis is followed by from easier to harder pattern. Specially, it begins from individual link, multiple links network to the entire study region as shown in figure 1-4. The links variables are estimated from fixed detectors data source which is offered by Stockholm traffic department and Royal Institute of Technology (KTH) traffic and logistics division. The whole region variables are deduced from taxis GPS data source that is provided by KTH logistics division. Pair of flow, speed and density are plotted and compared with reference diagrams. The neat and similar curves testify non-congestion part of MFD established in Stockholm urban region. At last, with the conjecture of whole MFD existence, it proposes one application in maximizing trip completion by controlling of the region density.

1.4 Thesis Organization

The other segments of this thesis will be organized as follow:

Chapter 2

Review some of relevant literatures. It includes the significant papers that make important progress in traffic flow theory development and the major reference paper, which is issued by Geroliminis and Daganzo’s (2007).

Chapter 3

State the methodology theory of this thesis. It contains the traffic models selection, traffic data sources description, data reliability analysis. In addition, the study region delimitation and reasons why this is chosen as study region are stated in the section. Two different data process charts are shown and explained at last.

Chapter 4

Most important section, it gives meticulous verification process of the non-congestion part of MFD existence by analyzing both the fixed detectors data and taxis GPS data. Furthermore, it suggests one simple urban traffic density control application.
Chapter 5

Summarize the results of this thesis. It concludes that non-congestion part of MFD exists in the Stockholm prescribed region. Besides, it points out the limitations of paper and indicates future work orientation.

All the references papers are listed in the bibliography. Finally the compulsory codes and calculation procedure are attached in the appendix.
2 Literature review

Earlier works of the flow theory provide the essential frame of reference. Greenshield (1934) observed 100 vehicle groups each of which contain 10 vehicles in one direction on a two lanes two way road. A camera was used to take pictures by the side of the road as shown in figure 2-1. With the observation data and basic relation

\[
\text{Flow} = \text{density} \times \text{speed}
\]

He postulated linear decline relationship between speed and density as shown in figure 2-2. Then the speed-flow model with parabolic shape is deduced from above relations as shown in figure 2-3. Since the term “flow” has not been defined in his period, he called it “Density-vehicles per hour”. Even though there were some problems, such as the research road was not freeway and the collection time was holiday, this parabolic shape curve was accepted as the proper shape for decades.

2-1Greenshields Measurements at a Point near Highway (Greenshields, 1934)
In addition, Drake et al. (1967) fitted new speed-density curve with data from Chicago expressway. The measured data contained number of vehicles, time mean...
speed and occupancy for 1224 data points. Density was calculated from volume and time mean speed. They fitted the speed-density function and then transformed this function into speed-flow function. They compared the seven speed-density hypotheses statistically and chose the best fit one. As a result, Edie’s (1961) discontinuous exponential form presented the best estimation of the fundamental parameters. Because minor changes in the speed-density function led to major changes in the speed-flow function, the original speed-flow data did not fit very well.

Another weighty research in macroscopic flow and speed relations was made by Thomson (1967). He used the vehicles average speed circulating through central London on designed route every two years of sum of 14 years period. Average flows were firstly converted measured link flows into equivalent passenger car units, then weighted by their respective link lengths. Linear models of speed and flow were constructed for each for eight years. In 1968, Wardrop added street width and stopped time in the average speed to modify this linear model.

Prigogine and Herman (1979) proposed two-fluid model in the multilane highway traffic kinetic theory. In this case, cars were divided into two parts. One part was moving cars. And the other was stopped cars. But the parking vehicles were not considered as stopped cars. The two-fluid model provided a method to evaluate the traffic service quality in macroscopic view. While the computer simulation technology was used in the traffic flow theory, Williams et al. (1987) made the simulation in the closed Central Business District (CBD) type street network based on a postulated relationship between the average fraction of vehicles stopped and the network concentration from the two-fluid theory. He derived fundamental variables models from these simulation results as shown in figure 2-4.
2-4 Simulation Results in a Closed CBD-type Street Network (Williams in 1987)

Review of speed, flow and density curves’ historical developments is instructive to current understanding. Recently, Geroliminis and Daganzo (2007) found macroscopic fundamental diagram (MFD) existing evidence in a field experiment of Yokohama (Japan). Yokohama which has 3.6 million populations is the second largest city in Japan. The study region was characterized as 10 square km triangle region in the center of Yokohama.

The data was from two different sources. One was 500 fixed detectors that offer vehicle accounts and occupancy measurements. The other was 140 taxis as probes that equipped with GPS. The analysis of data was designed in three stages. First was the individual link detector. Second was all the links with detectors. At last, it came to the whole region. In the detector data analysis part, one weekday sample in 5 minutes intervals was elected for analysis. The parameters, such as average flow and weighted flow, were based on Thomson (1967) data analysis of the central London. And the average space mean effective vehicle length was set as 5.5 meters in the transform calculation between density and occupancy. The results of first two stages showed that traffic variables such as flow, speed and density performed more rational relation in urban scale than for individual link. The smooth and small deviation curves in diagrams demonstrated that the link
network located detectors existed MFD. However, due to the detectors only measured their propinquity parts; this result did not represent the whole region.

