EMPIRICAL ESTIMATION OF A MACROSCOPIC FUNDAMENTAL DIAGRAM (MFD) FOR THE CITY OF CAPE TOWN FREEWAY NETWORK

Minor Dissertation

Master of Engineering, Transport Studies

By

John Koketso Rammutla

In partial fulfilment of requirements for the Master of Engineering in Transport Studies

Centre for Transport Studies
Department of Civil Engineering
Faculty of Engineering & Built Environment
The University of Cape Town, South Africa
May 2020

KEY WORDS

Macroscopic Fundamental Diagrams (MFD), Empirical Data, Fixed Loop Detectors, City of Cape Town. Freeways. Hysteresis Loops. Congestion.

ABSTRACT

The City of Cape Town is the most congested city in South Africa, with Johannesburg coming in second. Capetonians are spending 75% more time in traffic because of the congestion during peak hours, thus reducing time spent on leisure and other activities. Due to population growth, increasing car ownership and declining capacity of rail infrastructure, Cape Town's road infrastructure will continue to be under severe pressure if the status quo is maintained.

Research shows that congestion levels in urban areas are key factors in determining the effectiveness and productivity of the transport system. Traffic congestion poses a threat to the economy and the environment. Increasing corridors' capacity by increasing the number of lanes does not necessarily solve the problem. Effective urban traffic management and efficient utilization of existing infrastructure are critical in creating sustainable solutions to congestion problems. To achieve this, it is important that appropriate urban-scale models and monitoring strategies are put in place.

Effective traffic management and monitoring strategies require accurate characterization of the traffic state of an urban-scale network. Several approaches, including kinetic wave theory and cell transmission models or macroscopic traffic simulation models, have been proposed and developed to describe the traffic state of an urban-scale network. However, these approaches are limited and require significant amounts of computational time and effort.

The application of macroscopic fundamental diagram (herein referred to as MFD) to characterize the state of an urban-scale network has thus far proven to be more effective than other approaches. MFD represents the state of urban traffic by defining the traffic throughput of an area at given traffic densities. It describes the characteristics and dynamics of urban-scale traffic conditions, allowing for improved and sustainable urban scale traffic management and monitoring strategies.

Against this backdrop, the existence of MFD for the City of Cape Town (CoCT) urban-scale network is yet to be established and the implications yet to be understood, as in other parts of the world. The main aim of this research was, therefore, to empirically estimate the macroscopic fundamental diagram for the CoCT's freeway network and analyse its observed features. To achieve this, observed data of 5 minutes periods for the month of May 2019 was used to estimate the MFD.

The results confirmed that when the chaotic scatter-plots of flow and density from individual fixed loop detectors were aggregated the scatter nearly disappeared and points grouped neatly to form a clearly defined free-flow state, critical state and the formation of hysteresis loops past the critical density corresponding with the network observed maximum flow. Further analysis of the MFDs showed that a single hysteresis loop always forms past the critical density during the evening peak in a weekday MFD. However, it was inconclusive during the morning peak period in weekday MFDs. Lastly, an explicit hysteresis loop seldom appears in a Saturday MFD when the peak of traffic demand is lower than on weekdays.

In order to understand the dynamics of the congestion spread, the freeway network was partitioned into penetrating highways network and the ring highway network. The results showed that the maximum flows observed for the two sub-networks were significantly different (943 veh/hr/lane for the penetrating highways network and 1539 veh/hr/lane for the ring highway network). The penetrating highways network's MFD indicated the presence of congestion in the network whereas the ring highway network indicated only the free-flow state (no indication of congestion) during peak periods.

The congestion seen on the penetrating highways network was found not to be sufficiently spread on those highways. On the 24th May, congestion on the penetrating highway network was observed during both the morning and evening peak periods, whereas on the 31st May congestion was observed mainly during the evening peak period, with hysteresis-like shape. These observations confirmed that congestion during peak periods is not homogenously spread across the entire network, certain areas are more congested than others, hence the observed formation of hysteresis loops and slight scatters.

Lastly, the hysteresis loops observed in the penetrating highways network's MFD was further characterized in terms of their shape and size. First, the results showed that the slight scatter and hysteresis patterns observed in penetrating highways network MFD's vary in size and shape across different days.

The shapes of the hysteresis loops observed during both the morning and evening peak periods, were type H2 hysteresis loops, signifying a stable recovery of the network with the average network flow remaining unchanged as average network density decreases during the recovery.

Characterization of the size of the observed hysteresis loops showed that the drop of the hysteresis (an indicator of network level of instability during recovery phase) was smaller, signifying a more stable network traffic and homogenous distribution of congestion during the recovery phase.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

LIST OF ACRONYMS

Term	Definition
MFD	Macroscopic Fundamental Diagram
SANRAL	The South African National Roads Agency Limited
FMS	Freeway Management System
CoCT	The City of Cape Town
GMFD	Generalized Macroscopic Fundamental Diagram
OD	Origin-Destination
LOS	Level of Service
HCM	Highway Capacity Manual
PHF	Peak Hour Factor
V/C	Volume to Capacity

PLAGIRISM DECLARATION

I know the meaning of plagiarism and declare that all the work in the document, save for that which is properly acknowledged, is my own. This thesis/dissertation has been submitted to the Turnitin module and I confirm that my supervisor has seen my report and any concerns revealed by such have been resolved with my supervisor

Student Number: rmmkok001 **Signature:** Signed by candidate

ACKNOWLEDGEMENTS

The Lord is gracious and kind. I give all thanks and credits to Him for the strength, courage and wisdom to carry out this research work.

I would like to express my heartfelt gratitude to my supervisor, Professor M.H.P Zuidgeest, for the opportunity to work with him on this topic. His insight, guidance, encouragement and enthusiastic supervision made this research possible. I would also like to express my sincere appreciation to the South African National Roads Agency Limited (SANRAL) for granting me access to their Freeway Management System (FMS) data which proved to be instrumental to this research.

Special thanks to Minette Coetsee for additional CTO data. I wish to extend my appreciation, love and thanks to my lovely wife for her consistent encouragement and assistance throughout the period of this research. Thank you, B. A big thank you to Theophilus Dzingai for his consistent encouragement and support throughout. Finally, my appreciation goes to my SANRAL colleagues, Mike Vinello-Lippert for the assistance when I required it most.

Chapter 1 Introduction

This chapter is dedicated to setting out the research framework. It commences with the background and justification for the research. Following is the identification of the research problem and definition of the research objectives and questions. Lastly, research design is presented, providing an overview of how the research will be carried out.

1.1. Background and Justification

According to TomTom (2020) traffic data, the City of Cape Town is the most congested city in South Africa followed by the City of Johannesburg. Due to an ever-increasing population figures, increasing car ownership and declining capacity of rail infrastructure, Cape Town's road infrastructure is expected to be under severe pressure if the status quo is maintained. Research shows that congestion levels in urban areas are key factors in determining the effectiveness and productivity of the transport system (Venables, 2007).

The understanding of urban congestion is key for effective traffic management strategies, interventions, policies and regulations. Historically, freeway traffic flow analysis has been based on microscopic models concerned with single road segments. Predictions of traffic state variables such as traffic volume, travel time and density were estimated using these microscopic models. The relationship between density and the corresponding flow of traffic on a highway is generally referred to as the fundamental diagram. The fundamental diagram can be used to characterize the traffic condition of a major arterial or highway.

According to the diagram, (a) when the density on the highway is 0, the flow is also 0 because there are no vehicles on the road, (b) as the density increases, the flow also increases, (c) however, when the density reaches its maximum, generally referred to as *jam density*, the flow must be 0, and (d) it follows that as density increases from 0, the flow will also initially increase from 0 to a maximum reduction of the flow.

The application of these models was limited to major arterial routes and motorways. However, effective traffic management and monitoring strategies require accurate characterization of the traffic state of an urban-scale network. The models were limited in this regard, hence these could not be used to characterize the traffic condition of an urban-scale networks (Tsubota, 2014).

Loop detectors and/or probe vehicles sensors are installed in urban road networks to monitor and measure traffic congestion on indivual links or road segments. It is, however, still difficult to accurately understand and evaluate network-wide traffic state or network performance based on disaggregated data of individual links. This is mainly because disaggregated traffic flow dynamics on congested networks are complex due to interactions between drivers' route choice behaviours and nonlinear traffic flow phenomena.

Against this backdrop, Daganzo (2007) proposed macroscopic congestion measure called Macroscopic Fundamental Diagram (MFD), as an alternative to the disaggregate approach. In essence, the MFD captures the state of urban traffic network at an aggregate level by relating the space-mean density (number of vehicles in an area) to the space-mean flow (area's trip completion rate), provided that the network is homogeneously congested and demand evolves slowly over time. Geroliminis and Daganzo (2008) showed that a well-defined reproducible MFD, which is invariant when demand changes, exists in Yokohama, Japan.

A well-defined MFD can be used to establish robust traffic management strategies because it makes it possible to understand the traffic state of an urban network using average densities only and real-time information can also be used to flexibly adjust control parameters. For example, a road network can be prevented from reaching a congested state by regulating area inflow rates to maintain average density that is less than the MFD critical density.

Given that the City of Cape Town is currently the most congested city in South Africa, it seems logical and necessary to establish the existence and implications of the MFD for the city's freeway network. It is for this reason, that this research study seeks to establish the existence and implications of MFD for the City of Cape Town road network using empirical data gathered from SANRAL Freeway Management System (FMS). Instrumental and central to this research is the reliable and accurate source of data provided by SANRAL. For the purpose of this study, only data from fixed loop detectors will be used to confirm the existence of the MFD – this is primarily due to both time and budget constraints.

Fixed loop detectors are positioned at various points of the City of Cape Town's road network for vehicle counting, traffic signal control and congestion monitoring purposes. Data from these loops will be used to collect flow information i.e. number of vehicles passing each loop every five (5) minutes and the average densities. Recently, fixed loop detectors' data have been used to estimate MFDs in the cities of Toulouse (Buisson and Ladier, 2009), Sendai (Wang *et al.*, 2015) and Chania (Ampountolas and Kouvelas, 2015), to name a few.

Lastly and interestingly, some empirical studies have demonstrated that MFDs with high scatters and hysteresis loops exists for certain network types such as freeways or heterogeneous network (Buisson and Ladier, 2009; Geroliminis and Sun, 2011b; Saberi and Mahmassani, 2013). These findings are considered important to this research study, given that the City of Cape Town's network study area consist of freeways.

1.2. Research Problem

The Cape Town Metropolitan Area experiences perturbing traffic congestion problems in certain areas, particularly on major National and Provincial Routes within a 10-20km radius of the CBD of Cape Town. Currently, the main routes servicing the Cape Metropolitan Area are the National Routes N1, N2, N7 & R300 together with heavily utilized lower order arterial roads. Cape Town is the most congested city in South Africa, resulting in adverse economic (drivers lost productive time in traffic congestion), social and environmental impacts on the city and South Africa at large.

Macroscopic Fundamental Diagrams have proven to assist with accurate characterization of the traffic state in a road network, monitoring and evaluation of effect of traffic management (Xu et al. 2013). To date, an empirical estimation and analysis of Macroscopic Fundamental Diagrams for the City of Cape Town (CoCT) road network has not yet been established. This research study aims to establish the existence of Macroscopic Fundamental Diagrams for the CoCT's freeway network and carry out basic analysis of the diagrams. Given the CoCT's highway network structure it is expected that such MFD would exists.

1.3. Research Objectives and Questions

1.3.1. Main Research Objective

The main objective of this research is to empirically estimate the Macroscopic Fundamental Diagram for the City of Cape Town's freeway network and to explore its observed features. To achieve this, observed data sourced from fixed loop detectors of 5 minutes periods over one month (May 2019), will be used to estimate the MFDs for the City of Cape Town highway road network.

The outcomes of this research may be used to better understand the City of Cape Town freeway network's performance in terms of traffic flow. This knowledge may in turn be used to inform decisions regarding the appropriate traffic management strategies and interventions (in form of policies) for implementation in order to relieve congestion during morning and evening peak periods.

1.3.2. Specific Research Objectives

The research purpose to achieve the following key objectives;

a) Derive and analyze whole month MFD for the CoCT.
b) Derive and analyze daily (Fridays) MFDs for the CoCT.
c) Analyze the impact of traffic demand of the shape of MFD.
d) Analyze the impact of partitioning the CoCT freeway network on the MFD shape
e) Based on the outcomes of the research, provide recommendations of future research needs.

1.3.3. Research Questions

The following research questions have been formulated to operationalize the above research objectives.

Table 1-1: Research Objectives and Questions

No	Research Objective	Research Questions
1	Derive and analyze whole month MFD for the CoCT	Given the available Empirical data, does a whole month MFD exist for the CoCT freeway network?
2	Derive and analyze daily (Fridays) MFDs for the CoCT	Given the available Empirical data, does a daily MFD exist for the CoCT freeway network?
3	Analyze the impact of traffic demand of the shape of the daily MFD	What is the impact of changes in traffic demand patterns on the shape of the daily MFD for the CoCT?
4	Analyze the impact of partitioning the CoCT freeway network on the MFD shape	What would be the impact of partitioning the freeway network of the shape of the daily MFD?
5	Based on the outcomes of the research, provide recommendations of future research needs	Based on the research outcomes, what are the recommended future research needs

1.4. Research Design

The processes employed in this research are summarised in Figure 1-1 . The research problem was founded upon extensive consultation with relevant literature on the subject matter across various cities across the globe. The fact that there is currently no evidence that confirms the existence of urban-scale macroscopic fundamental diagram for the CoCT's urban-network, contributed significantly to the formulation of the research problem.

The formulation of the research problem made it possible to set out research objectives and questions. SANRAL Freeway Management System (FMS) data produced by fixed loop detectors was sourced and organized in the preferred format for analysis purposes. The data was sourced and available in xls and pdf formats. The study area consists of fixed loop detectors positioned at various locations along the freeways.

Attempts were made (to no success) to access data produced by some of the major arterials (under the jurisdiction of the CoCT Metropolitan Municipality) located within the study area. As a result, only data from loop detectors located along the freeways were used for this research. Sorting, organizing and analyzing the data using excel spreadsheet was an extensive process. Due to time and financial constraints, no other software packages were used to organize and analyze the mass data, which would have probably made the process less extensive.

Various assumptions were made during various stages of the analysis process. These steps gave effect to the CoCT MFD which may be used to determine the performance of the urban network. Based on the research findings, conclusions were drawn, and recommendations proposed.

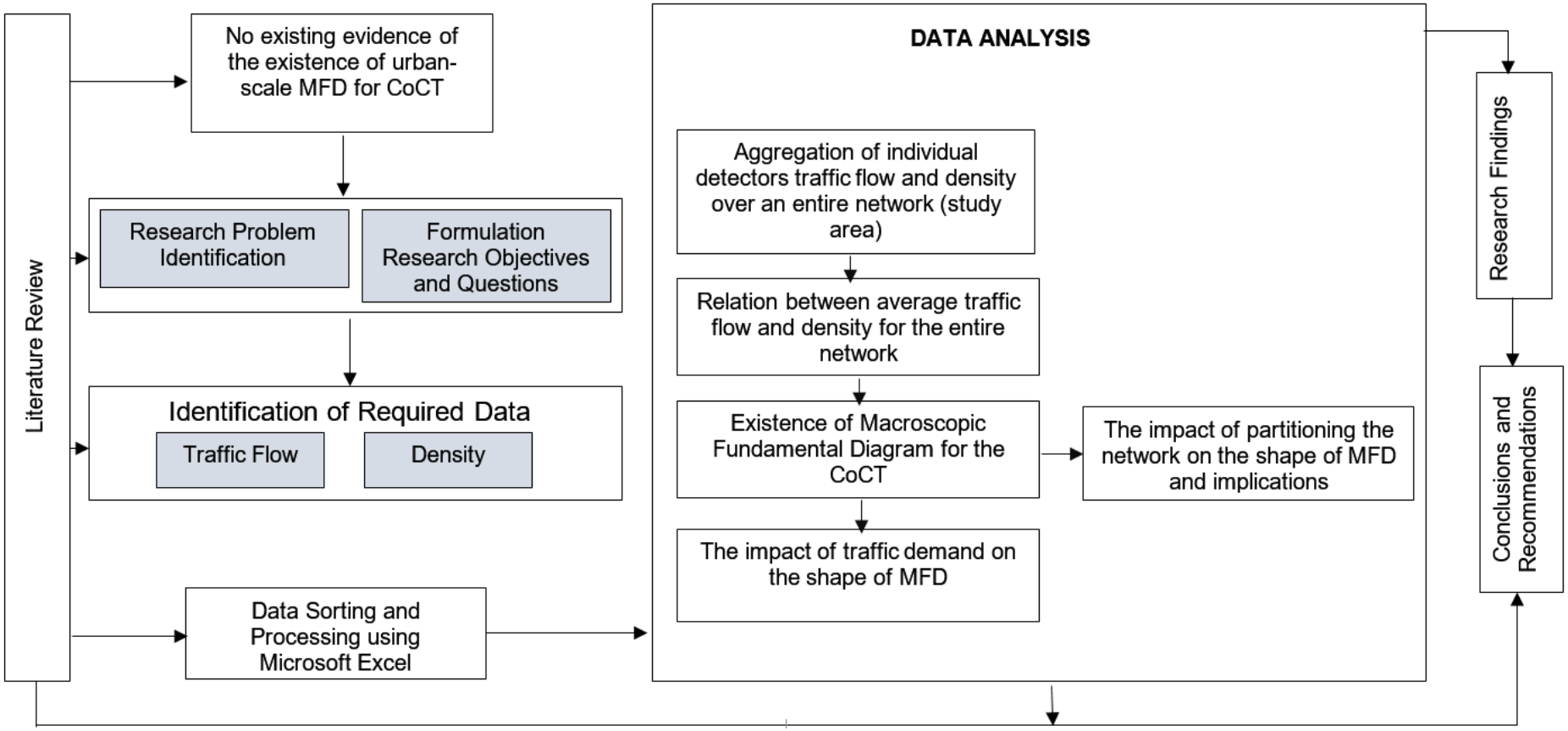


Figure 1-1: Research Design

1.5. Research Benefits

The research findings are expected to yield the following benefits to the local transport planning and traffic management authorities and the body of research knowledge in the field of MFDs.

a) The confirmation of the existence of macroscopic fundamental diagram for the CoCT freeway network will enable the local transport planning and traffic management authorities to gain better understanding of the state of traffic on an urban scale.

b) The authorities including SANRAL, Western Cape Government and CoCT will be able to use the understanding of the state of traffic on uran-scale in terms of production and accumulation, to implement strategies and policies that may result in significant reduction in the high levels of congestions.

c) The research findings will set a precedent for other researchers to carry out further related research work based on this research's recommendations and by so doing, contributing to the local body of research knowledge in the field of MFDs for major South African cities.

1.6. Thesis Outline

The thesis is organized into six chapters and have been outlined as follows;

Chapter 1: Introduction

Chapter 1 presents a comprehensive introduction to the research study, commencing with the research background and justification. This is followed by the identification of the research problem and definition of research objectives and questions. Finally, the chapter presents research design, outlining how the research will achieve its objectives.

Chapter 2: Literature Review

This chapter is dedicated to making use of reputable data sources to provide a theoretical background to some of the key and relevant concepts covered in the research study. It commences with a brief review of the conventional measures of traffic congestion, focusing on Highway Capacity Measures (Levels of Services), travel time measures and delay measures in section 2.1. This is followed by a review of the traffic flow theory and relations between traffic flow parameters in section 2.2.

Section 2.3 provides an overview of the hysteresis phenomenon in traffic flow. The chapter goes further to provide a review of the macroscopic observations of traffic networks in section 2.4. This is followed by a review of the key characteristic of macroscopic fundamental diagrams in section 2.5. The chapter goes further to present an overview of MFD estimations in major cities around the globe in section 2.6. A review of the applications of MFD in urban-scale traffic control and monitoring in sub-section 2.7. Finally, section 2.8 will provide a summary and concluding remarks of the chapter.

Chapter 3: Study Area

This chapter provides an introduction of the study area, the City of Cape Town freeway network in South Africa. The chapter commences with the presentation of a brief overview of the geographical, demographic and traffic dynamics of the city. It also provides an overview the city's traffic management and operations dynamics. Finally, a detailed description of the freeways forming part of the study area is provided.

Chapter 4: Research Methodology

This chapter present the methodology adopted in carrying out the research. It presents the research methodological framework in section in 4.1. This is followed by the research data requirement and data collection method in section 4.2. The data limitations are briefly discussed in section 4.3. Lastly, the data analysis method is outlined in section 4.4.

Chapter 5: Results, Analyses and Discussions

This chapter is dedicated to providing discussions and detailed analysis of the whole month and daily MFDs for the CoCT freeway network. The impact of traffic demand on the shape of the daily MFDs will be explored here. This will be followed by an exploration of the implications of partitioning the network on the shape of the MFD. The chapter concludes by proving a summary of the results and discussions.

Chapter 6: Conclusions and Recommendations

This chapter provides the research conclusions and recommendations based on findings in Chapter 5. Section 6.1 of the chapter presents a brief overview of the aims and objectives as set out in Chapter 1. Section 6.2. provides a summary of key findings and conclusions. Finally, based on the research key findings and conclusions, the chapter presents recommendations and future research needs.

Chapter 2 Literature Review

This section of the thesis provides a theoretical background to some of the key and relevant concepts covered in the research study, commencing with a brief review of the conventional measures of traffic congestion in section 2.1. This is followed by a review of the traffic theory & relations between traffic flow parameters in section 2.2. Section 2.3 provides a comprehensive background of the hysteresis phenomenon in traffic flow.

The macroscopic observations of traffic networks are discussed in section 2.4. This is followed by a review of the key characteristic of macroscopic fundamental diagrams in section 2.5. The chapter goes further to present an overview of MFD estimations in major cities around the globe in section 2.6. An overview of the applications of MFD in urban-scale traffic control and monitoring is presented in sub-section 2.7. Finally, section 2.8 provides a summary and concluding remarks of the chapter.

2.1. Conventional Traffic Congestion Measures

2.1.1. Highway Capacity Measures – Level of Service (LOS)

The Highway Capacity Manual makes provisions for principles and procedures for calculating capacity and evaluation of Level of Services (LOS) of arterial roads, highways and intersections. Quantitative indicators such as volume-to-capacity ratio, traffic density, average speed, travel time etc. depending on the type of facility are used to characterize operational traffic conditions. These quantitative indicators are typically converted to Level of Services as shown in the Table 2-1 below (Federal Highway Administration, 1986).

Table 2-1: Level of Services

LOS	Traffic Condition
A	Represents a free flow. Individual users are virtually unaffected by others in the traffic streams. Freedom to select desired speeds and to maneuver within the traffic stream is extremely high.
B	Represents the range of stable flow but the presence of other users in the traffic stream begins to be noticeable. Freedom to select desired speeds is relatively unaffected but there is a slight decline in the freedom to maneuver within the traffic stream from LOS A.
C	Represents the range of stable flow but the selection of speed is affected by the presence of others. Maneuvering within the traffic stream requires substantial vigilance on the part of the user.
D	Represents high-density but stable flow. Speed and freedom to maneuver are severely restricted.
E	Represents operating conditions at or near capacity level. All speeds are reduced to a low but relatively uniform value.
F	Represents forced or breakdown flow

2.1.2. Travel Time Measures

Conventional measures of congestion based on travel time define the levels of congestion with travel time. To quantify the levels of congestion, the cost of travel is explicitly represented by travel time, an important variable in the monitoring of road performance.

Lomax *et al.*, (1997) proposed a travel rate (TR) indicator defined in Equation (2-1) as the amount of time required to travel a unit section length. Travel rate estimates may be compared to an ideal value that depicts the dividing line between unacceptable and acceptable congestion levels.

$$TR = \frac{TT}{SL} \qquad (2\text{-}1)$$

Where TR represent the travel rate, TT represent travel time (in min) and the SL is the section length (in km). Another useful measure is the travel rate index and section travel time by the Texas Transportation Institute (2005).

Travel rate index (TRI) compares peak period travel time and the target traffic conditions (e.g. free flow) as shown in equation (2-2), whereas travel time represents the section travel time weighted with traffic volume (in vehicle-minutes) as shown in equation (2-3).

$$TRI = \frac{TR_a}{TR_t} \qquad (2\text{-}2)$$

$$TT = TR_a \; x \; L \; x \; V \qquad (2\text{-}3)$$

Where TRI is the travel time index, TRa and TRt indicate actual travel rate and target travel rate (in min/km) respectively, L represents the road section length (in km) and V is the vehicle volume (in veh).

2.1.3. Delay Measures

(a) Congestion Threshold Definition

The real impact of congestion is the resultant delay experienced by drivers. Delay is defined as the additional time spend by drivers, in comparison to the free flow or acceptable travel time. For the estimation of delay, several threshold values have been proposed for the start of delay.

Lindley (1987) defined threshold as the volume-to-capacity ratio of 0.77, a number equivalent to the average speed of 88 km/hr. This boundary correlates to the traffic condition Level of Service C-D as shown in the Table 2-1.

Equation *(2-1)* define the travel rate *(TR)* indicator introduced by Lomax, *et al* (1997). The indicator measures the time (in min) required to travel a measured distance. In order to ascertain the threshold for the commence of congestion, the paper introduces the maximum allowable time to travel a road section without having to experience undesirable levels of congestion using desirable travel rates for different roadway categories.

(b) Quantification of Congestion Delay Measures

Delay is defined in Equation (2-4) as the difference between the threshold congestion and the actual travel time (Lomax *et al.*, 1997).

$$D = (TT_a - TT_t)\, x\, V \qquad (2\text{-}4)$$

Where D represents the delay (in veh.min) while TTa and TTt represent actual travel time and target travel time (in min), respectively and V is the traffic volume (in veh).

Total delay for wide-area evaluation is calculated by aggregating delay over a set area. Lomax, *et al.*, (1997) introduced three indicators associated with delay i.e. delay rate DR, relative delay RDR and delay ratio DRA defined by Equations (2-5), (2-6) and (27) respectively, founded upon the travel rate define in equation (2-1).

$$DR = TR_a - TR_t \quad (2\text{-}5)$$

$$RDR = \frac{DR}{TR_t} \quad (2\text{-}6)$$

$$DRA = \frac{DR}{TR_a} \quad (2\text{-}7)$$

Where TRa and TRt indicate the acceptable and actual travel rate (in min/km), respectively. Delay rate measures the levels of congestion as the additional time experienced by motorists above the tolerable level of congestions.

The delay ratio and the relative delay rate are the normalized delay rate, which can be used to compare congestion levels across various corridor categories. The delay ratio is used to compare the congestion levels to the actual travel time, whereas the relative delay ate is the comparison of congestion levels to the level of acceptable congestion.

In summary, though the definition, measurement and quantification of levels of congestion for the City of Cape does not form part of the scope of this research, it is however important to understand and put into perspective the potential role that the MFDs will play in this regard. MFDs have been proven to be effective measures of congestion on cities' urban-scale networks. Sub-sections 2.2, 2.3 and 2.4 explain this phenomenon in detail.

2.2. Traffic Flow Theory

2.2.1. Historical Perspective

This section of the dissertation aims to provide a brief historical perspective on the traffic flow theories as an introduction to the concept of Macroscopic Fundamental Diagrams. Traffic flow theories aim to describe the interactions and relations among drivers, vehicles and the surrounding infrastructure.

Traffic flow theories are indispensable and integral elements of all traffic modelling and analysis used in the design and functioning of highways. Scientific studies and research of traffic flow dating back to the 1930s, commenced mainly with the implementation of probability theory to the characterization of road traffic (Adams, 1936) and with the ground-breaking studies carried out by Greenshields (1935) on the study of traffic models comparing speed and volume and the investigation of intersections' traffic performance.

The culmination of World War II was followed by unprecedented rise in the use and ownership of automobiles and the subsequent further development and improvement of the highway system and infrastructure. These developments led to an increase in the research and study of characteristics of traffic. To date, the body of scientific knowledge around traffic-flow theory continues to increase and easily characterized through advanced computation technology.

The 1950s saw advancements in the theoretical developments based on different types of approaches including traffic wave theory, car following theory and queuing theory. Research studies during that period included works by Wardrop (1952) and Herman (1992). It was by 1959 that the developments in traffic flow theories appeared desirable to hold an international symposium (Gartner, *et al.,* 1992).

In late 1959, the first “international symposium theory of traffic flow” was held in Michigan. The symposium was followed by a “series of triennial symposia on the theory of traffic flow and transportation” (Gartner, *et al.*,1992). Since then, several other symposia conferences are being held focusing on a variety of relevant and related topics

Even though over the years, the field of traffic flow theory has diffused into many other sub-topics, the fundamental principles of the theory are just as important today as they were in the times past. These principles remain the backbone of many of the theories, procedures and techniques that are being used in the planning, design, implementation of modern transportation systems (Gartner, *et al.,* 1992).

2.2.2. Traffic Flow Parameters

As discussed in the previous sub-section 2.2.1, though the traffic flow theory has evolved and diffused into many other sub-topics, the fundamental principles remains the same and form the basis of many of the theories being applied today in modern transportation systems.

Today, traffic flow theory is applied by traffic engineers to characterizes traffic flow. At any given time, there are thousands of vehicles on roadways. The interactions among these vehicles has an impact on the overall traffic movement. These fundamental principles of traffic flow theory are described by the following parameters;

Speed (v)

The vehicle average speed is calculated by dividing the vehicle distance travelled (in km) by the average time travelled (in hrs.). In most cases, the vehicle travelling on a given road section will likely travel at a varying speed from other vehicles in the vicinity.

Density (k)

Traffic density is defined as the total number of vehicles observed on a given length of the roadway (in veh/km). High density signifies a condition where the distance between individual vehicles is short, whereas low density signifies longer distance between individual vehicles.

Flow (q)

Traffic flow is defined as individual vehicles' rate at which they pass through a given point and usually expressed in vehicles per hour. It is calculated by multiplying the measured density by the measured travel speed using equation (2-8).

2.2.4. Speed-Flow-Density Relationship

The relationships between density and speed are easily observed in experimental space, whereas the effects of density and speed on flow are not as clear. Equation (2-8) below can be used to define the relationship between speed, flow and density, given uninterrupted traffic flow conditions.

$$v = \frac{q}{k} \qquad (2\text{-}8)$$

Where q is the traffic flow (in veh/hr), v is the speed (in km/hr) and k represents density (in veh/km).

This relationship between density, flow and speed represent the basic equation of traffic flow. Provided that the relationship between two of the three variables is established, Equation (2-8) can be used to establish the relationship of the third variable. According to the definition, traffic volume is required to be zero when density is zero and when traffic reaches maximum density.

The traffic flow is equal to zero when speed and/or density is zero. Two common traffic conditions are used to illustrates these points; (a) traffic jam condition characterized by very low speeds and very high densities. Second condition is observed at very low densities, and motorists can travel in free flow speeds. The low density makes up for the high free flow speed, resulting in very low traffic flows.

2.2.5. Greenshield's Macroscopic Stream Model

Greenshields (1935) carried out ground-breaking studies of traffic models comparing speed and volume and investigating the performance of traffic at intersections. Greenshields (1935) assumed that, given uninterrupted traffic flow conditions, density and speed are linearly related.

Greenshields (1935) expressed this relationship mathematically and graphically by Equation (2-9) and Figure 2-1, respectively.

$$v = u_f - \frac{u_f k}{k_j} \qquad (2\text{-}9)$$

Where v is the mean speed (in km/hr) at density k (in veh//km), k_j and u_f are jam density and free flow speed, respectively. This equation (2-9) is often referred to as Greenshields model (Greenshields, 1935). It shows that when density becomes zero, speed approached free flow speed (i.e. $v \longrightarrow u_f$, when $k \longrightarrow 0$).

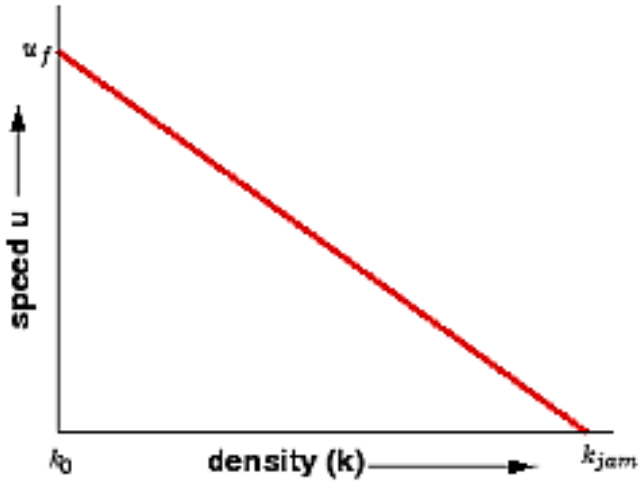


Figure 2-1: Speed and Density Linear Relationship

Substituting Greenshields's speed and density relation (see Equation (2-9)) into Equation (2-8) results in the following mathematical equation;

$$q = \left(k_j - \frac{u_f\, k}{k_j}\right) k \qquad (2\text{-}10)$$

Where q is the traffic flow (in veh/hr), v is the speed (in km/hr) and k represents density (in veh/km). Figure 2-2 below illustrates this relationship between density and flow.