In the second part, the analysis of taxis GPS data which covers the entire region would help to achieve the final goal. Total 140 taxis provided GPS data logger of position, status, distance, brake and turn movements. The estimation of new data was based on the assumption that full taxis were used as an unbiased estimator of the normal vehicles. Given that the taxis data did not contain ‘meter on’ or ‘free’ tags, the full taxis were selected out by imitating the feature moves such as longer travel distance and less circuitous routes. In order to assess this scale factor, 10 taxis were traced manually for 24 hours. Then, author calculated the number of full taxis in the study region with total travel time and time slice. All cars flow was inferred by scaling up full taxis flow. Finally, all cars flow and average speed diagram showed that MFD existed in whole Yokohama research region.

All these above papers prepared solid theoretical foundation for this thesis. Because some similarities as shown in the table 2-1, the study of Yokohama case made by Geroliminis and Daganzo (2007) is chosen as the major reference.

<table>
<thead>
<tr>
<th>Research Goal</th>
<th>Yokohama Case</th>
<th>Stockholm Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Available Data source</td>
<td>500 fixed detectors</td>
<td>192 fixed detectors</td>
</tr>
<tr>
<td></td>
<td>140 taxis with GPS</td>
<td>About 500 taxis with GPS</td>
</tr>
</tbody>
</table>

2-1 Similarity between Geroliminis and Daganzo (2007) Yokohama Case and this thesis

Stockholm case
3 Methodology

3.1 Traffic modeling

Conventionally, models are classified into microscopic models, macroscopic models and mesoscopic models according to the level of research details.

Microscope models are used to simulate the state of individual vehicles. They comprise the records of individual vehicle motions, such as acceleration, deceleration and the drivers’ behavior such as car following, lane changing. Their analysis result is useful to comprehend the traffic at a detailed level. The increasing interests in the microscope models are due to two reasons. One is the requirement for prediction of dynamic traffic condition and extremely unpredictable drivers’ behavior. The second is that, as the powerful computers utilizations in the transport simulation process, precise output which means high calculation costs in the past is attained in more efficient way. However, this kind of model is replied on the massive data input from in Origin/Destination (OD) matrix or the vehicle movements at intersection. The small inaccuracy in the input data leads expected errors in results.

Mesoscopic models are the properties combination of both microscopic and macroscopic models. The simulated objects are individual vehicles, but mesoscopic models describe their activities and interactions on aggregate relations. Typically, these models are used to evaluate the traveler information system.

The macroscopic models are aggregated of the activities performed in the microscopic models. Cumulative traffic stream characteristics such as flow, speed and density and their relations to each other are commonly employed to study in macroscopic models. These models are applied to large scale networks such as freeways, corridors and rural highways rather than individual vehicles. They can be used to evaluate the spatial extent of congestion caused by traffic demand in urban scale.
With aim of analysis of the whole Stockholm urban region, macroscopic models are chosen in this thesis. Since macroscopic models need input variables (flow, speed and density) and these variables are derived from different data sources, so it is necessary to assess data sources reliability. The data sources description and reliability assessment are explained in the following part.

### 3.2 Data sources

Based on different sources, the data is usually divided into primary data and secondary data. The primary data is observed or collected from firsthand experience; in contrast, secondary data is from existing source and gathered by others. Because both the fixed detectors data and taxis GPS data employed in this thesis are obtained from other parts, therefore these data sources are belong to secondary data.

The fixed detectors data source comes from *city of Stockholm traffic administration*. In terms of collection system, it can be divided into two types. One is from congestion charge system. The other is from links’ fixed detectors system.

Data from congestion charge system contains 18 tolls at the congestion charge region border. Each toll captures every passing vehicle’s information without stopping it by making use of laser beams and camera technology. There is lots of information such as individual vehicle license plate, time, date and tax amount recorded in the database. However, not all the information is required in this thesis. According to the demand of macroscopic models, the aggregate numbers of vehicles are exacted out for further study. One piece of data sample is shown in table 3-1. It contains the stations’ id where the first number is the toll’s location and second number means direction (1 is in, 2 is out). Date period spans one month (October, 2009). The aggregation interval is 30 minutes. For example, this sample indicates that there are 85 vehicles travelling in the region during the time period from 01:00 to 01:30 on 1st Oct 2009.
The other data is from links’ fixed detectors system. There are total about 192 detectors distributions around the big Stockholm region. These detectors include both inductive loops and microwave beams, which record the number of vehicle, type of vehicle, and average speed for both directions. Based on the requirement of macroscopic model analysis, some essential data are picked up for estimation. One piece of fixed detector data sample is shown in table 3-2. It contains the detectors’ id that means different location. Date period lasts for three months (September, October and November, 2009). The aggregation time interval is 15 minutes. R1 means volume of vehicles in one direction and R2 is opposite direction. R1 speed is the vehicle average time mean speed.

<table>
<thead>
<tr>
<th>ID</th>
<th>Date</th>
<th>Time</th>
<th>R1</th>
<th>R2</th>
<th>R1 Speed</th>
<th>R2 Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2009-9-17</td>
<td>00:15</td>
<td>121</td>
<td>64</td>
<td>62.4</td>
<td>59.3</td>
</tr>
</tbody>
</table>

Taxis GPS data is offered by Royal Institute of Technology (KTH) Traffic and Logistics Division. Global positioning system technology is grouped into Intelligent Transportation System (ITS) method. ITS data collection method entails the use of data interactions between specially equipped vehicles and a central system. In this thesis, 500 operation taxis equipped with GPS travel around the Stockholm city. For each taxi, it transmits a GPS probe record every 120 seconds that includes taxi identification, location and status. A data broker receives all these GPS data information and transferred them to KTH in packet. The data transmission flow is illustrated more visually in figure 3-1 as follows:
One piece of GPS data sample is shown in table 3-3. First element is ID and it is accumulated when the record receiving. The vehicle_id is unique and identify the different taxi. Pos_N means the longitude value. As well, Pos_E is the latitude value. Status indicates taxis carry on passengers or not. Measure_time shows when this piece of information sending out from taxi. For each taxi, it sends out one piece of data approximately every 2 minutes. Therefore, this data source accumulates millions pieces of data in one month.