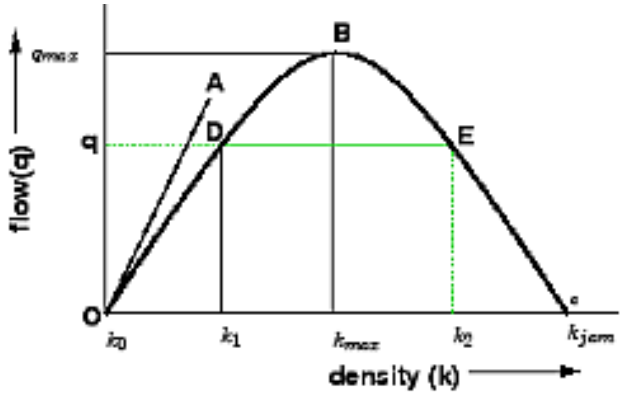


Figure 2-2: Flow - Density Relationship

The observed relationship between density and flow makes it possible to determine the density at which flow is maximized, as shown below;

$$\frac{dq}{dk} = k_j - \frac{2u_f\, k}{k_j} \qquad (2\text{-}11)$$

$$\frac{dq}{dk} = 0, \text{yields } k = \frac{k_j}{2} \qquad (2\text{-}12)$$

Substituting the maximum value of k (see Equation (2-12)) into Equation (2-9) yields the maximum flow;

$$q_{max} = u_f\left(\frac{k_j}{2}\right) - \frac{u_f}{k_j}\, x\, \frac{k_j^2}{4}$$

$$q_{max} = \frac{k_j u_f}{4} \qquad (2\text{-}13)$$

Therefore, the maximum traffic flow is 25% of the product of jam density and free flow speed. In order to get the speed at the maximum flow, $u0$, Equation (2-13) is substituted in Equation (2-9) to get the following Equation (2-14), showing that the speed at maximum flow is equivalent to half the free flow speed.

$$u_0 = u_f - \frac{u_f}{k_j} \; x \; \frac{k_j}{2}$$

$$u_0 = \frac{u_f}{2} \qquad (2\text{-}14)$$

The relationship between speed and flow is established by substituting the density k (see Equation (2-8)) into Equation (2-9) to yield the following mathematical relation (see Equation (2-15) and graphical representation in Figure 2-3.

$$q = k_j u - \frac{u^2 k_j}{u_f} \qquad (2\text{-}15)$$

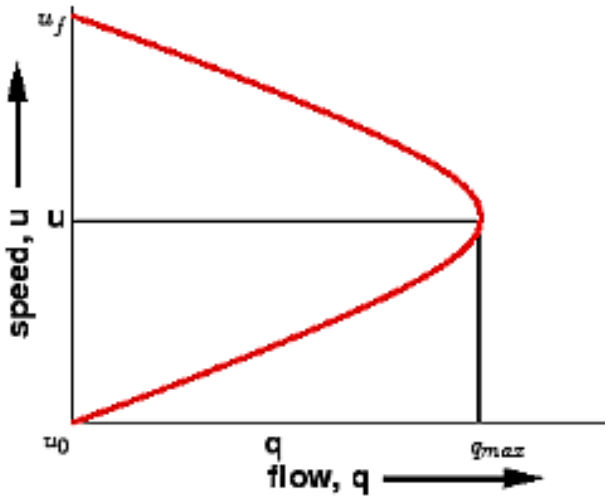


Figure 2-3: Speed and flow relationship

Greenshields (1935)'s macroscopic stream model can be summarised into the following points;

a) When traffic density is zero, it means there are not vehicles on the roadway and therefore traffic flow is zero.
b) When the density increases, the flow also increases up to a maximum flow (at critical density)
c) When the density reaches the maximum (jam density), the flow is zero (parking lot condition)

2.2.5. Shock Waves

Another key component of traffic flow theory is the shock wave phenomenon that is observed in traffic flow. A shockwave propagates along a platoon of vehicles in reaction to disturbed traffic conditions. Changes in conditions may result from a variety of external factors including collisions and sudden speed increasing resulting from entering free flow conditions. The following equations (2-16) is used to calculate the velocity at which the shock propagates in response to changing conditions;

$$v_{sw} = (q_B - q_A)/(k_B - k_A) \qquad (2\text{-}16)$$

Where v_{sw} is the shockwave's propagation velocity (in km/hr), qB is the traffic flow conditions before the disturbance (in veh/hr), qA is the traffic flow conditions after the disturbance (in veh/hr), kB is the conditions of the traffic density before the disturbance (in veh/km) and kA is the conditions of the traffic density after the disturbance (in veh/km). The relationship between these variables can also be explained by the following flow-density curve (see Figure 2-4).

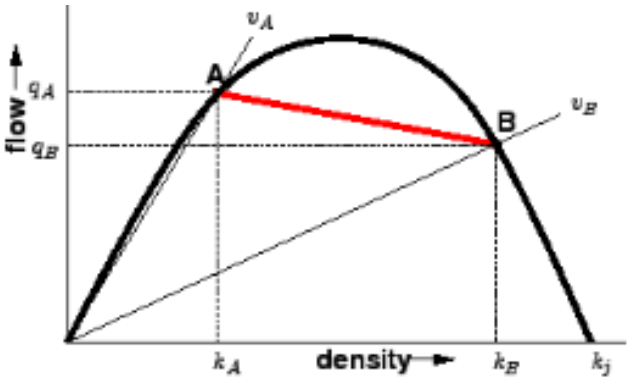


Figure 2-4: Shock wave, flow-density curve

Positive shock wave velocity implies that the shock wave is travelling in the same direction, whereas a negative velocity implies that the shock wave is moving in the opposite direction.

2.3. Understanding Hysteresis Phenomenon

This section of the chapter presents a review of the hysteresis phenomenon, starting with an overview of microscopic observations of hysteresis phenomenon in traffic flow in sub-section 2.3.1. This is followed by a review of the macroscopic observations of hysteresis in freeway networks in sub-section 2.3.2. The section goes further to look at some of the characteristics of hysteresis loops observed in urban-scale freeway networks in sub-section 2.3.3

.

2.3.1. Hysteresis in Traffic Flow – Microscopic Observations

Congestion and hysteresis phenomenon in traffic flow remain a key focus area in the traffic operations literature. According to Laval (2009), when a group of vehicles in a traffic stream experiences a disturbance, the conditions before the disturbance are not reinstated instantly; the delay in reclaiming speed when vehicles exit the traffic disturbance is referred to as traffic hysteresis.

Previous research makes a connection between the hysteresis phenomenon and the asymmetry seen in the deceleration and acceleration diagrams (Herman and Potts, 1961; Newell, 1965). Edie (1961) Introduced a traffic flow model that presupposed that hysteresis is seen during a change from traffic congestion state to free flow state.

Treiterer and Myer (1974) observed the hysteresis phenomenon using aerial photographs. They argued that the observed hysteresis phenomenon was as a result of driver's behaviour when approaching and exiting the traffic disturbance, resulting in the observed behavioural asymmetry. The observed asymmetry was explained and described using the deceleration and acceleration rates. However, later results by Maes (1979) did not support the two loops that were observed.

Zhang (1999) developed mathematical models to formulate hysteresis by introducing a model for explaining the change between various regimes of traffic flow based on the behaviour of the driver. This development was followed by Zhang and Kim (2005)'s proposal of a "car-following theory for multiphase vehicular traffic flow" which showed that models can be used to describe the hysteresis, capacity drop and the waves observed during a shift from congested to free flow regime.

Recently, Laval (2009) used Edie (1963)'s definition for volume, speed and density to analyse the congested data and identified the following hysteresis levels founded upon the difference between the traffic volumes in and traffic volumes of the hysteresis at given densities; (a) strong hysteresis – observed when the difference between the volumes in and volumes out of the hysteresis is greater than 300 veh/hr, (b) weak hysteresis – observed when the difference is not more than 300 veh/hr, (c) negligible hysteresis – observed when the difference is insignificant and (d) negative hysteresis when acceleration branch of the hysteresis is above the deceleration.

Laval (2010)'s went further to analyse the above findings and found out that varying driver behaviour during acceleration and deceleration may have a strong impact in defining the type of hysteresis loop. The findings showed that the aggressive and timid driver behaviour of a following vehicle results in counterclockwise, negative hysteresis loop on the volume-density diagram (see Figure 2-5).

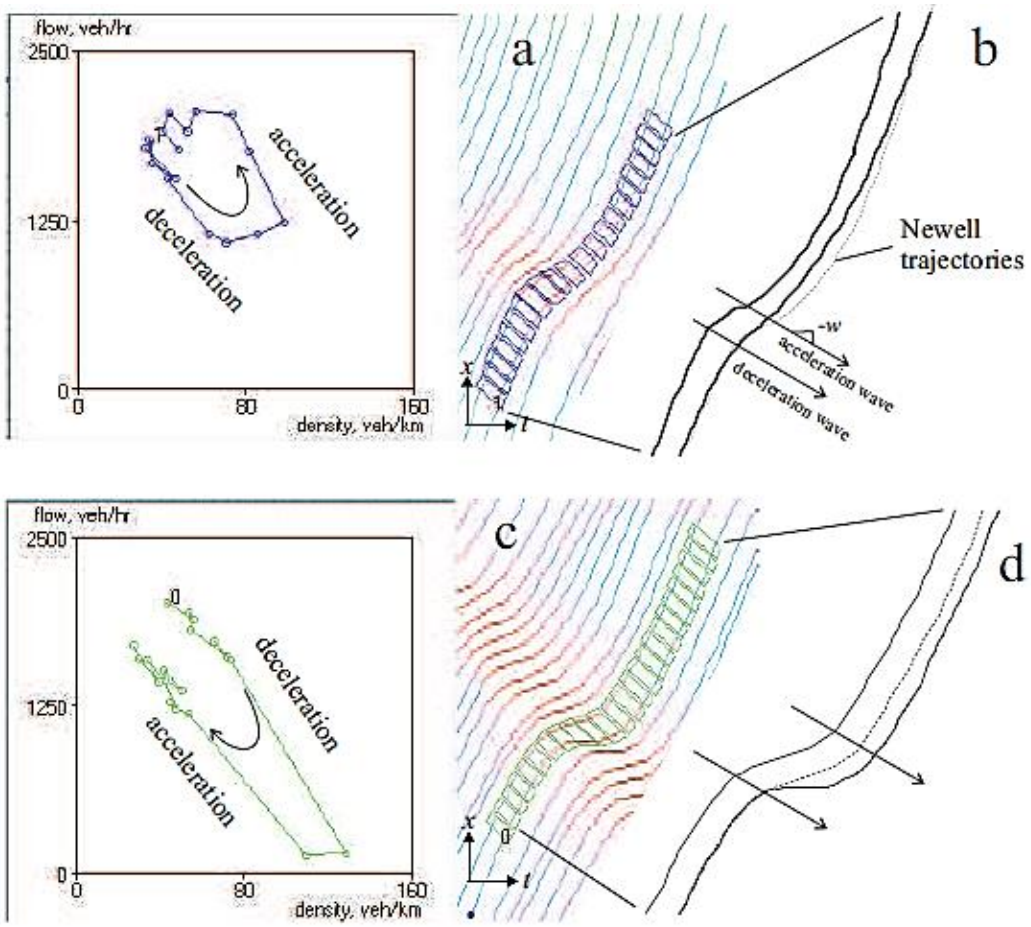


Figure 2-5: Trajectory data's illustration of how showing how timid and aggressive driver behaviour results in negative (a-b) and positive (c-d) hysteresis loops.

Source: (Laval, 2010)

Although the formation of hysteresis in traffic flow is well studied, its macroscopic characteristics such as the magnitude (weak or strong) and type (clockwise or counterclockwise) have thus far been related statistically only to the volume difference between the onset and the end of the phenomenon. Section 2.3.2 goes further to provide macroscopic observations of hysteresis in freeway networks. The characteristics of the hysteresis observed in urban-scale networks will be covered in section 2.3.3.

2.3.2. Hysteresis in Freeway Networks – Macroscopic Observations

Research has shown that the scatter seen in freeway MFDs follows a hysteresis pattern (Buisson and Ladier, 2009; Gayah and Daganzo, 2011; Saberi and Mahmassani, 2012). Buisson and Ladier (2009) used data obtained from the Toulouse freeway network's loop detectors and observed MFD with hysteresis-like pattern when there was a significant disturbance of the network traffic.

Gayah and Daganzo (2011) demonstrated that the hysteresis pattern in MFDs are observed when the network experiences traffic disturbance. They managed to use the two-bin model to establish that the hysteresis patterns are unlikely to form when drivers can change their routes adaptively in order to avoid congestion. Geroliminis and Sun (2011b) used empirical data obtained from Minneapolis, Minnesota freeway networks' loop detectors to establish that the derived MFD was not well-defined in the same way as the Yokohama MFD, this due to the presence of hysteresis effects.

Recently, Saberi and Mahmassani (2012) proposed a characterization of hysteresis pattern that is path-dependent in freeway networks. They used data obtained from Oregon freeway network's loop detectors to study the temporal and spatial distribution of congestion's impact on the MFD properties, with focus on the properties of the hysteresis patterns and conditions under which they form. Furthermore, Gayah and Daganzo (2011) had earlier characterized the hysteresis loops by whether they are clockwise or counterclockwise.

In summary, the above findings of macroscopic observations of hysteresis loops in freeway networks give rise to a need for further confirmation, characterization and interpretation of the hysteresis phenomenon in freeway networks. Further characterization and interpretation of the hysteresis loops in the following subsection 2.3.3 will provide better understanding of the underlying principles that resulted in the formation of the observed patterns.

2.3.3. Characterization of Hysteresis in Urban-Scale Freeway Networks

As established in the previous sub-section 2.3.2, there is a need for further confirmation and characterization of the MFD's hysteretic behavior in freeway networks. This further detail analysis of the hysteretic behavior will allow for better understanding of the underlying principles and strategies that can be devised to control them. Against this backdrop, this sub-section goes a level deeper to focus on the empirical characterization of hysteresis loops and capacity drop observed in freeway networks for the cities of Chicago, Portland and Irvine in the United States (Saberi and Mahmassani, 2013).

(a) Comparison: Freeway Network Fundamental Diagram

Saberi and Mahmassani (2013) used loop detector data to conduct a study aimed at characterizing the hysteresis and capacity drop phenomenon for the freeway networks of Chicago (Illinois), Portland (Oregon) and Irvine (California) as shown in Figure 2-6 below. The observed MFDs for the three networks were compared in order to explore the effects of variations in network topology and size on the properties of the MFD.

Figure 2-6: Study locations: (a) Chicago freeway subnetwork, (b) Portland freeway network, and (c) Irvine freeway network (maps in different scales).

Source: (Saberi and Mahmassani, 2013)

Saberi and Mahmassani (2013) first investigated whether Buisson and Ladier (2009), Geroliminis and Sun (2011b) and Saberi and Mahmassani (2012)'s research findings and observations hold for other freeway networks. Figure 2-7 illustrates that for the Portland freeway network's MFD, a consistent and well-defined relationship is not observed due to the hysteresis phenomenon. Moreover, the observed hysteresis and scatter pattern were found to differ in size and shape.

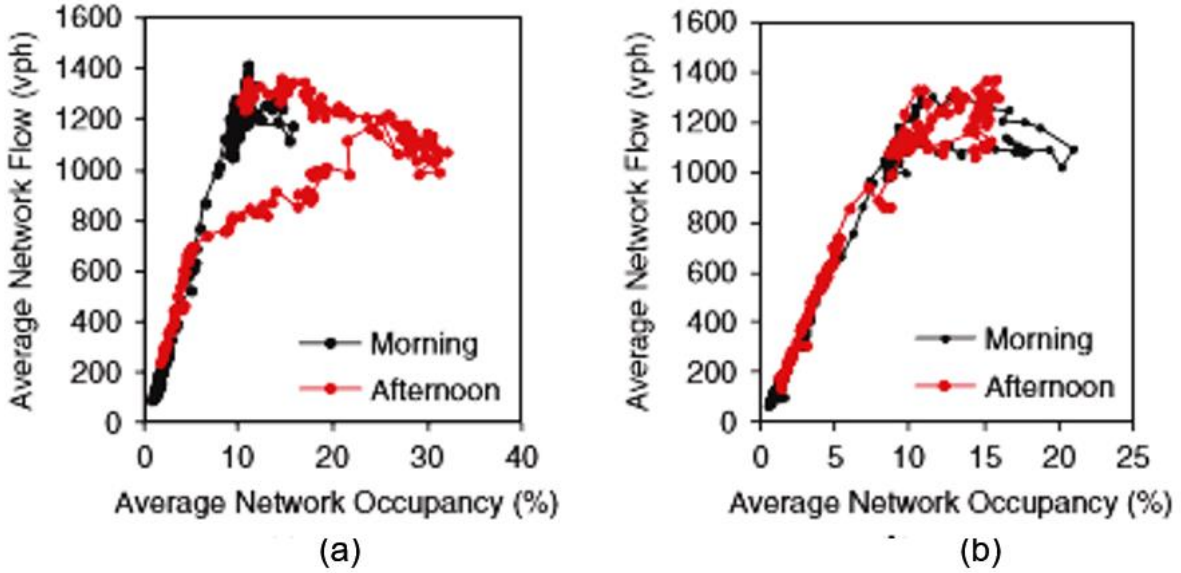


Figure 2-7: Portland freeway network's MFD for the; (a) 22nd April (Friday), (b) 25th April (Monday) 2011.

Source: (Saberi and Mahmassani, 2013)

Figure 2-8 shows the MFD for the Irvine freeway network for the 6th and 16th February 2009. Similarly, Figure 2-9 shows network wide MFD for Chicago freeway network for the 4th and 5th August 2009. The comparative analysis of the observed MFDs across the three networks with varying size and topology showed that the freeway networks that are congested are unlikely to produce low-scatter and well-defined MFDs and that hysteresis loops are observed in freeway networks. In addition, the results show that the hysteresis phenomenon does not follow a similar repetitive pattern but differ across the networks. on how congestion is spatially spread, freeway networks produce hysteretic patterns that are not consistent in size and shape.

In agreement with Saberi and Mahmassani (2012) and Geroliminis and Sun (2011b)'s findings, these results confirm that at a given average occupancy, the average flow is lower during recovery phase than during the loading phase for every cycle.

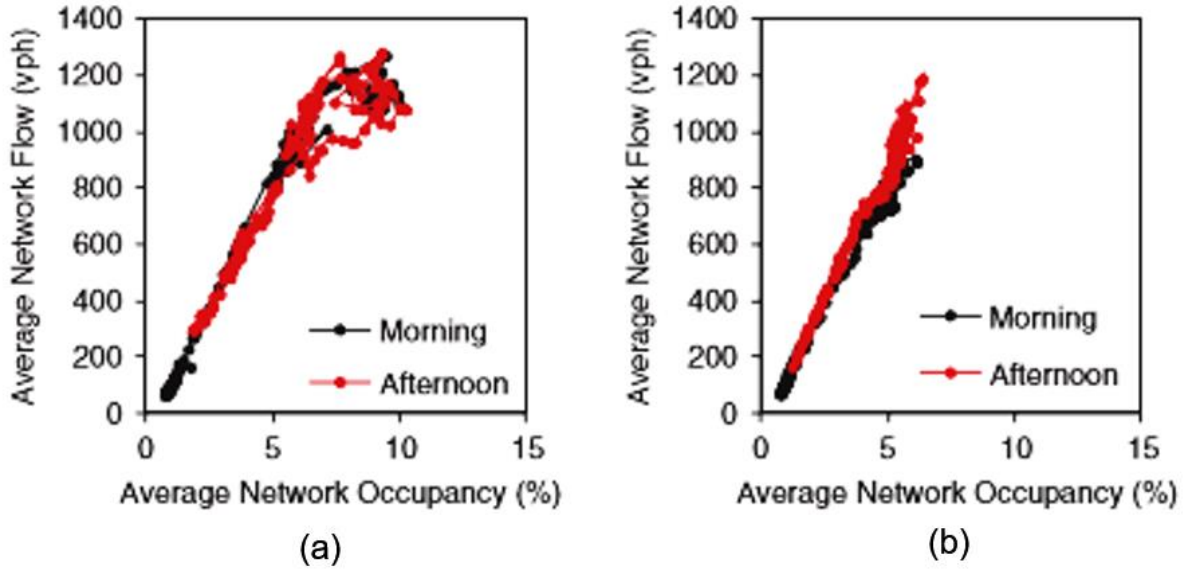


Figure 2-8: Irvine freeway network's MFD for the: (a) 6th February (Friday) and (b) 16th February (Monday) 2009.

Source: (Saberi and Mahmassani, 2013)

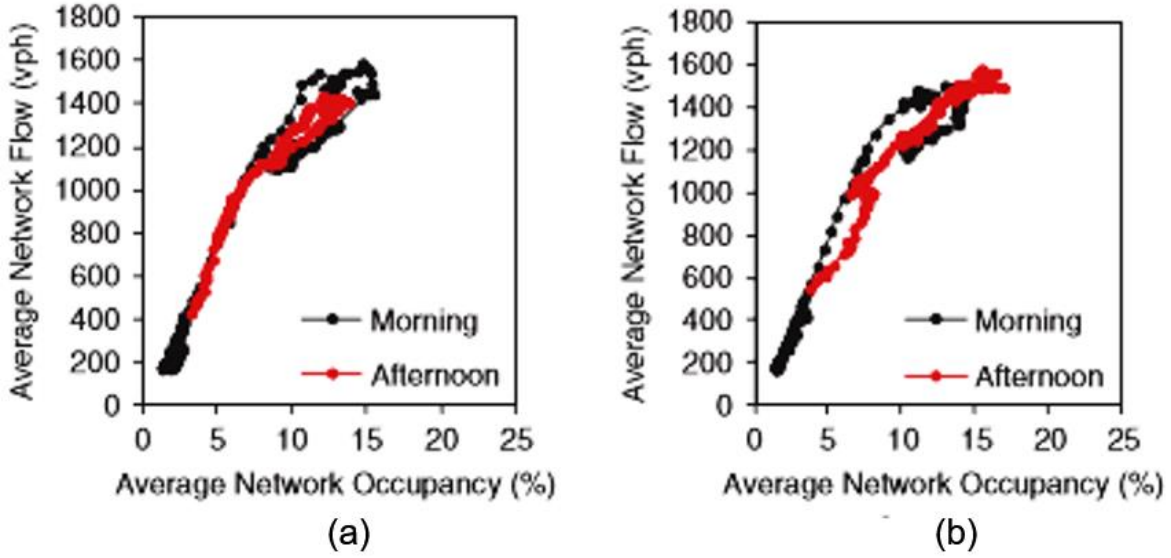


Figure 2-9: Chicago freeway network's MFD for the; (a) 4th August (Tuesday) and (b) 5th August (Wednesday) 2009.

Source: (Saberi and Mahmassani, 2013)

In addition, the hysteresis phenomenon observed in the MFDs often exhibit, consecutively, morning and afternoon hysteresis loops of the loading-recovering cycles. Saberi and Mahmassani (2013) showed that the network morning loading recovery cycle sometimes produces a hysteresis loop that is not complete and often open at the end of the morning recovery phase, whereas the afternoon hysteresis loop is often open at the beginning of the loading phase.

These observations confirmed Geroliminis and Sun (2011b)'s statement that the "evolution of traffic states follow different paths during the onset and offset of the morning and evening peaks". Figure 2-10 illustrates the unloading-reloading hysteresis in freeway networks.

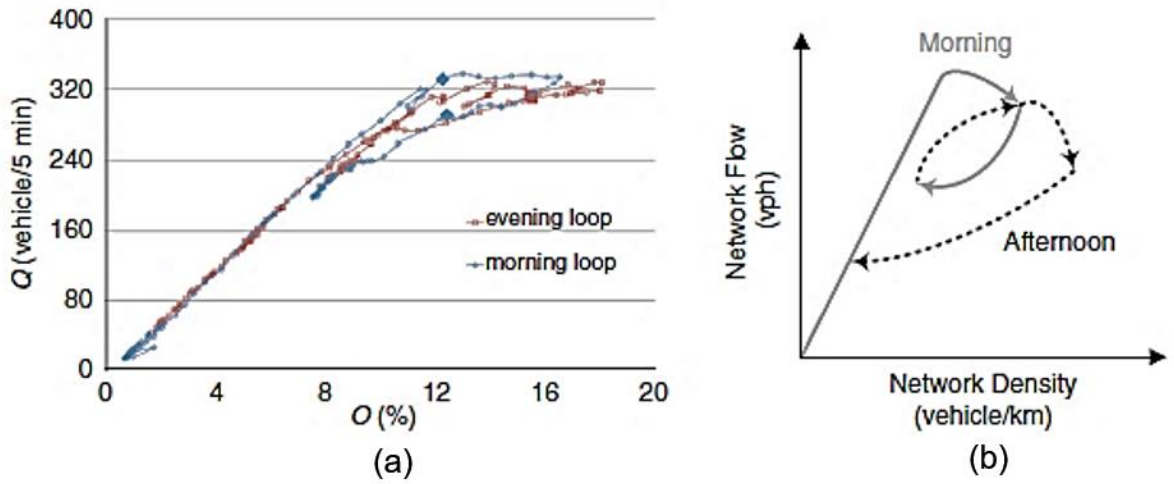


Figure 2-10: (a) Minneapolis freeway network's MFD for the 22nd May 2007, (b) demonstration of unloading-reloading hysteresis.

Source: (Saberi and Mahmassani, 2013)

(b) Shapes of Clockwise Hysteresis Loops

The most common hysteresis in MFDs of freeway networks are clockwise hysteresis. Saberi and Mahmassani (2013) identified two types of these clockwise hysteresis loops and named them type H-1 and type H-2.

The type H-1 loop is observed when there is an unstable recovery of a network following a congestion peak period in which there is heterogeneous distribution of traffic during the recovery and the average flow decreases constantly with decreasing average occupancy. The MFD exhibit the type H2 hysteresis loop during network recovering when the average flow remains unchanged when the average occupancy declines.

Figure 2-11 illustrates an example of type H-1 and H-2 clockwise hysteresis observed in the Portland freeway network. Table 2-2 presents a summary of the types of clockwise hysteresis loops observed during morning and afternoon peaks across the different freeway networks. The results show that the type H-1 hysteresis loops are observed more than the type H-2.

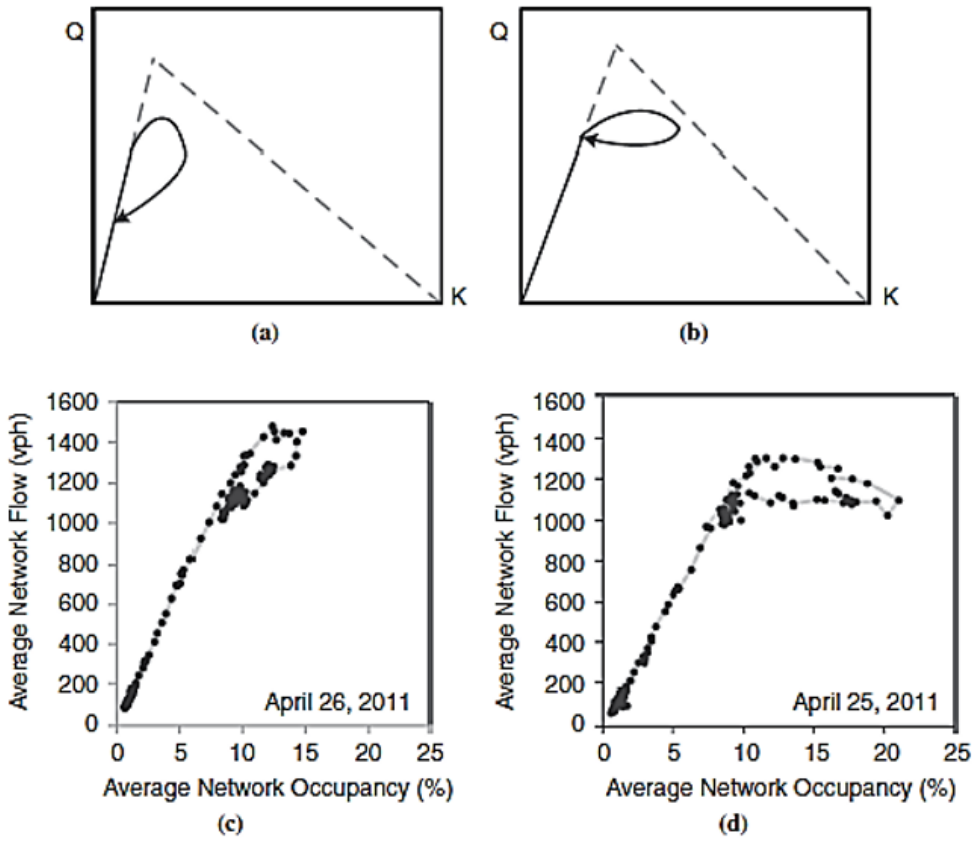


Figure 2-11**:** Hypothetical (a) Type H1 & (b) Type H2 hysteresis loops, in network-wide flow–density plane, and the observed (c) Type H-1 and (d) Type H2 hysteresis loops both exhibited by the Portland freeway network's MFD.

Source: (Saberi and Mahmassani, 2013)

Table 2-2: Classification of Observed Hysteresis Loops

Network Location	Date	Shape of Hysteresis Loop	
		Morning	Afternoon
Portland	22 April 2011	H1	H1
	25 April 2011	H2	H2
	26 April 2011	H1	H1
	29 April 2011	H1	H2
Irvine	6 February 2009	H1	H1
	7 December 2009	H1	H1
Chicago	4 August 2009	H1	H1
	5 August 2009	H1	-

(c) Quantifying the size of the hysteresis loop

Another identified key characteristic of the hysteresis loop was the size of the loop which could be defined by its height and width. Equation (2-17) shows that the size of a loop (SH) is described by the ordered pair of its width (change in average occupancy measured in veh/km/lane) and height (change in average flow measured in veh/hr/lane).

$$S_H = (\Delta K, \Delta Q) \qquad (2\text{-}17)$$

Figure 2-12 (a) illustrates the hypothetical loop with the width (ΔK) and height (ΔQ), whereas Figure 2-12 (b) shows that the size of the hysteresis loop observed on the 25th April 2011 can be expressed as (7.7, 178). The height ΔQ of a hysteresis loop signifies the level of instability in the network and measures of efficient is the recovery phase relative to the loading phase.

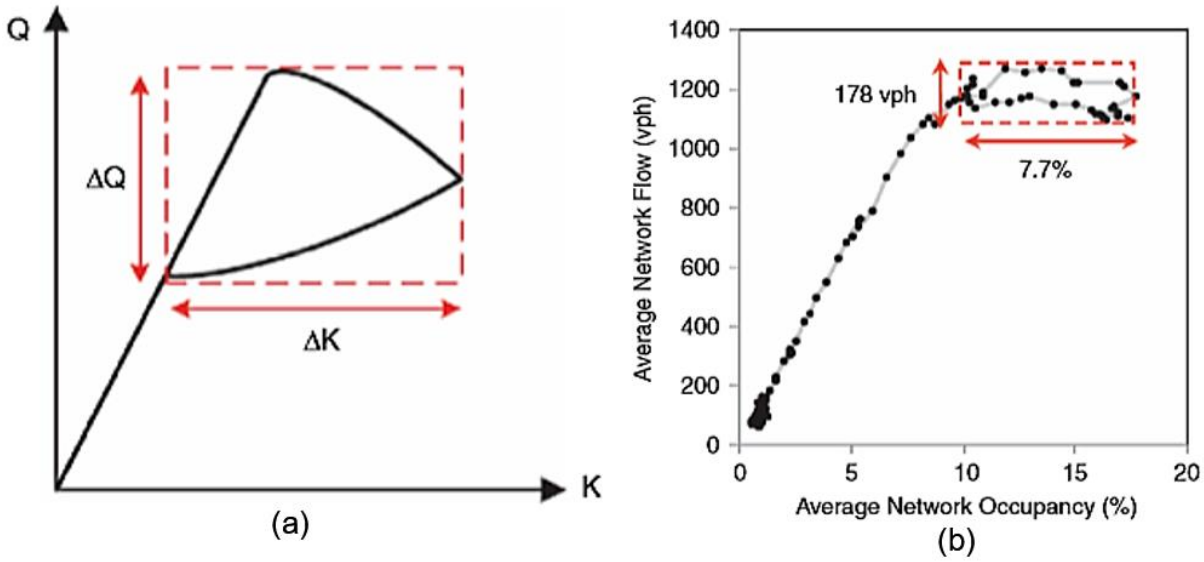


Figure 2-12: Size (a) Hysteresis loop hypothetical size expressed in terms of its width (ΔK) and height (ΔQ), and (b) the size of the hysteresis loop observed in the MFD of the 24th April 2011.

Source: (Saberi and Mahmassani, 2013)

Larger ΔQ signifies relatively unstable network traffic and more heterogenous distribution of congestion during the recovery phase. Hysteresis loops with smaller ΔQ represents a more homogenous and stable recovery. Earlier, Saberi and Mahmassani (2012) proposed a possible link between heterogeneity of congestion distribution and the size of a hysteresis loop. Saberi and Mahmassani (2013) showed that there are instances from the Portland freeway network where there is a correlation between the size of the hysteresis loop and the standard deviation of the average occupancy during the recovery phase.

(d) Capacity Drop in Freeway Networks

Saberi and Mahmassani (2013) confirmed that the capacity drop phenomenon exists in freeway networks, the same way it does in freeway segments. They showed that this phenomenon often results from the instability of the network traffic. There are two types of capacity drops that have been identified.

Capacity drop type-1 is seen when the freeway network is not able to sustain its flow at its maximum for a longer period, causing the capacity to drop while the network is still loading, and the demand is still high. Capacity drop type-2 is seen when network is reloading after an incomplete recovery from the initial loading causing an instability in the network.

Figure 2-13 demonstrates occurrence of capacity drop type-1. Figure 2-13(a) shows that the capacity of the Chicago freeway subnetwork on the 4th August 2009 could not sustain itself for too long before it started dropping from 1550 veh/hr/lane to 1450 veh/hr/lane.

The initial observed capacity drop didn't last long before it reinstated and reached the 1550 veh/hr/lane level again after which another significant drop was observed. Figure 2-13 (b) also shows that the Portland freeway network drops capacity from 1400 veh/hr/lane to 1000 veh/hr/lane while the network is still loading.

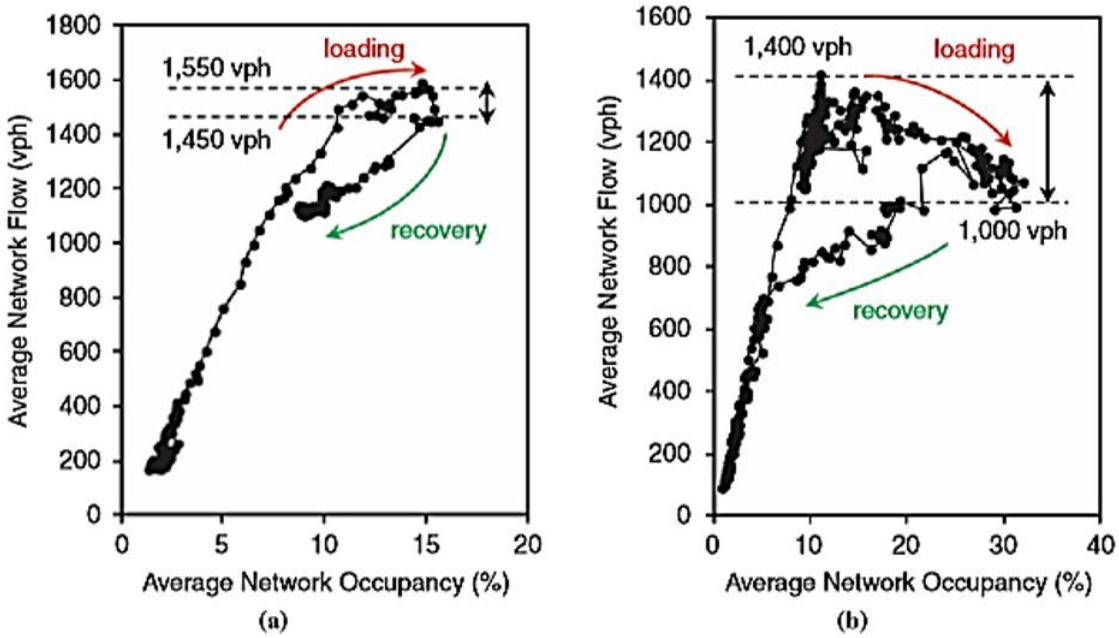


Figure 2-13: An example of capacity drop type11 observed in the MFD of (a) Chicago freeway network on the 4th August 2009 and (b) Portland freeway network observed on the 22th April 2011

Source: (Saberi and Mahmassani, 2013)

Figure 2-14(a) illustrates the zoomed-in portioned of the observed morning and afternoon MFDs of the Chicago freeway subnetwork on the 4th August 2009. The capacity drop type 2 is observed. On same day, the network morning capacity was approximately 1560 veh/hr/lane and dropped to 1420 veh/hr/lane in the afternoon. In the same way, Figure 2-14 (b) illustrates the zoomed-in portion of the Portland freeway network's observed MFD for the morning and evening peak periods on 28th April 2011. The capacity drops from 1430 veh/hr/lane in the morning to 1293 veh/hr/lane in the afternoon.