<table>
<thead>
<tr>
<th>ID</th>
<th>Vehicle_id</th>
<th>Pos_N</th>
<th>Pos_E</th>
<th>Status</th>
<th>Measure_time</th>
</tr>
</thead>
<tbody>
<tr>
<td>34003180</td>
<td>11571</td>
<td>59.30546</td>
<td>18.077938</td>
<td>Free</td>
<td>2010-04-11 23:59:49</td>
</tr>
</tbody>
</table>

3.3 Data Reliability

Since the accurate estimation of variables is depended on the data reliability, therefore, it is obligatory to evaluate the data sources’ reliability before the data manipulation. There are two kinds of data sources, thus the reliability is assessed separately.

First of all, the assessment starts from data loss fraction in fixed detectors data source. It misses few flow and speed records in some detectors because of the sensors’ malfunction or the collection omission. For instance, fixed detector ‘11009’ loses number of vehicles data from 18:00 to 23:45 on 17th Oct 2009. Compared with the whole day number of vehicles data; it only takes 1.19% fraction. Since the flow is calculated by the average of all fixed detectors’ flow, this small missing portion has minor effect on the estimation result.

After that, it changes to the taxis GPS data. There are 500 taxis with GPS running in Stockholm. They send out GPS record about every two minutes. Therefore, the database has massive amount data. Since the estimation of macroscopic model is in one single day, it only necessary to exact one day from database. With the aim of calculation whole region’s flow and speed, the GPS data should be assessed from location error, transmission frequency and temporal and spatial coverage.
Professors and PHD students in logistics division have already made some impressive researches with the taxis GPS data.

*Location error:* the GPS location is used to calculate vehicle space mean speed. These coordinate’s records have 20 meters error at most due to the GPS equipments accuracy characteristic.

*Transmission frequency:* the transmission frequency affects the speed calculation accuracy and movement estimation. According to Mahmood’s (2010) paper, he selected five days data source from 1st to 5th Mar 2010 to assess transmission frequency. The average transmission frequency is about 110 seconds varied by the taxis status.

*Temporal coverage:* temporal coverage ensures estimation is significant important for the whole day. Based on Mahmood’s study, he aggregated number of probes every 15 minutes between 6:00 and 20:00 from 18th Jan to 8th Mar 2010. The result confirms that there is significant quantity and evident flow fluctuation in both the weekday and weekend day.

*Spatial coverage:* Spatial coverage makes sure that the data can be used for the whole region estimation. From his paper, the probes cover 1900 links in the morning from 7:00-7:15 and 2200 links from 8:00-8:15 by investigating aggregation data in 5 days. These links cover up most of important motorways and main roads in the inner city.

### 3.4 Study Region Delimitation

The city of Stockholm contains 188 square km region. Its road network includes roads of various types, such as highway E4 and urban streets, which have 2 to 4 lanes. Moreover the speed limitation of the motorway roads is 90km/hour, 50km/hour in the residential region and 30km/hour in school region. The determination of research region is followed by some rules. First rule is the research region should cover the major city center region where most of trips occur. The second rule is that the selection region envelops more detectors and easier for further analysis.
In this thesis, the study region is delimited the same region as congestion charge system region as shown in the figure 3-2. It is approximately a 46 square km polygon region, which covers most of city centre region such as Södermalm, Norrmalm, Östermalm, Vasastaden, Kungsholmen, Stora Essingen, Lilla Essingen and Djurgården (http://www.transportstyrelsen.se). Moreover, because it is the same as congestion charge system region, together with the fact that Stockholm was built by a series of islands; these 18 tolls record all the traffic flow in and out of the study region. It is very helpful for exit flow estimation. These tolls details from fixed detector data source are listed in table 3-4.