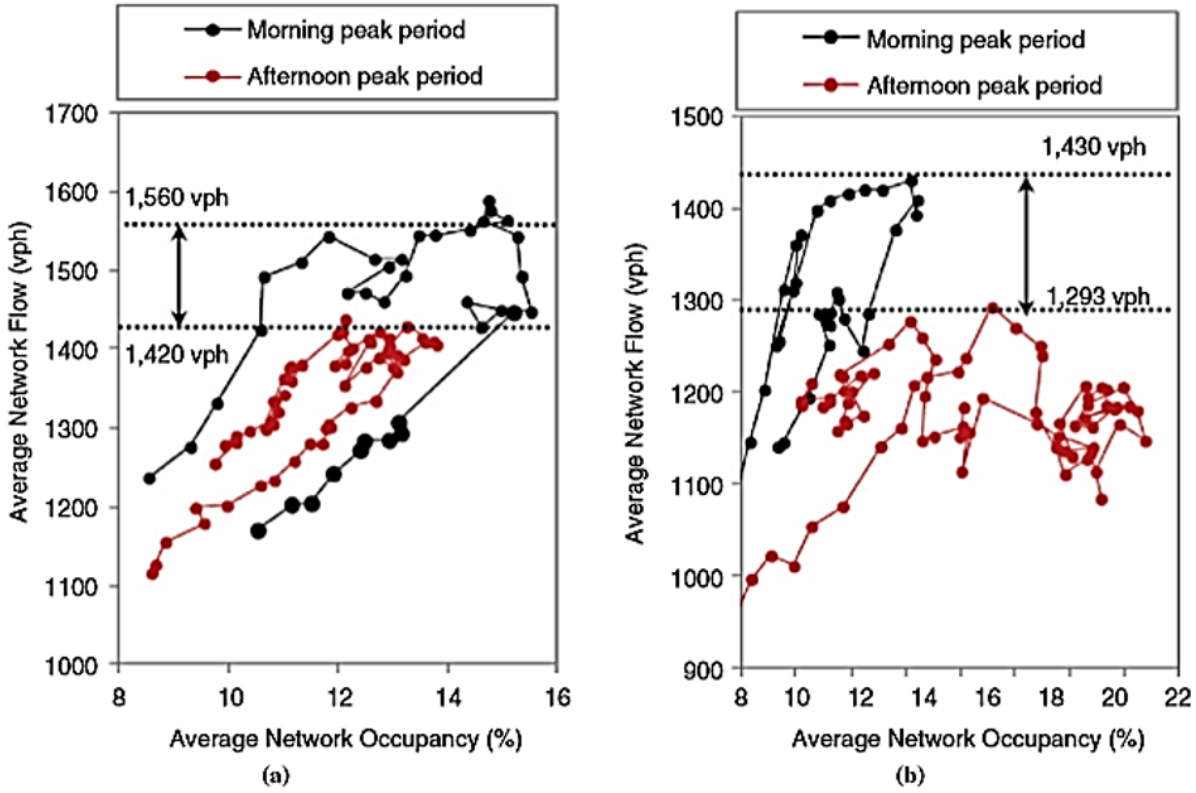


Figure 2-14: An example of capacity drop type-2 observed in the MFD of (a) Chicago freeway network on the 4th Aug 2009 and (b) Portland freeway network observed on the 22th Apr 2011

Source: (Saberi and Mahmassani, 2013)

More analysis of the Portland freeway network capacity drop observed on the 28th April 2011, revealed that the average flow observed during the afternoon peak period couldn't reach the maximum value in the morning.

2.4. Traffic Network Observation on Microscopic Level

The previous section provided a brief overview of the hysteresis phenomenon. This section provides a brief review of earlier research work with empirical observations of traffic networks in subsection 2.4.1, the two-fluid model in subsection 2.4.2 and the recent development on "input-output" description of the system.

2.4.1. Earlier Empirical Observations

During the past 50 years, there has been numerous research studies in the area of network wide performance evaluation. Most of these research studies were empirical in nature and observed traffic characteristics and used formulae to explain their relationships on macroscopic level. These studies were able to reveal dependencies among various network traffic variables. Their findings led to the development of various macroscopic models.

Wardrop (1952) and Smeed (1967) earlier developed macroscopic models to cater for streets and later adjusted models to cater for wide networks. It was Smeed (1967) who introduced a function relating the city area, the traction area and number of vehicles able to enter the city's centre. The function was used to model various traffic conditions for various cities. Although the model output suggested that the average speed affects the vehicle numbers which can enter the network, it did not manage to explain it for overcrowded cities.

Thomson (1967) used data he collected over a period of 14 years in the city of London and discovered that the speed and flow function was decreasing linearly. Wardrop (1952) had earlier introduced another linearly decreasing function between average flow and speed by taking into account average signal spacing and average road width.

Zahavi (1972) described how average speed, road density and the traffic intensity are related to each other. His study revealed inverse proportion relationship between flow and speed. This monotonic model was found not suitable to describe the performance of a highly congested road network. Godfrey (1969) proposed the first non-monotonic model which also required further validation because its estimations were based on few data observations.

2.4.2. Two-Fluid Model

A more generic model, the two-fluid model was later introduced primarily to set the principles and theoretical background of macroscopic traffic conditions in urban networks. The two-fluid model was applied more widely than the other models discussed in 2.4.1.

Herman and Prigogine (1979) developed a macroscopic model on the premise that the network traffic can be split into two parts; (a) one corresponding to stopping vehicles as a result of traffic control measures, incidences etc. and (b) the other to moving vehicles. The model was developed and proposed to define and characterize the urban network traffic by linking stop time and the trip time.

Herman and Prigogine (1979) model was later further enhanced and validated using empirical data obtained from different cites in the US. The influence and impact of the driver type to the two-fluid model parameters was investigated and findings revealed that the model is sensitive to driver behaviour (normal, conservative and aggressive) (Herman, *et al.,* 1988).

Mahmassani *et al*., (1987) made a notable observation concerning the urban-network level relationship between density, flow, and speed. The observations were in line with the framework of the two-fluid model. They discovered that the relationships among the traffic variables on a network level is similar to the one for individual road sections.

Mahmassani *et al*., (1987) verified the fundamental relationship as defined by Equation *(2-8)*, for urban-scale aggregated variables. Mahmassani et al., (1987) carried out a comparison between the simulated values and Herman *et al.,* (1988)'s derived fomalae. The simulated values were found to be fitting well with the emulated values.

These findings were found to be consistent with the two-fluid model as being a theoretical base to characterise the state of an urban network. These research studies by Herman *et al,.* (1979), Herman *et al,.* (1988) and Mahmassani et al,.(1987) complemented each other well to prove that the two-fluid model can be used to characterize the state of traffic condition on a network level.

However, the model's validity under different demand patterns was its key limitation (Daganzo, 2007). In other words, it was sensitive to exogenous conditions such as Origin-Destination (OD) patterns. A suitable or ideal model would be required to be insensitive to exogenous conditions for the robust and reliable monitoring and control of the urban network.

2.4.3. Recent Observations of Traffic Networks

Recently, Daganzo (2007) proposed the use of macroscopic variables under an "inputoutput" description to model an urban gridlock. By relating vehicle accumulation and the rate at which vehicles exit the system (exit flow rate), the reference indicated a system description and control based on accumulation for gridlock dynamics and peak traffic. The proposed concept was argued to hold for homogenously congested networks.

A year later, Geroliminis and Daganzo used empirical data to confirm the macroscopic relationship between traffic 'production' and 'accumulation' (Geroliminis and Daganzo, 2008). The estimation of MFD was confirmed to consist of three regimes i.e. free flow regime, critical regime and congested regime as seen in the fundamental diagrams for single road segments (Geroliminis and Daganzo, 2008).

Microscopic simulation studies conducted in San Francisco and empirical studies in Yokohama revealed that a well-defined Macroscopic Fundamental Diagram exists for cities that are highly congested but the relationship between flow and density for individual links shows scattered plots (Geroliminis and Daganzo, 2008). In addition, the studies showed that the relationships remain constant for various traffic demand scenarios, suggesting that MFD are independent of network traffic demand.

2.4.4. Summary of Macroscopic Observations of Traffic Networks

Table 2-3 below summarises, in a chronological order, the development of macroscopic observations of traffic networks over the years as discussed under subsections 2.4.1, 2.4.2 and 2.4.3. The subsequent section 2.5 takes a step further to look at some of the key characteristics of macroscopic fundamental diagrams.

Table 2-3: Summary of Macroscopic Observation for Traffic Networks

	Period	Authors	Research Developments
Early Work with Empirical Observations	1952 1966	Wardrop Smeed	Developed macroscopic models for streets and highways later extended to network wide models.
	1966	Smeed	Introduced a mathematical function relating the area of a city with vehicle accumulation. The model revealed that the number of circulating vehicles in area is dependent on their average speed. However, the model couldn't be used to describe observations in overcrowded and large cities.
	1967	Thomson	Empirical data obtained from the streets of Central London was used to establish a decreasing linear function of average flow and speed.
	1968	Wardrop	Proposed an average speed and average flow functional relation by taking into account the signal spacing and street width. Findings revealed monotonic decreasing relationship.

	Period	Authors	Research Developments
	1969	Godfrey	Proposed a non-monotonic model capable of explaining the relation between vehicle concentration and the space-mean speed and between the hourly vehicle-kilometer and average speed. The model was however estimated with few data points and as a result needed further validation.
	1972	Zahavi	Proposed a relationship between the road density, traffic intensity and average speed. Findings revealed monotonic function that inversely relates speed and flow. However, the models could not describe heavily congested network performance.
Two-Fluid Model	1972 1979	Development of two-fluid model for the purpose of establishing the theoretical background and principles of macroscopic traffic relations in large cities.	
	1979	Herman Prigogine	Proposed a model able to characterize the quality of urban network traffic by linking the stop time to the trip time.
	1988	Herman Prigogine	The 1979 model was developed further and validated with empirical data from number of US cities.
	1988	Mahmassani Williams	Validated the fundamental relationship for urban-scale aggregated variables. The findings supported the two-fluid model as being a theoretical base for defining the state of an urban network.
	2007	Daganzo	Two-fluid model was found to have one key limitation. The model was found not valid under different demand patterns.
Recent Development	2007	Daganzo	Re-introduced modelling of urban gridlock using macroscopic variables under an “input-output” system description. The proposed model was argued to be valid only for homogeneously congested networks.

	Period	Authors	Research Developments
	2008	Geroliminis Daganzo	Empirically observed macroscopic relationships between aggregated variable, ‘production’ and ‘accumulation’. The existence of urban-scale MFD was proven to exhibit three distinct regimes; free flow regime, critical regime and congested regime.

2.5. Understanding MFD Characteristics

The previous section 2.3 provided a brief review of some of the earlier key empirical macroscopic observations of traffic networks including the review of the two-fluid model and the recent empirical observations of macroscopic fundamental diagrams for urban-scale networks. This section takes a step further to provide a brief review of some of the key characteristics of macroscopic fundamental diagrams. The hysteresis phenomenon covered in section 2.3 is further explained here under sub-section 2.5.2.

2.5.1. The MFD Estimation Under Ideal Conditions

Since the confirmation of MFD for urban-scale networks, there has been significant research focus on the features of the MFD and the implications thereof. Geroliminis and Daganzo (2008) proposed analytical formulae by introducing moving observers (Daganzo, 2005a, 2005b) and applying the variational theory. The proposed method was restricted to urban-scale networks with strong regularities in signal setting and topology.

Geroliminis and Daganzo (2008)'s proposed analytical model was later further developed by Leclercq and Geroliminis (2013) into a generic model able to relax the conditions for regulating topologies of networks. In addition, the study investigated various features of MFD under various principles of route choice (discussed in detail in section 2.4.3).

The three methods for the computation of macroscopic fundamental diagram, (a) the analytical formulae proposed by Geroliminis and Daganzo (2008), (b) the loop detector method and (c) the trajectory-based method influence by Edie (1963)'s method, were compared to each other by Courbon and Lecrercq (2011). The loop detector method records data at the location of loop, whereas the trajectory method takes into account different traffic states during the trip (by probe vehicles fitted with GPS) over the network to produce average flow and density.

.

Courbon and Lecrercq (2011) used Geroliminis and Daganzo (2008)'s analytical method as a reference during the comparison study and deduced that the probe vehicle data collection method was best suited to estimate the MFD. In addition, the study revealed that the combination of probe vehicle data and loop detector data produce the most accurate results. The trajectory-based estimation method was also further developed by Saberi *et al.*, (2014) from single corridor analysis to an urban scale network.

Saberi *et al.*, (2014)'s research study developed the conventional time-space diagram used in Edie (1963)'s method into a three-dimensional diagram by introducing another dimension in space. Saberi *et al.*, (2014) verified, through a series of simulation tests, that their proposed method is capable of estimating network level variables, consistent with the MFD concept.

2.5.2. Congestion Homogeneity's Impact on the MFD

Congestion homogeneity over a network is a key requirement for a well-defined MFD (Geroliminis and Daganzo, 2007, 2008). Buisson and Ladier (2009) were however first to demonstrate empirical demonstrations of the MFD in Toulouse (France) where congestion was not homogeneously distributed over the network (see Figure 2-15). Their research paper was devoted to relaxing Geroliminis and Daganzo (2007, 2008)'s explicit homogeneity conditions in gathering empirical data that led them to their well-defined MFD.

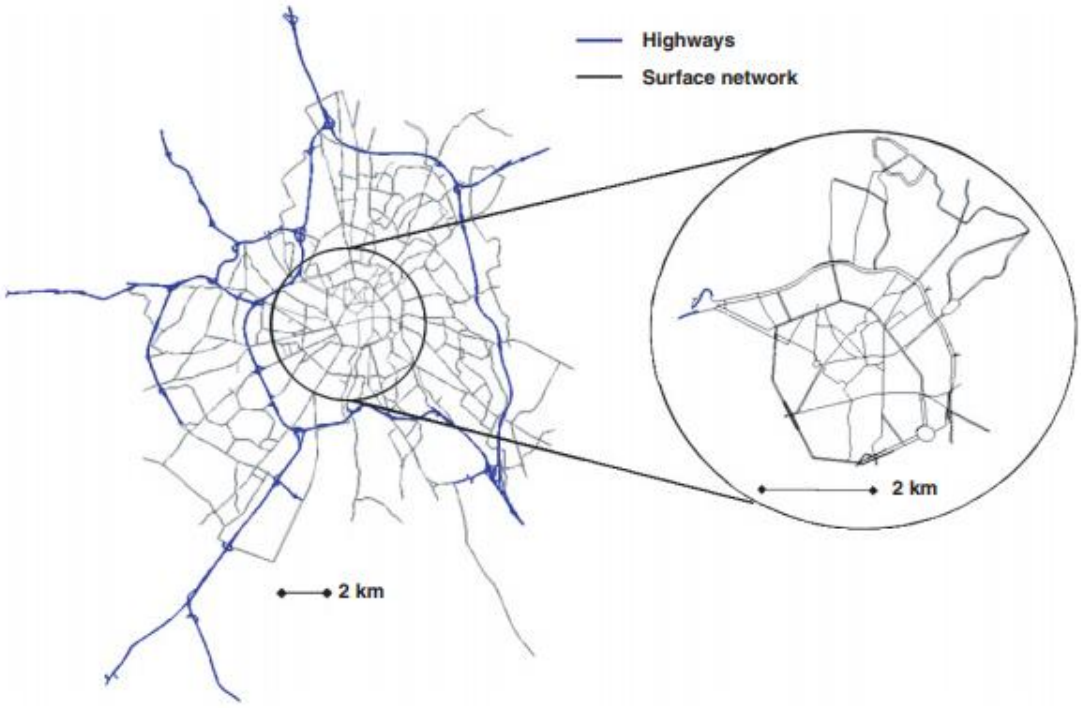


Figure 2-15: Toulouse (France) Highways and Surface Road Network

Source: (Buisson and Ladier, 2009)

Buisson and Ladier (2009) research study reached the following conclusions from their data analysis; First, there is a strong correlation between the road-type and the MFD shape. Their study showed that for Toulouse highway networks, the estimated MFD did not exhibit trapezoidal shape as shown in Figure 2-16 (b). As a result, dividing the network into zones should consider the type of road

Second, in an arterial road network, the distance between the detectors and downstream traffic signal strongly impact the observed features of the MFD as shown in Figure 2-17. It was therefore recommended that the MFD should be estimated with homogeneously located detectors. Lastly, a hysteresis loop was observed. This observation could have been due to spatially non-homogeneous evolution of congestion as shown in Figure 2-18.