3-2 Study Region same as the Congestion Charge System (http://www.transportstyrelsen.se)
<table>
<thead>
<tr>
<th>Toll id</th>
<th>Toll Name</th>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Danvikstull</td>
<td>59°18'49.93&quot;N</td>
<td>18° 6'13.94&quot;E</td>
</tr>
<tr>
<td>2</td>
<td>Skansbron</td>
<td>59°18'14.65&quot;N</td>
<td>18° 4'46.05&quot;E</td>
</tr>
<tr>
<td>3</td>
<td>Skanstullsbron</td>
<td>59°18'22.88&quot;N</td>
<td>18° 4'38.71&quot;E</td>
</tr>
<tr>
<td>4</td>
<td>Johannesshovsbron</td>
<td>59°18'13.00&quot;N</td>
<td>18° 4'38.38&quot;E</td>
</tr>
<tr>
<td>5</td>
<td>Liljeholmsbron</td>
<td>59°18'43.17&quot;N</td>
<td>18° 1'45.16&quot;E</td>
</tr>
<tr>
<td>6</td>
<td>St Essingen</td>
<td>59°19'19.44&quot;N</td>
<td>17°59'47.25&quot;E</td>
</tr>
<tr>
<td>7</td>
<td>Lilla Essingen</td>
<td>59°19'30.12&quot;N</td>
<td>18° 0'13.99&quot;E</td>
</tr>
<tr>
<td>8</td>
<td>Drottningholmsvägen</td>
<td>59°19'53.25&quot;N</td>
<td>18° 0'39.04&quot;E</td>
</tr>
<tr>
<td>9</td>
<td>Tpl Lindhagensgatan</td>
<td>59°20'5.96&quot;N</td>
<td>18° 0'42.00&quot;E</td>
</tr>
<tr>
<td>10</td>
<td>Klarastrandsleden</td>
<td>59°20'26.10&quot;N</td>
<td>18° 0'46.75&quot;E</td>
</tr>
<tr>
<td>11</td>
<td>Ekelundsbron</td>
<td>59°20'19.44&quot;N</td>
<td>18° 1'48.13&quot;E</td>
</tr>
<tr>
<td>12</td>
<td>Tpl Karlberg</td>
<td>59°20'37.06&quot;N</td>
<td>18° 1'37.61&quot;E</td>
</tr>
<tr>
<td>13</td>
<td>Solnabron</td>
<td>59°20'45.95&quot;N</td>
<td>18° 1'59.29&quot;E</td>
</tr>
<tr>
<td>14</td>
<td>Norrtull N Stationsg</td>
<td>59°21'0.97&quot;N</td>
<td>18° 2'42.80&quot;E</td>
</tr>
<tr>
<td>15</td>
<td>Roslagsvägen</td>
<td>59°22'30.89&quot;N</td>
<td>18° 2'49.44&quot;E</td>
</tr>
<tr>
<td>16</td>
<td>Gasverksvägen</td>
<td>59°21'28.56&quot;N</td>
<td>18° 6'1.70&quot;E</td>
</tr>
<tr>
<td>17</td>
<td>Lidingövägen</td>
<td>59°21'20.38&quot;N</td>
<td>18° 6'16.76&quot;E</td>
</tr>
<tr>
<td>18</td>
<td>Norra Hamnvägen</td>
<td>59°21'19.56&quot;N</td>
<td>18° 6'24.01&quot;E</td>
</tr>
</tbody>
</table>

3-4 Congestion Charge System 18 Tolls Details
3.5 Data Process Flow

In this part, two simple flow charts are depicted to explain how these two kinds of data sources are processed to get the desire variables. As shown in the flowchart 3-3, the analysis starts from fixed detectors. First of all, locate all the detectors positions in the digit map and find out which one are our study target detectors. Then randomly pick up one weekday and weekend day data from three months database. The fundamental diagrams illustrating flow, speed and density relation are built from single link to aggregation variables in certain links network. Comparison with reference diagrams leads to the confirmation of MFD in links network. A simple density control policy is proposed based on average flow-density relation and average flow-exit flow proportion relation.

3-3 Fixed Detectors Data Process Flow
Afterwards, it turns to GPS data source as shown in flowchart 3-4. To correspond with prior study, one weekday taxis GPS data is chosen as sample from one month database. Filter the full taxis by the status value. Once these full taxis are selected out, on one hand, it estimates the number and average speed of full taxis in the study region in every time interval as 30 minutes. On the other hand, it weighs number of general vehicles across the region boundary against number of full taxis that travel in matching movements. According to the constant ratio between vehicles and full taxis, the fundamental diagrams for entire region are constructed. Likewise, the similarity with reference diagrams and links net work diagrams validates the conclusion that non-congestion part of MFD exist in the entire Stockholm study region.

3-4 Taxis GPS Data Process Flow
4 Data Analysis and Results

This section gives you an idea about how to manipulate the traffic data into identical flow, speed and density relation. It will begin with the analysis of easier single link. Afterwards, all fixed detectors data are aggregated for inference. Finally, the taxis GPS data will be filtered and estimated to get whole region variables. Since the construction and inference process is similar as Yokohama case study, the verification of MFD is made by comparison of Stockholm variables relation curves with reference curves in each analysis phase.

4.1 Fixed Detectors Data Analysis

There are total 192 fixed detectors, but not all of them locate in study region. Therefore, it is necessary to find out which detectors belong to the region before further analysis of these fixed detectors data.

The selection method of objective detectors is explained in two steps as following:

At first, transform the detectors’ coordinates and locate them in the digit map. The coordinates need to be transformed since they are in different form from the digit map. The coordinates’ type in the data source is SWEREF 99 form, which is the Swedish geodesy reference system implemented in 1999. Nevertheless, the digit map software only accepts WGS 84 form that is a reference coordinate system widely used in the digit map organization. Many tools could transform the coordinates SWEREF 99 to WGS84. In this thesis, CoordTrans, which is able to transform the popular types of coordinates straightforwardly into each other, is adopted. Once the coordinates are transformed, these detectors locations are imported into Google Earth. The details are shown in figure 4-1. The blue labels are fixed loops and microwave beams. The pink labels are the congestion charge system tolls. In details, each label contains the detectors’ id and location description.
Second, the study region is depicted in the same digit map with the location of congestion tolls. As shown in figure 4-2, shadow coverage polygon is the study region. It is easy to observe that 88 detectors are belonging to this region from the map.