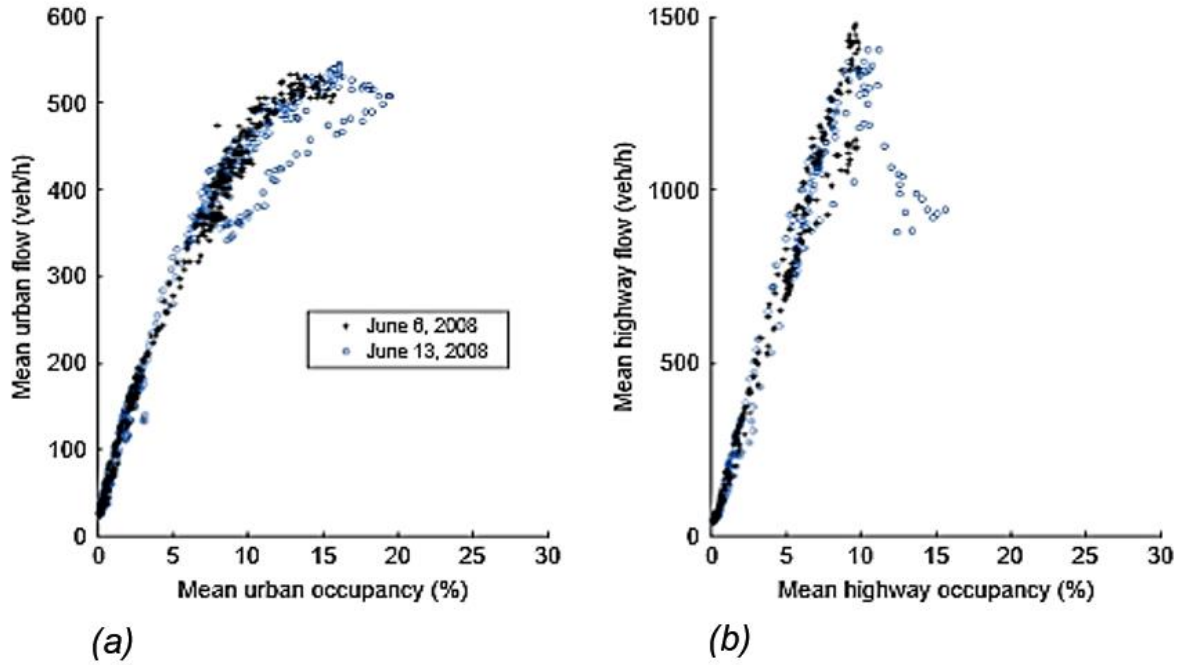


Figure 2-16: Toulouse (France)'s (a) surface street MFD for the 6th and 13th June and (b) highway MFD for the 6th and 13th June 2008

Source: (Buisson and Ladier, 2009)

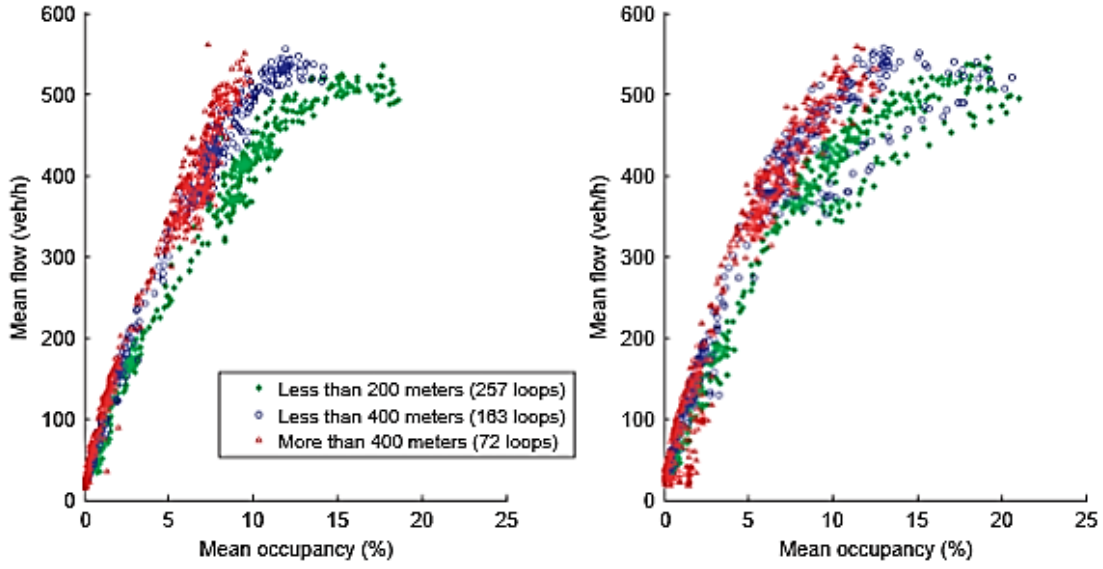


Figure 2-17: Toulouse surface street MFD illustration of the impact of the distance between fixed loop detectors and downstream traffic signal for the (a) 6th June 2008 and (b) 13th June 2008.

Source: (Buisson and Ladier, 2009)

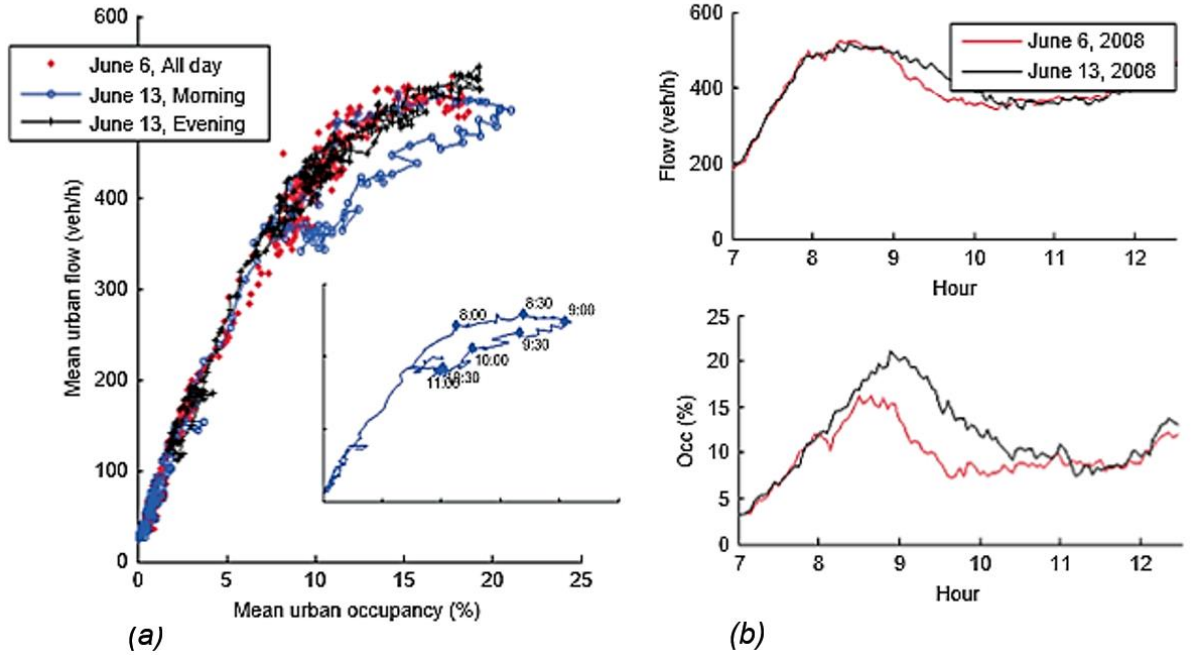


Figure 2-18: Toulouse global data for street loop detectors with less than 200m from the downstream traffic signal; (a) MFD illustrating the onset and offset evolution of congestion, (b) observed time-series of average flow and occupancy during the morning of the 6th and 13th June 2008

Source: (Buisson and Ladier, 2009)

Observation of the ovulation loop in Figure 2-18 shows that the onset of congestion observed on 13th June (morning) is consistent with all other MFDs, seen either on the evening of 13th June or 6th June. On the other hand, the fading away of congestion (after 9:00 in the morning) is observed at a lower mean flow than its appearance. These observations may have been as a result non-homogeneous spread of congestion.

Buisson and Ladier (2009)'s showed that the hysteresis loop that had lower average flow for the same level of density in the offset period and higher during the onset period, were regenerated in simulation studies of the Amsterdam road network by Ji *et al.,* (2010). They managed to verify the contour plot of time-space density and showed that the source of heterogeneity may be as a result of the observed uneven congestion during the onset and offset periods, resulting in the formation of the observed hysteresis loops.

Cassidy *et al.,* (2011) went further to study the theory and observation of freeway network's MFD. The research concluded that the relations at macroscopic level should be observed if the data is obtained during times when traffic lanes on all links over the network are in either uncongested or congested state. Their theory concerns freeways of any size including those that are heterogeneously congested.

Cassidy *et al.,* (2011) research further confirmed that the MFD for freeway networks could be estimated using loop detectors, given that each link in the network consist of a minimum of one loop detector and that the data is filtered to satisfy the requirement for single regime. In cases where the data set isn't properly filtered, chaotic scatters are observed below the MFD. Finally, the study revealed that policies and regulations to allow for even spread of congestion over the freeway network can be prove to useful in maximizing the trip rates.

Congestion spatial variability's effect on the features of MFD was investigated further. Mazloumian *et al.,* (2010) and Knoop *et al.,* (2011) confirmed that one of the contributing factors to the observed scatters results from the spatial distribution of link densities. They accomplished this by deploying a model for the simulation for the grid network with a boundary condition that there should be a constant number of vehicles in the network.

The simulation involved indefinite circulation within the network while choosing their route randomly. The simulation model showed the link densities variance over various locations is key aspect required for the determination of an overall network performance. Mazloumian *et al.,* (2010) and Knoop *et al.,* (2011) explained that nonhomogeneous spatial distribution of vehicle density increases the chances of spill over, resulting in decreased average network flow.

Knoop and Hoogendoorn (2013) later formulated what they termed Generalized Macroscopic Fundamental Diagrams (GMFD), a quantitative description of the network average flow as function of the spatial spread of density and accumulation. They used empirical data obtained from the Amsterdam freeway ring road to estimate GMFD.

The analysis of the Amsterdam freeway revealed that the GMFD was a continuous function which increases as accumulation increases, in the same way as the fundamental diagram. The GMFD explained the spread in the production than the MFD, especially close to this maximum observed production. This description of lean traffic state could therefore be utilized in setting a target for traffic monitoring and control (Knoop and Hoogendoorn, 2013).

The effect of the non-homogenous distribution of traffic on the features of the MFD was studied and confirmed by Geroliminis and Sun (2011a, 2011b). They carried out statistical tests on how link densities of signalized arterial network were spatially distributed. They conducted these tests in downtown Yokohama where a well-defined MFD was estimated and Minnesota freeway network where the MFD exhibited hysteresis loops.

The statistical tests results showed that whenever the link densities' spatial distribution for two different periods with the equal vehicle numbers in the area, produces similar histograms, the same network average flows should be seen for these different periods. The signalized network was found to satisfy this condition whereas the Minnesota freeway network did not satisfy the condition (Geroliminis and Sun, 2011a, 2011b). As a result, a significant scatter is observed under freeway MFDs.

Saberi and Mahmassani (2012, 2013) also carried out an empirical interpretation and characterization of hysteresis phenomenon for the freeway networks in Portland, Chicago and Irvine. Their research findings revealed that the observed high variations of densities during the recovery period results in low traffic flow, forming hysteresis loop (Saberi and Mahmassani, 2012).

In addition, Saberi and Mahmassani (2012) conducted an analysis of different highway networks in Chicago, Irvine and Portland and confirmed that the drop in freeway capacity was caused by, (a) the network traffic's instability during an evening peak following an incomplete recovery from the morning peak and (b) the inability of the freeway network to sustain its maximum flow for a period of time

Based on Saberi and Mahmassani (2012, 2013)'s research outcomes, Ji and Geroliminis (2012) went further to develop an algorithm for the partitioning of non-homogenously freeway congested networks in San Francisco. The objectives of the network partitioning were to (a) obtain minimal link densities' variance within a cluster (b) ensure that each cluster is spatially compact. The research findings showed that both objectives of spatial compactness and minimal variance can be realised with the proposed partitioning algorithm.

2.5.3. Turning Behaviour of Drivers' Impact on the MFD

Mazloumian, et al, (2010) observed that vehicle drivers have a tendency of distributing unevenly when there is a constant number of circulating vehicles in an area while having the choice to randomly choose their own routes. This observation is one of the causes of non-homogeneous conditions resulting in scatters below the MFD

The two-bin model proved that random turns at intersections contributes to uneven congestion of the system. This phenomenon results in the bifurcation phenomenon observed in the MFD at maximum flows, causing the system to gridlock. Gayah and Daganzo (2011) confirmed this phenomenon with a grid network and ring road network.

Daganzo *et al.,* (2011)'s paper confirmed that allowing drivers to adaptively choose their routes reduces the tendency towards unevenness and increases bifurcation point density, reduces the probability of the system reaching gridlock. Gayah and Daganzo (2011) used Daganzo *et al.,* (2011)'s two-bin model to further explore the hysteresis loops.

To account for network traffic dynamics during peak periods, Gayah and Daganzo (2011) generalized the model by introducing exogenous inflow and outflow. Their research study confirmed that the system is not sustainable during the recovery phase because of the slow recovery of congestion in highly congested areas.

The clockwise loops observed in the MFD is caused by this unstable phenomenon; the lower network average flow observed during the recovery phase instead of during the loading phase. In addition, the study showed that if adequate number of vehicles could adaptively choose their routes in reaction to congestion, then the instability during the recovery phase could be mitigated (Gayah and Daganzo, 2011).

Mahmassani *et al.,* (2013) and Tsubota *et al.*, (2013) carried out an investigation of the dynamic information provision's impact on the features of the MFD using simulation models. Mahmassani *et al.,* (2013) confirmed that the size of the gridlock reduces as more drivers adaptively choose their routes in response to congestion.

Tsubota *et al.*, (2013) carried out a simulation of a grid network using various adaptive drivers' ratios and showed that the dynamic information does have an impact on the distribution of drivers over the network. They confirmed that traffic in the network spreads homogeneously when there is sufficient number of adaptive drivers, resulting in high network average flow. These observations were found to be consistent with the theoretical observations by Gayah and Daganzo (2011).

Leclercq and Geroliminis (2013) went further to investigate the impact of different route choices principles on the shape of the MFD. They derived the MFD using Wardrop (1952)'s user equilibrium principle, the system optimum condition and the logit model. Their findings were consistent with Mazloumian *et al.,* (2010)'s observation; the link densities' variance is the key variable in the determination of the MFD shape.

2.6. Overview of MFD Estimation Processes in Major Cities

The previous section 2.5 focused on the key characteristics of macroscopic fundamental diagrams in general. This section presents a brief review of some of the key processes followed during the estimation or derivation of MFDs for the urban-scale networks of Yokohama (Japan), Sendai (China), Zurich (Switzerland) and Toulouse (France).

2.6.1. Data Provision for MFD

The estimation of the MFD was initially based on the key assumption that congestion spreads homogeneously across the network and that it is independent of demand patterns as long as the average travelled distance remains the same (Geroliminis and Daganzo, 2007). However, different research findings challenged these assumptions (Buisson and Ladier, 2009). Urban networks are not always homogeneously congested. As a result, efforts were made to partition networks according to the homogeneity of congestion (Ji and Geroliminis, 2012).

In addition, recent research studies show that the MFD is variant to changes in the origin-destination (OD) matrix (Lecrercq *et al.*, 2015). Given such limitations, the question arises; how can a well-defined and reproducible MFD be estimated from the available empirical data. There are two viable sources of empirical data for the estimation of MFD i.e. (i) fixed loop detector data and (ii) probe vehicle data.

They typically record the traffic variables; (i) flows and occupancy. Loop detectors are used to count vehicles, detect congestion and control traffic signals. They have been used as input data for the estimation of MFD empirically and through simulations.

Probe vehicle data is collected from probe vehicles transmitting the data through a trajectory measurement device such as GPS or cellphones. Probe vehicle data requires a matching of the GPS trajectories to the road network. Floating vehicle data has also been used to estimate the MFD empirically and through simulation (Tsubota *et al.*, 2013). Important issues to consider with the use of floating vehicle data are the probe penetrations rate and its spatial distribution

Different methods used to estimate MFDs were evaluated and it was found that a combination of fixed loop detector data and GPS probe vehicles data provided the most valid and reliable results in simulation studies (Leclercq *et. al.,* 2014). Table 2-4 below shows the empirical data sources for MFD estimations in the cities of Yokohama (Japan), Toulouse (France), Sendai (Japan) and Zurich (Switzerland).

Table 2-4: Empirical studies on urban MFD estimation

Country, City	**Data**	**Reference**
Japan, Yokohama	Loop Detector Data Floating Vehicle Data	(Geroliminis & Daganzo, Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings, 2009)
France, Toulouse	Fixed Loop Detectors	(Buisson & Ladier & Ladier,2009)
Japan, Sendai	Fixed Loop Detectors	(Wada & Hara, 2015)
Zurich, Switzerland	Loop Detector Data Floating Vehicle Data	**Invalid source specified.**

2.6.2. Definition of MFD

This sub-section provides a review of the analytical formulae used in the estimation of the MFDs in Yokohama (Japan), Sendai (China), Zurich (Switzerland) and Toulouse (France).

(a) Yokohama (Japan) MFD

Daganzo and Geroliminis (2008) conducted an empirical study in Yokohama (Japan) that showed that MFD that links network average flow, speed and density exists for a large urban-scale network. Data obtained from fixed loop detectors and floating vehicles for one month (December 2001) were used to estimate a well-defined MFD using the following formulae;

$$q^w = \frac{\Sigma_i q_i l_i}{\Sigma_i l_i} \qquad (2\text{-}18)$$

$$o^w = \frac{\Sigma_i o_i l_i}{\Sigma_i l_i} \qquad (2\text{-}19)$$

Where q^w is the weighted average flow (in veh/hr/lane), o^w is the weighted average occupancy (in %), i and l_i are road lane segment between intersections and its length (in km), respectively. And q_i and o_i , the flow (in veh/hr/lane) and occupancy (in veh/km/lane) measured by the corresponding loop detector within an aggregation time.

The results showed that when a high scatter of flow-density from individual fixed loop detectors (see Figure 2-19a) were aggregated over the network, the scatter faded and individual data points grouped and followed a smooth declining curve (see Figure 2-19b). This observation suggested, but not proved, that an MFD exist for the complete and whole network because the loops only captured localized conditions which may not represent the entire network.

As a result, the quality of data analysis process was further enhanced by data from floating vehicles, which managed to cover the whole network. Using the relevant parts of the loop detector and probe vehicle data sets, the network average densities and speeds at various times of the day were estimated.

The estimates were also observed to be lying close to low-scatter declining curve. The data analysis process also confirmed a relation between the average flows on the entire network, which are estimate given the existence of a well-defined MFD and the trip completion rates (Daganzo and Geroliminis, 2008)..

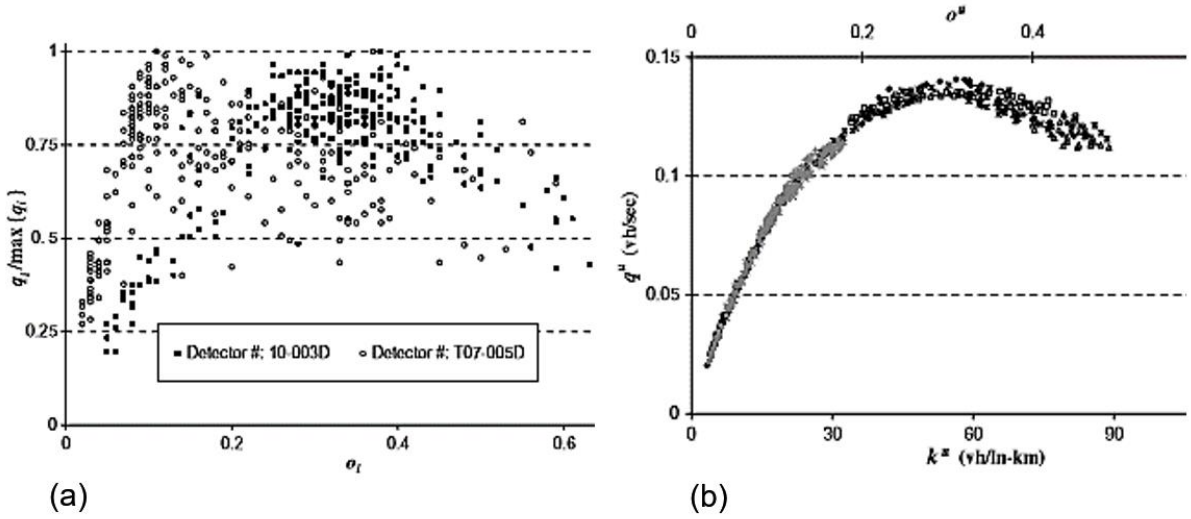


Figure 2-19: Yokohama: (a) Weekday flow-occupancy plot for two individual loop detectors and (b) the observed MFD for the entire network.

Source: (Daganzo and Geroliminis, 2008)

(b) Zurich (Switzerland) MFD

Ambühl *et al.*, (2016) estimated and analyzed MFD for the city of Zurich in Switzerland (see Figure 2-20) using the both loop detector and probe vehicle data sets. Their study showed that each data source produced a well-defined and reproducible MFD. However, they have observed differences in the shape and feature of MFDs for the two data sets, this difference could be attributed to limitations of the data source.

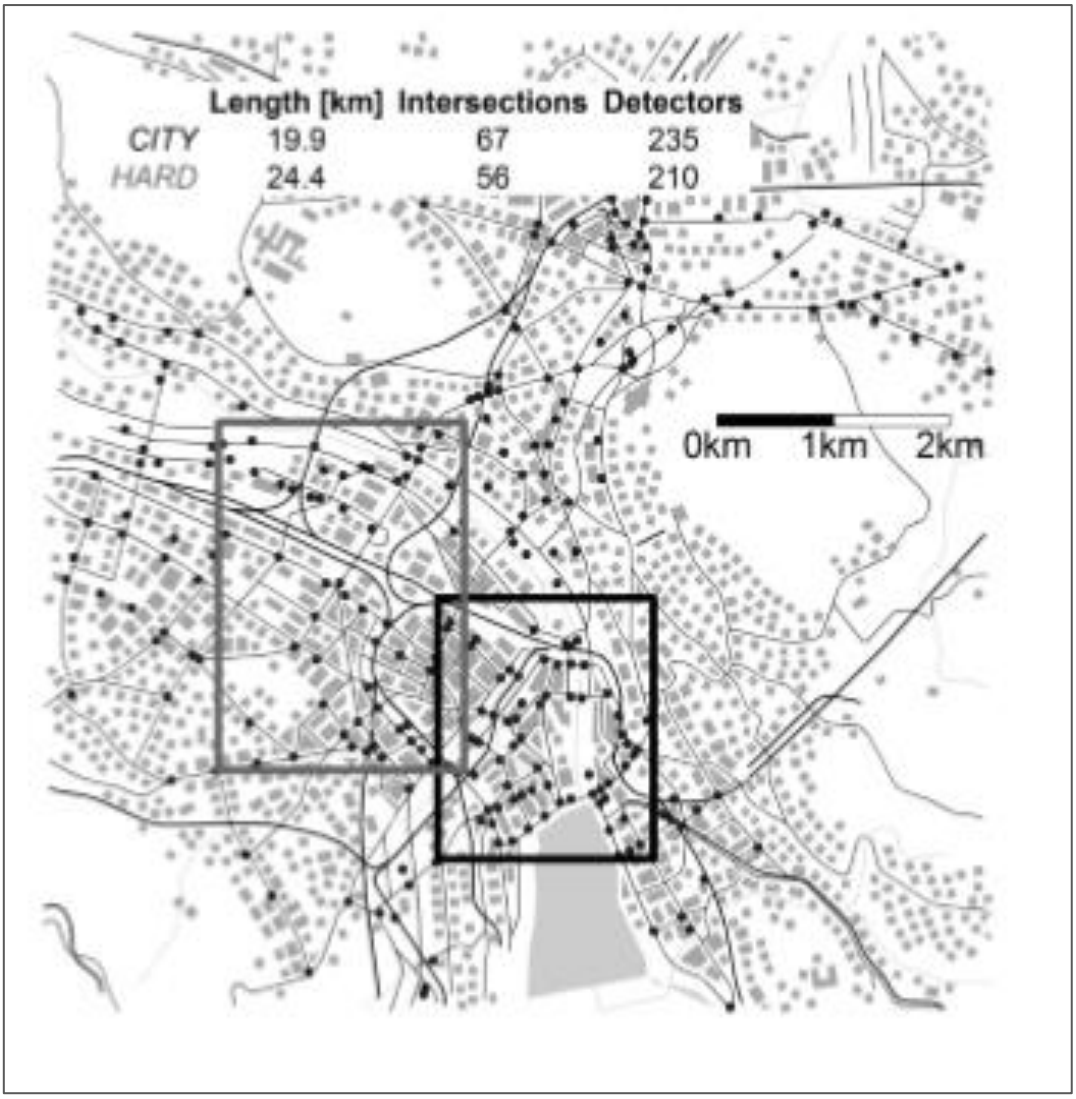


Figure 2-20: Case Study Network, Zurich.

Source: (Ambühl et al., 2016)

The study identified a link selection bias and placement bias for the loop detector data, which led to an overestimation of density values or occupancy, respectively. These biases where mitigated by a developed methodology accounting for the relative position of a loop detector on links and their frequency at that position.

In addition, Ambühl *et al.*, (2016) combined loop detector data flows and probe vehicles data speeds to estimate the MFD, which partly eliminated key drawbacks of both data sources. Ambühl *et al.*, (2016)'s loop detector data based MFD was estimated in accordance with Daganzo and Geroliminis (2008).

Accordingly, they calculated the network flow and network occupancy, including all loop detectors, as follows;

$$q_{LDD} = \frac{\Sigma_i q_i l_i}{\Sigma_i l_i} \qquad (2\text{-}20)$$

$$o_{LDD} = \frac{\Sigma_i o_i l_i}{\Sigma_i l_i} \qquad (2\text{-}21)$$

Where l_i represents length of the link i, q_{LDD} (in veh/hour/lane) represents loop detector data network flow and o_{LDD} (in %) represents loop detector data network occupancy, while q_i and o_i represents individual loop's average flow and occupancy during the 5min aggregation period, respectively (Buisson and Ladier, 2009).

Since the fixed loop detectors were used for traffic control and monitoring, a subsample of detectors was located on turning pockets. Figure 2-11(a) illustrates how the effect of excluding turns increased the maximum flow by 15%.

Buisson and Ladier (2009) research study showed that the placement of loop detectors does have a significant impact on the shape of the MFD. Ambühl *et al.*, (2016) confirmed Buisson and Ladier (2009) findings by restricting the location of loop detectors to more than 20m upstream of a traffic signal in **Figure 2-21** (b).

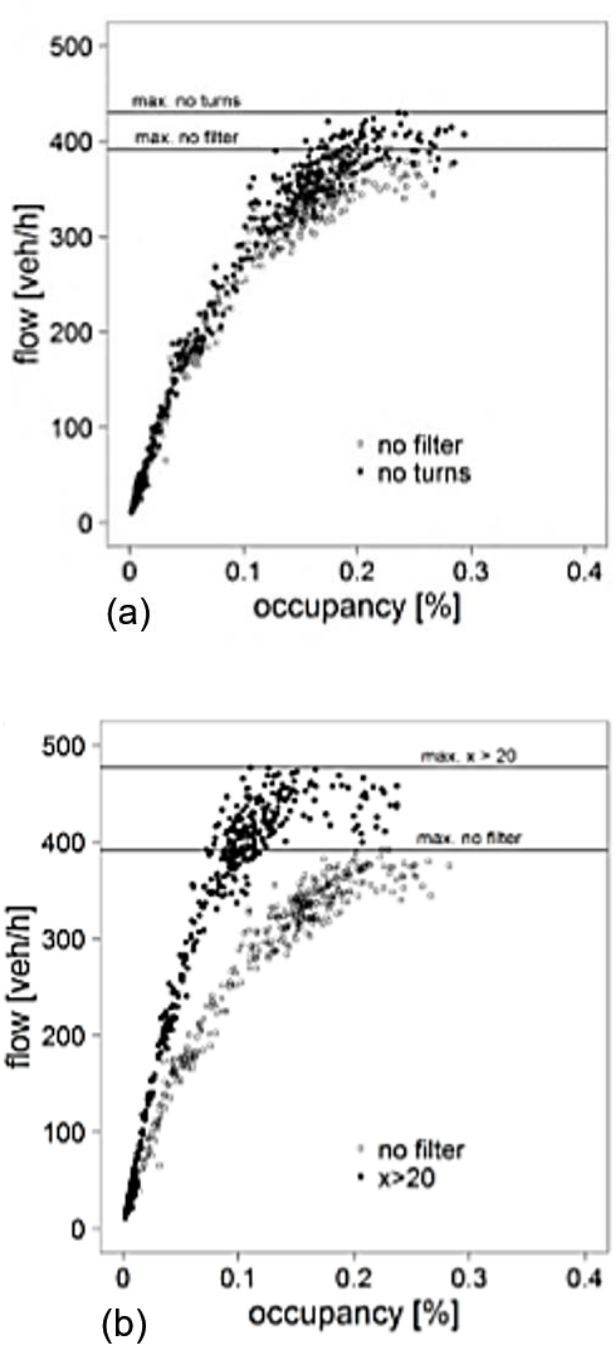


Figure 2-21: (a) Loop detector data filtering turns (b) Loop detector data filtering loop detector position.

Source: (Ambühl et al., 2016)

Ambühl *et al.*, (2016) also estimated the MFD using the probe vehicle data. Probe vehicle data provided 2-year daily averages of hits and speeds during 5min intervals. They analyzed the macroscopic relations using the following equations;

$$\hat{q}_{PVD} = \frac{\sum \max\left(H_i\, l_i, v_i\, T\right)}{T \sum l_i} \qquad (2\text{-}22)$$

$$\hat{v}_{PVD} = \frac{\sum \bar{v}_i l_i}{\sum l_i} \qquad (2\text{-}23)$$

Where $\hat{q}_{PVD}$ and $\hat{v}_{PVD}$ are network flow and network speed, respectively. H_i is the average number of probe vehicles during observed time T on link i of length $l_i \bar{v}_i$ is the average speed of probe vehicles on link i of length l_i over an aggregation period of 5min. Figure 2-12 below illustrates the estimated MFD based on floating car data (FCD).

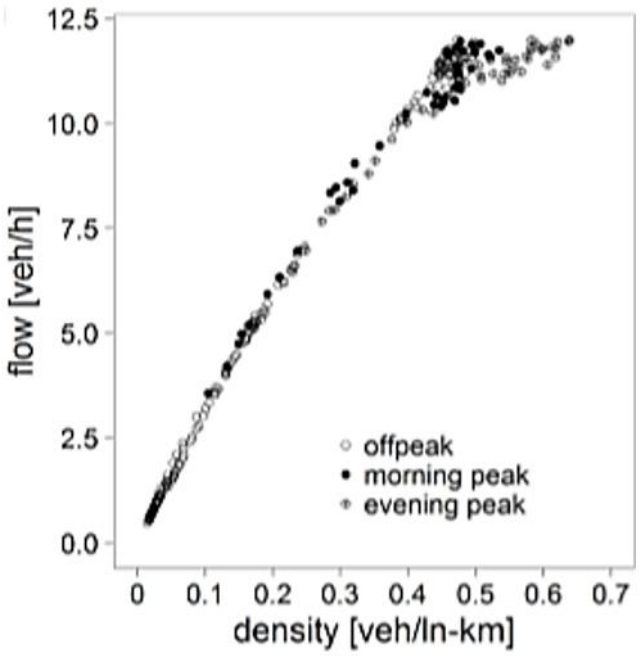


Figure 2-22: Probe vehicle data MFD, Zurich.

Source: (Ambühl et al., 2016)

The density was calculated by dividing the MFD flow by the MFD speed. The flow and density values were found to be low due to the sample size. Nevertheless, a low-scatter MFD was apparent. It was interesting to observe that the MFD did not show strong indication of congestion.

(c) Sendai (Japan) MFD

Wang *et al.*, (2015) estimated and explored the features of the MFD for Sendai road network. They used observed data of 5min periods for one year (01 May 2012 – 30 April 2013) to draw the MFDs for Sendai road network. Each plotted point is the pair (average density $\bar{k}_t$, average flow $\bar{q}_t$) observed during each period t for each day. The average density and average flow are defined by Equations (2-24) and (2-25), respectively.

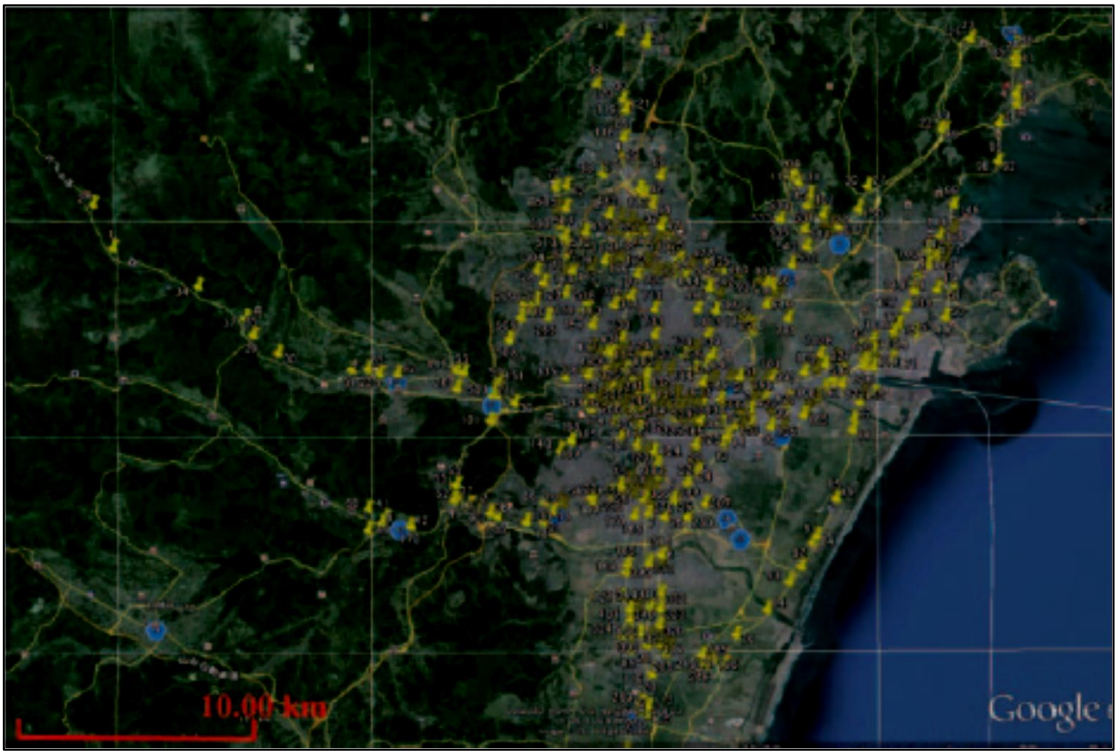


Figure 2-23: Locations of loop detectors within Sendai road network.