4-1 All the fixed Data Collection Detectors in Stockholm city (By Google Earth)
4-2 All the Fixed Data Collection Detectors in the Study Region ((By Google Earth)

Next, given that the objective detectors have been figured out; it is time to define the parameters and estimate variables. Some parameters, such as average flow, average time-mean speed, directly come from data source. One exceptional parameter is lane length, which is defined as length of lane where lay the fixed detectors. These lengths are obtained by elaborately measured on the digit map. In this thesis, they are designated as following:

- $n$ is total number of detectors
- $i$ is a road lane with the fixed detectors
- $l_i$ is the lane length between two intersections
- $q_i$ and $v_i$ are the corresponding lane vehicles flow and average time mean speed
- $v_s$ is the space mean speed
- $A$ is the whole study region
- $A_f$ is the subsidiary links with fixed detectors
Based on the above basic parameters, aggregation variables, such as weighted average flow, density and speed, are calculated in below formulas. Weighted flow is with reference to Thomson’s (1967) definition in London study case. Time mean speed, space mean speed and density calculation are from fundamental flow theory.

Unweighted average flow: \[ q^u = \frac{\sum q_i}{n} \]

Unweighted average density: \[ k^u = \frac{q^u}{v_s} \]

Weighted average flow: \[ q^w = \frac{\sum q_i l_i}{\sum l_i} \]

Weighted average density: \[ k^w = \frac{q^w}{v_s} \]

Space mean speed: \[ v^s = \frac{n}{\sum (1 / v_i)} \]

Time mean speed: \[ v^t = \frac{\sum v_i}{n} \]

The unweighted average flow is the mean value of all detectors’ flow. It is restricted application in the detectors parts of the lane. In the same way, the unweighted density is representative of the average density within the detectors segments of the lane.

By contrast, the weighted average flow is calculated by sum of all detectors’ flow multiplying each length of lane where the detector locates dividing sum of all
length. Since the calculation process takes the length of lane into account, therefore it characterizes whole link’s traffic condition. The assumption is based on the fact that the flow measured by the detector is able to be representative of the whole link not only the detector’s part. The main reason is that the frequency of the flow is 15 minutes, which is much more than the standard traffic cycle. The long time period can eliminate the signal’s affection and ensure that the detector’s part flow is taken for the whole lane flow. However, this inference procedure is only suit for traffic flow variable. Because the speed measured by the detector is restricted used in the detector’s part.

### 4.1.1 MFD Existence in the Links with Detectors

Since the single detector has less data and easier to process, the analysis starts from single detector. First, two independent detectors, id 29 and 53 are chosen randomly from 88 detectors in the region. The locations of these two detectors are shown in figure 4-3. Time period is chosen as Monday (2009-09-21) from three months. Pattern of speed against flow in the whole day is plotted in figure 4-4. These plots show in apparent disarray way. For instance, link flows distribute in wide range from 20 vehicles per hour to 760 vehicles per hour at the same speed 33 km per hour. This phenomenon is explained by the fact that the individual detector tends to be affected by the fleet fluctuation in short period.
Afterwards, the analysis is extended from single detector to multiple links. Weekday (2009-10-15) and weekend day (2009-10-17) are both selected to aggregate of the unweighted average flow. Figure 4-5 shows the time sequence unweighted flow in each day. There are two obvious flow peaks at A.M 8:30 and P.M 18:15 on the weekday. In comparison, there is only one long last rush hour period from A.M 11:45 to P.M 17:00 on the weekend day. More to the point, the weekend flow peak value is less than the weekday’s climax as anticipated.

Besides the flow, speed is the other critical variable. Generally, speed is divided into time mean speed and space mean speed based on different reference. The raw data source contains only time mean speed in each detector every 15 minutes.
With previous formulas, the time mean speed can be transformed into space mean speed.

Figure 4-6 illustrates one weekday (2009-10-15) time mean speed as triangle and space mean speed as plus sign change with time. These two speed curves perform similar sharp decay in the morning rush hour. As anticipated, the space mean speed is always lower than the time mean speed.

Next, pair of flow, speed and density is plotted. As shown in figure 4-8, both the space mean speed and time mean speed go slightly down while the unweighted average flow increasing. But the least speed is over 30 km per hour on weekday (2009-10-15). Compared with Yokohama curve in figure 4-7, where the higher speed regime pointed out in ellipse, they represent analogous feature. Another figure 4-10 shows the density linearly ascending with unweighted average flow rising. It behaves similar linear relation as the reference figure 4-9 in the lower density regime where the density value is lower than 30 vehicles per kilometers. According to earlier indication, the unweighted flow and density are only representative of the detectors parts. The higher ordering of the points and close correlation with reference curves prove the non-congestion part of MFD existence in the detectors’ parts.
At last, it extends the scale from detectors’ parts to the whole links. The weighted flow is assumed as the whole link flow. Compared with average flow, here the length of links are inducted into the weighted flow computation. Therefore, the weighted flow is more likely to be considered as the full links’ flow. The scatter points for pair of weighted flow and weighted density in figure 4-12 perform comparable pattern as reference Yokohama study curve’s low density regime in figure 4-11. In the same way, although the space mean speed could not be simply considered as the whole link speed, weighted flow and space mean speed in figure 4-13 is still another significant indication. It behaves decline linear relation while
the links’ flow rising. Both of these comparison results indicate non-congestion part of MFD existence in the network of links with detectors.

4-12 Weighted Flow vs. Density Aggregated in one day

4-11 Weighted Flow vs. Density in Yokohama by Geroliminis in 2007

4-13 Weighted Flow vs. Space Mean Speed of one Day

Both the Stockholm case and Yokohama case pick up one day to construct the fundamental diagrams. Flow, speed and density diagrams are compared in pairs. The similarity shape and change trend of these diagrams shows that the non-congestion part of MFD are existence in the links with detectors.

There are several different parts in the comparisons. First is the different scale. For instance, in the flow-speed relation, average free flow speed is 52 km per hour
in Stockholm. In contrast, Yokohama case is 32 km per hour. The difference is caused by different speed limits between two city regions. Second, only low flow, speed and density regimes in both the flow-speed and flow-density reference diagrams reappear. This is plainly explained that the Stockholm traffic situation is better than Yokohama case. But it is not the whole story. There is also congestion problem in Stockholm urban region. For instance, the detector 513 shows the congestion regime in five weekdays’ (2010-10-02, 2010-10-08) aggregation in figure 4-14.

Because the unweighted and weighted flows are calculated by the average of all detectors flow, this individual congestion is counteracted by the aggregation effect. These detectors locate in both saturation roads and under saturation roads. The low burden roads underestimate the whole region congestion situation.

According to the Yokohama’s density flow plots, there is a crucial point that the growth of flow achieves to the peak and once exceeding this point, even the density increasing, the flow still dropping. Based on this speculation, in the next part, it proposes one uncomplicated policy of traffic control application.
4.1.2 Maximize Trip Completion by controlling density

In this part, a simple density control strategy will be proposed based on the previous analysis results. The fundamental assumption is that the vehicles arriving to or leaving the boundary are considered as the trip completion. It is attained by the exit flow in the congestion charge system. The purpose is to maximize trip completion. If there is a stable ratio between the exit flow and the weighted average links flow, it is possible to control the exit flow by keeping the density at the decisive value, which is corresponding to the highest average weighted links flow.

Because these two flows have different time interval, the weighted average flow with 15 minutes frequency is adjusted to the same interval as the exit flow with 30 minutes frequency. Afterwards, these two flows are plots in time series for a weekday (2009-10-15). As shown in the figure 4-15, the left scale with 5000 vehicles interval is for exit flow and right scale with 400 vehicles interval is for weighted flow. These two curves show correlation, for example, the similar ascending flow in the morning rush hour from 5:30 A.M to 8:30 A.M.

Furthermore, the ratio of two flows for each point is calculated and compared with the average ratio. As shown in figure 4-16, the straight line is the average ratio, whose value is 0.046. Although each point ratio fluctuates along the average line, the divergence is in small scale. The standard deviation, which measures how the sample deviates from the average value, is 0.007.
In summary, a low standard deviation indicates that these data points tend to be very close to the mean of sample. It verifies that there is constant ratio between the region exiting flow and network links’ weighted flow. Therefore, it is guarantee to utilize the controlling density policy to realize the maximum of trip completion. The control strategy is explained by that author establishes the connection between exit flow and density using average weighted flow as intermediary. The management of density could maximize the networks’ weighted flow. Meanwhile, because of the stable ratio, it will finally realize the purpose of maximizing region exit flow.
The controlling density policy can be illustrated in the terms of unit of density. Since the unit of density is vehicles per km, there are two parameters, which can be adjusted to keep the density under the critical points. One method is decreasing the number of vehicles. The other is increasing the length of roads. Increasing length of roads means to construct more roads and it requires more investments and time. In contrast, the method of decreasing the number of vehicles is easily to be implemented. There are some executed paradigms, for instance 2008 Beijing Olympic traffic restriction policy. It implemented the traffic control policy by banning parts of vehicles. In details, the vehicles with license plates ending in odd numbers are banned from roads on even-numbered calendar days. This policy reduced 600 thousands of vehicles flow per day and as result, the average speed during rush hour increased 15%. The accessibility was improved during that period.

4.2 Taxis Data Analysis

The preceding investigations and conclusions apply for links network with detectors. Afterwards, because the final purpose of this thesis is verification of MFD existence in the whole region, taxis GPS (Global Positioning System) data, which cover the entire region, provide a practicable resource. Usually, the taxis drive dissimilar with normal vehicles for the reason that the taxis stop more frequently to pick up passengers and follow the circle route to look for passengers. Therefore, it is necessary to filter taxis, which drive similar as normal vehicles. In this thesis, because the taxis carrying on passengers drive directly to the destination via shortest path, it assumes that these full taxis are representative of normal vehicles.

4.2.1 Taxis Data Manipulation

In the taxis GPS spatial coverage reliability part, it points out that these data are able to cover up most of important motorways and main roads in the inner city. The study region defined here is the inner city parts as well. However, the prior data is for all taxis. It is still necessary to check the full taxis coverage and prove
that they can envelop the entire region. Figure 4-17 is one taxi id ‘11002’ travel locations of one single day (2010-04-12) when the taxi status is ‘meter on’. As figure 4-17 illustrates that the shadow region is the study district. And the labels designate the taxi’s location with time series. It shows out this taxi travels around and through the study region. Since these location labels are belonging to only one taxi, it is reasonable to surmise that 500 taxis cover arterials and most of major roads within the study region in one day.

4-17 Individual Taxi Trace in one Weekday (By Google Earth)

4.2.2 Estimation Ratio of Full Taxis to All Vehicles

Consistent with last part result, the descriptions of the full taxis variables represent the normal vehicles performance in this study region. Nonetheless, there is still a gap to cross how to estimate normal vehicles variables from the full taxis.
data. In this section, the connection between the taxis and normal vehicles is tried to be established. Detailed steps are shown in figure 4-18 below.

4-18 Taxis Flow Estimation Process

- One weekday (2010-04-12) is chosen randomly. There are about 38,300 pieces of information extracted from GPS database.

- Because only full taxis behave as normal vehicles, these data are selected out according to the parameter “status”.

- Afterwards, these data will be sorted firstly by measure time and then by vehicle id.

- Once the data is reorganized, the whole data should be separated in every 30 minutes which is synchronism with congestion charge system data time interval.
Since each piece of GPS information contains coordinates, it is possible to determine whether the taxis are in the research region or not with the help of “inploygon” command in Matlab. The location status is defined as ‘in’ or ‘out’.

At last, when the taxis travel from outside into the research region, their location status change from “out” to “in”. Each of this change accumulates the travel in flow. In the same way, when the taxis travel out, the out flow adds.

These full taxis in and out flows are accumulated in the every 30 minutes and compared with the same day congestion charge system data, which includes all vehicles in and out flows in the same time interval. Both in and out ratio for each time points are plotted in figure 4-19 in time sequence.

The plus sign is in flow ratio while the triangle is out ratio. As the ellipse circling out, it shows out that from morning 6:30 to afternoon 20:30 both in and out ratios perform concentrated. These disperse points in the late night and early morning are explained by the reason that full taxis operate more regularly than normal vehicle in this period. Exclude this abnormal time period, during the stable period, the average ratio is 0.02012 and standard deviation is 0.0029. The small standard deviation implies these data points are very close to the mean value. In addition, the regression between in ratio and time is analyzed as table 4-1. The coefficient
of time is statistically insignificant ($t<1.96$), which means the ratio does not change along the time. It sums up that the full taxis in the study region take account about 2% of the all vehicles.

<table>
<thead>
<tr>
<th></th>
<th>Coefficients</th>
<th>t Stat</th>
<th>P- value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.018</td>
<td>9.72</td>
<td>$2.58 \times 10^{-10}$</td>
</tr>
<tr>
<td>t</td>
<td>$7.18 \times 10^{-5}$</td>
<td>0.55</td>
<td>0.59</td>
</tr>
</tbody>
</table>

4.2.3 Estimation of the Speed and Flow

In view of the fact that the stable proportion between full taxis and all vehicles exist, it is time to estimate full taxis and vehicles’ flow, speed and density relations in the whole region.

For each time slice in 30 minutes, set variables as below:

$r$: ratio as full taxis comparing with all vehicles

$N_f$: Number of full taxis in the region

$N_v$: Number of vehicles in the region

$v_{sf}$: Speed of single full taxi

$s$: Direct distance between two uninterrupted taxis coordinates

$t$: Time difference between two continue ‘measure time’

$v_{st}$: Space mean speed for all full taxis

$v_s$: Space mean speed for all vehicles

$T$: Total travel time for all full taxis in the region within the time slice

$\Delta t$: Time slice as 30 minutes
There is one significant assumption that the full taxis’ speed is considered same as all vehicles speed. This assumption is supported by the reason that the full taxis’ travel habit is similar as the ordinary vehicles. The formula of number of full taxis is according to Geroliminis and Daganzo’s (2007) Yokohama’s case study.

\[ v_{ts} = v_s \]

\[ N_i = \frac{T}{\Delta t} \]

\[ N_s = \frac{N_i}{r} \]

\[ v_{ti} = \frac{s_i}{t_i} \]

\[ v_{ts} = \frac{n}{\sum_{i}^{n}(1/v_n)} \]

**4.2.4 MFD Existence in the Whole Region**

The whole region number of vehicles is speculated from above calculation formulas. Both growth and decay along the time in one weekday (2010-04-12) are shown in below figure 4-20. There are two noticeable flow peaks at 08:00 and 16:00. It performs comparable shape as preceding average links flow curve in figure 4-5.
Furthermore, the whole region number of vehicles and average speed plots create the quadratic model with R squared 0.9179 in figure 4-22. The R squared value indicates how well the regression line approximates the real data points. The higher R value means the regression line fits the data very well. Refer to the reference diagram with the same variables in figure 4-21, both of average speeds start from 30km per hour and the declining when the region number of vehicles increasing. They have the analogous performance. This indicates that from macroscopic view, traffic fundamental relations exist in the entire research region.

Nevertheless, there are some different parts required explanations. The first is the average free flow speed. In the links net work diagrams, the highest average speed is 46km per hour. One reason is that the detectors locate in the middle of road and the record average speed is higher than the real travel condition. The other reason is that during the estimation of the full taxis speed, which equals to the normal vehicle speed, some very low speed points, such as less than 2km per hour, are included. These low speed points are produced when full taxis stop at intersection or wait for signals. The second is the number of vehicles scale. In the reference Yokohama diagram, the high number is about 12,000. In contrast, Stockholm region is 23,000. This difference is clarified by the region area difference. The study region in Stockholm is four times than the reference region.
In brief, fundamental variables diagrams from the fixed detector and taxis traffic data are similar to Yokohama case’s corresponding diagrams. However, the Stockholm case diagrams only appear in the non-congestion part of reference diagrams. It provides evidence that non-congestion part of MFD exists in Stockholm urban region. It also suggests but not secures to pronounce that the macroscopic traffic state of Stockholm urban region is in non-congestion condition in that single day. Desire for more reliable traffic evaluation result, it requires more aggregation data of longer time period and precise estimation process.
5 Conclusion

Based on preceding chapters, author finds that the diagrams, which derived from Stockholm research region, verify that non-congestion part MFD exists in Stockholm center region. In this chapter, it will sum up the result of the thesis. Besides, it includes the insufficiency of the data analysis process and proposes continuous improvements and future potential developments.

5.1 Summary

The objective of this thesis is making use of the available data sources to find the evidence that prove MFD exist in Stockholm study region. In details, at first two individual detectors are randomly chosen to plot the flow against speed every fifteen minutes for twenty four hours. The scatter points do not perform as the expected classical relation between flow and speed and it implies that under the single detector and short temporal condition, the result is vague.

Therefore, it is required to analyze the aggregation data of multiple detectors. The aggregation flow and speed fluctuation suggest obvious A.M and P.M peak hour increasing and descending. Furthermore, aggregated average flow, weighted flow, speed and density are plotted in pairs and compared with reference Yokohama’s curves. The conformity of the comparison provides evidence that non-congestion part of MFD establishes during the links with detectors. Since the prediction parabola curve of density and flow has the maximum flow value, combining with the constant proportion between the region exit flow and average links flow, the traffic density controlling policy is recommended to achieve the highest accessibility.

Finally, it turns to taxis GPS data. The full taxis are assumed to drive in the similar way as general vehicles. Consequently the full taxis’ speed is equal to the vehicles’ speed. Based on this assumption, the whole region vehicles variables,
such as flow and speed, can be estimated from the full taxis GPS data. In the similar way, fundamental diagrams are constructed and matched up to fixed detectors and reference figures. The high likeness demonstrates also that non-congestion part of MFD is applicable in whole research region.

The thesis proves that from spatial vision, macroscopic scale analysis is better than the individual sections. But from temporal view, only one single day’s data is extracted for study. This short period is enough for the MFD verification, but could not guarantee evaluating the traffic situation.

Another imperfection is the estimation process. First is that during the of trip completion, only the vehicles arriving to or leaving the boundary movements are considered as trip completion. Actually, some vehicles finish their trips within the region as well. Second is the taxis speed calculation process. Since there is no digit map match, the travel distance uses the straight line distance between two continuous points, which is usually less than actual travel path distance. This causes the calculated average speed higher than the real condition.

What is more, there is the less quantitative analysis in the thesis. Author simply compares the shape with reference figures visually and does not work out how close they fit each other. At last, there is no particular model constructed and few statistical tests during the conjecture estimations.

### 5.2 Future work

There are several potential works, which can be progressed in future. With the constraint of the time and energy, author only picks one single day. If longer time period data is processed, the estimation result will be more realistic.

In addition, although the evident comparison is able to give readers a direct viewing, to make it more convincible, it is necessary to build the models among flow, speed and density and compute the similarity with reference models in further investigation. Besides, more statistics tests are needed to guarantee the estimation precision.
In the end, with the accurate model and long term results, it is possible to recommend more practical policies and more concrete execution procedures.
Bibliography


Carlos Daganzo and Nikolas Geroliminis (2008), an analytical approximation for the macroscopic fundamental diagram of urban traffic, Transportation Research Part B, 42(9):771–781.


Guillaume Leduc (2008), Road traffic data: collection methods and applications, JRC 47967.

Gunilla Pettersson (2008), The Stockholm solution for intelligent toll collection, IBM CeBIT.


ISA in Stockholm (2005), Results from trials and possibilities for implementation, Trafikkontoret.


Leslie Edie (1961), Car following and steady-state theory for non-congestion traffic, Operation Research, 9, pp.66-76.


Nissan Albania and Bang Karl (2006); Evaluation impacts of the motorway control system (MCS) in Stockholm, ETC. proceedings paper 1653.


www.stockholm.se

www.transportstyrelsen.se
Appendix

%Main part for how calculate the speed, total travel time of full taxi
m=xlsread('e:\412data\1.xlsx');
% read document includes 30minutes full taxi data ordered by time and vehicle id
[o,p]=size(m);
for i=1:o-1;
    if m(i,1)==m(i+1,1)
        d(i,1) = distance(m(i,2),m(i,3),m(i+1,2),m(i+1,3), 6378.1);
        % Calculate two points distance, unit by Km
        d(i,2) = m(i+1,4)-m(i,4);
        % time difference
    end
    if m(i,1)~= m(i+1,1)
        d(i,1) = 0;
        d(i,2) = 0;
        % Judge the vehicle id change
    end
end

d(:,3)=d(:,1)./(d(:,2)*24);
% speed calculation, since the time is in decimal number, time 24 to get unit
km/hour

d(1,4)=sum(d(:,2))*48;
% Total taxis number, total full taxis travel time/ time slice (0.5 hour)
d(1,5)=nanmean(d(:,3));
% average speed of full taxis
xlswrite('e:\412data\speed1.xls',d);