Source: (Wang et al., 2015)

$$\bar{k}_t = \frac{\sum_{i=1}^{I_0} k_t^i}{|I_0|} \qquad (2\text{-}24)$$

$$\bar{q}_t = \frac{\sum_{i=1}^{I_0} q_t^i}{|I_0|} \qquad (2\text{-}25)$$

Where |I0| in Equations (2-24) and (2-25) is the number of loop detectors in set I_0 (see Figure 2-23) They divided the 24 hours in each day into four different time periods i.e. 0:00 – 6:00, 6:00 – 13:00, 13:00 – 20:00 and 20:00 – 24:00. Each time period is shown with a different colour on the MFD (see Figure 2-24).

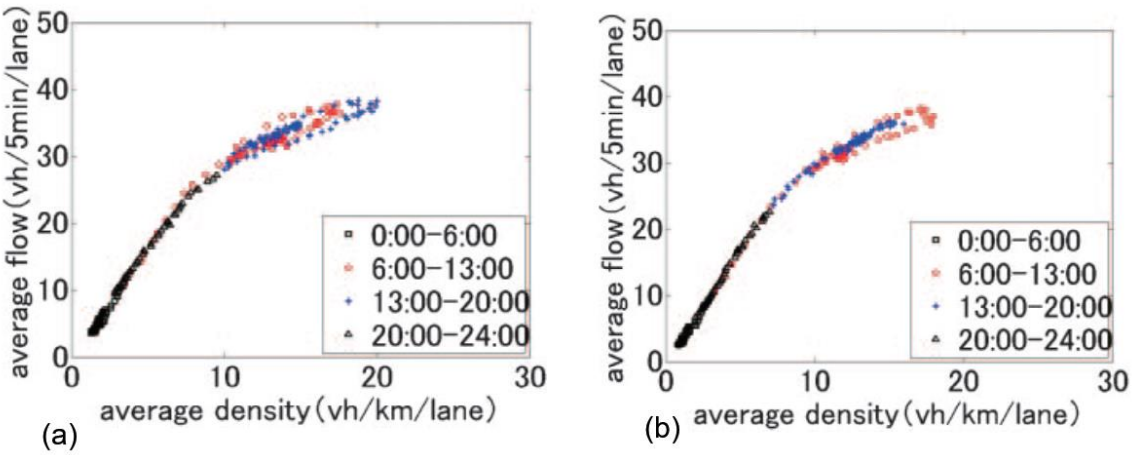


Figure 2-24: Sendai typical MFD for the (a) 14th September (Friday) and (b) 19th November (Monday) 2012.

Source: (Wang et al., 2015)

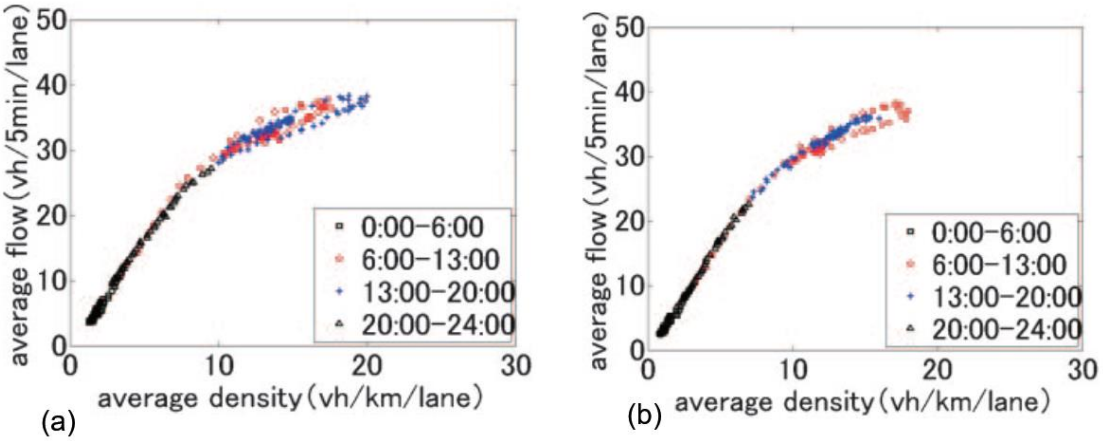


Figure 2-25: Sendai typical MFD for the (a) 14th September (Friday) and (b) 19th November (Monday) 2012.

Source: (Wang et al., 2015)

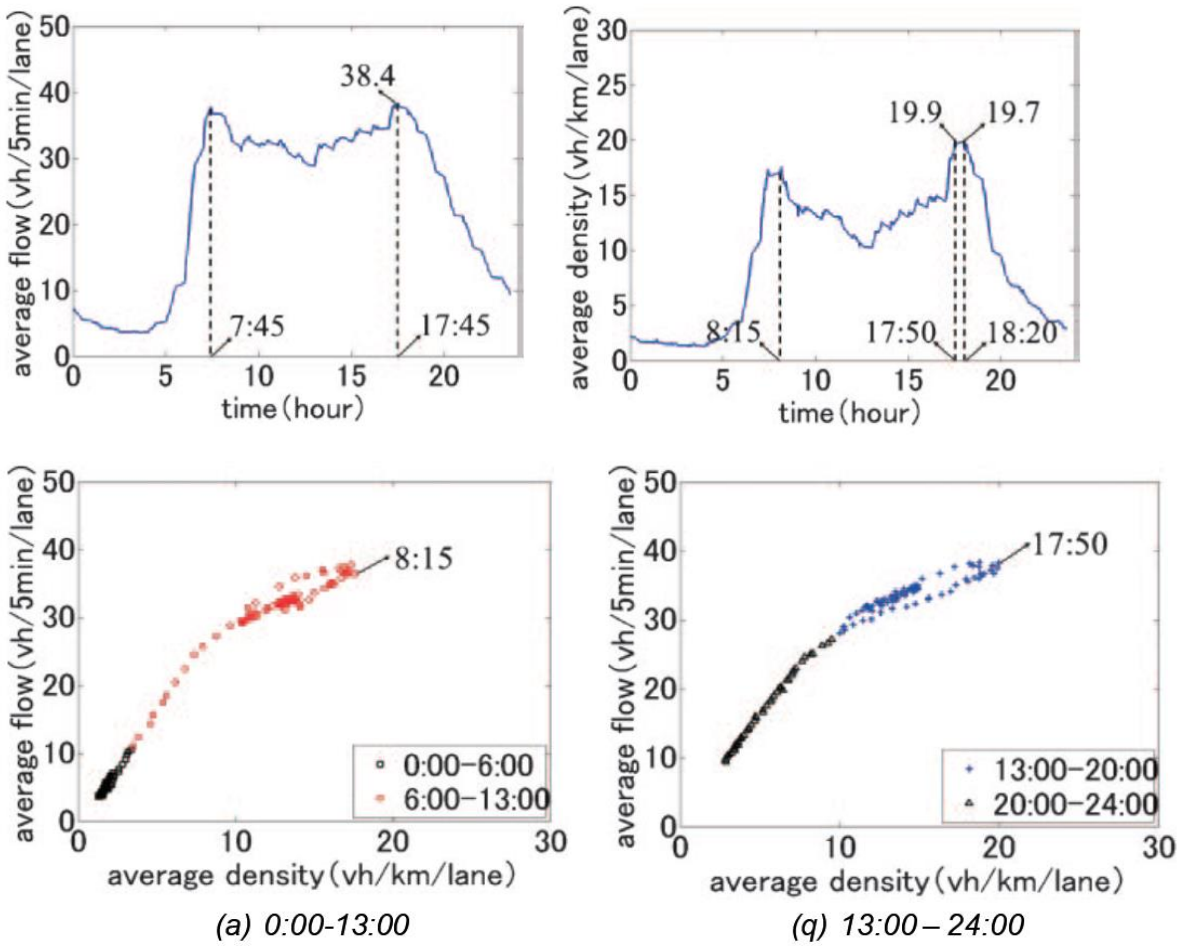
Numerous MFD observations revealed that the daily MFD throughout the year could be classified into several types depending on its shape. Among them, one was “typical MFD” that appeared on weekdays during good weather conditions (herein referred to as “standard condition MFD”), as shown in Figure 2-6.

All the standard condition MFDs exhibited the following two common features; (i) a hysteresis loop always exists and (ii) the MFD shape is invariant throughout the entire year. It was found that the significant feature of the standard condition MFD was the formation of a loop.

To express this feature more clearly, Wang *et al.*, (2015) divided the 24 hours into two parts i.e. the morning (0:00-13:00) and the evening (13:00-24:00) to draw the MFD and found that; (i) in the morning, a single loop always forms (see Figure 2-26 (a) &

Figure 2-28 (a) and (ii) in the evening, a loop may form (see Figure 2-26 (b) &

Figure 2-28 (b)).



(a) 0:00-13:00 *(q) 13:00 – 24:00*

Figure 2-26: MFD in the morning and evening of 14th September 2012

Source: (Wang et al., 2015)

Figure 2-27: Average flow and average density time series for the 14th September 2012.

Figure 2-28: MFD observed during the morning and evening of the 19th November 2012.

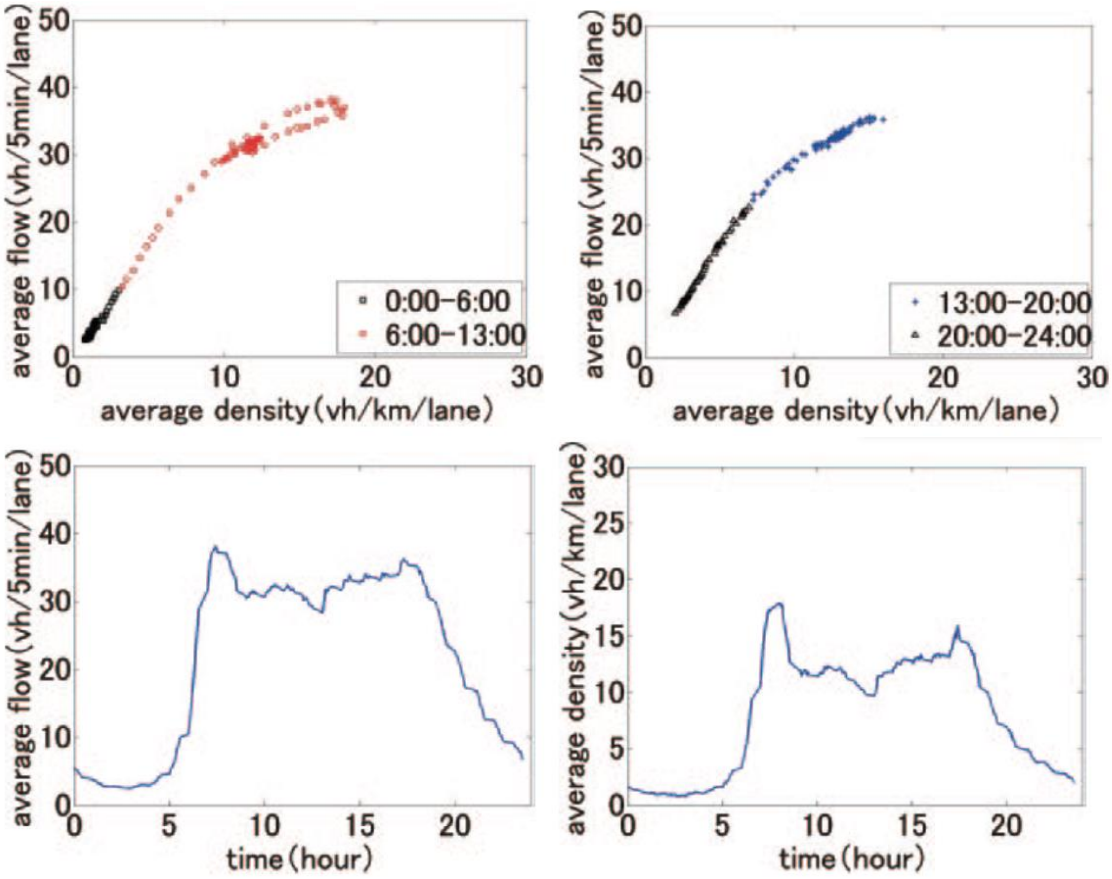


Figure 2-29: Average flow and density time series of 19th November 2012.

Source: (Wang et al., 2015)

The MFDs were explained in conjunction with the patterns of the network average flow and network average density time series. First, Figure 2-26 (a), Figure 2-26 (b), and Figure 2-27 (a), Figure 2-27 (b) showed that the coordinates of plots of the loop on the

MFD correspond to the average density and average flow in peak hours. It was found that the loop indeed form in the MFD during morning peak from 7:30-8:30 and in the evening from 17:30-18:30.

(d) Toulouse (France) MFD

Buisson and Ladier (2009) estimated and explored the impact of non-homogeneity on the MFD for the Toulouse (France) road network. The loop detector data set includes measurements on highways, urban centre streets and residential area streets. Heterogeneity was examined is several aspects; (a) difference between highway and surface network, (b) the distance between loop detector and downstream traffic signal's impact on the MFD and (c) the differences between ring roads and penetrating roads in the highway network.

They estimated the MFD using the following unweighted mean values (Equations (2-26) and (2-27)) suggested by Geroliminis and Daganzo (2008). Where q_i and t_i are the flow and occupancy rate for an aggregation period of 3 minutes for the i loop detector, n_i is the number of lanes, T is the occupancy rates (in %), Q is the average flow (in veh/hr/lane) and N is the total number of loop detectors i across the network.

$$Q = \frac{1}{N}\sum_{i=1}^{N}\frac{q_i}{n_i} \qquad (2\text{-}26)$$

$$T = \frac{1}{N}\sum_{i=1}^{N} t_i \qquad (2\text{-}27)$$

The authors drew the global mean flow times series of a total set of 528 loop detectors and observed the following features as shown in Figure 2-30 (a); (a) the morning and evening peaks are well pronounced, especially the evening peak and (b) At around 1 PM, two sudden peaks were observed.

.

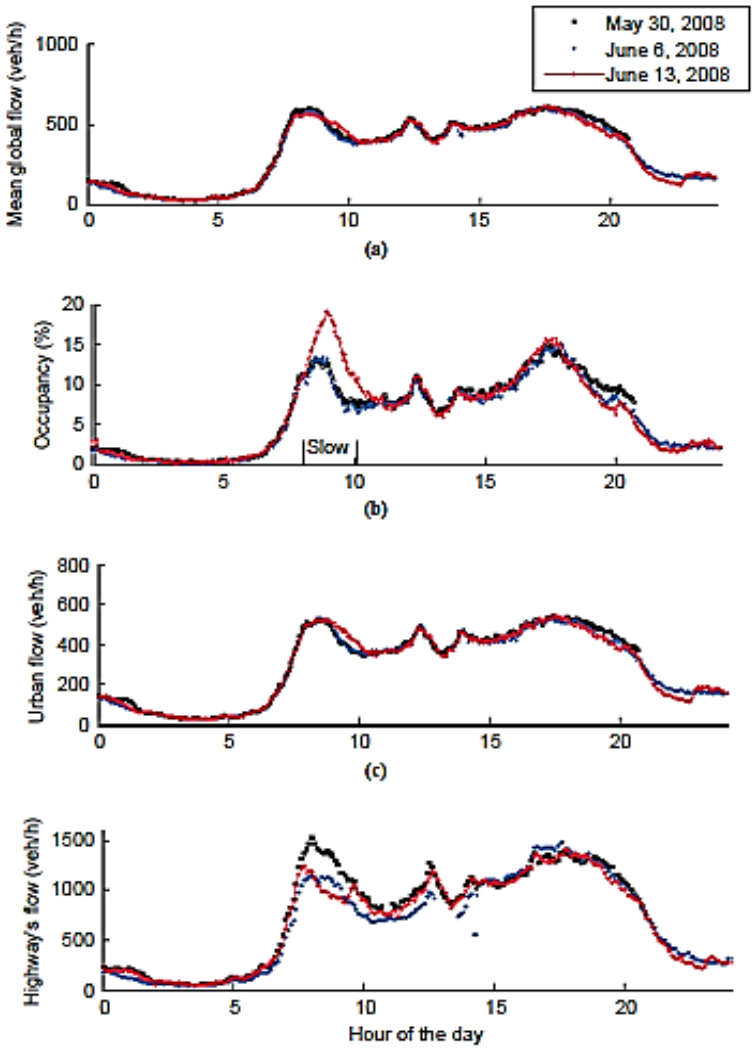


Figure 2-29: Toulouse's evolution of global variables for the 30th May (Friday), 6th June 9 (Friday), and 13th June (Friday) 2008: (a) global average flow, (b) global average occupancy, (c) surface network observe average flows, and (d) highway network observed average flows.

Source: Buisson and Ladier (2009)

The global MFD for the Toulouse street and highway networks is presented in Figure 2-30. The global MFD for Toulouse was compared with the equivalent Daganzo and Geroliminis (2008)'s MFD as shown in Figure 2-31. It was observed that the Toulouse MFD's maximum flow was in the same order of magnitude.

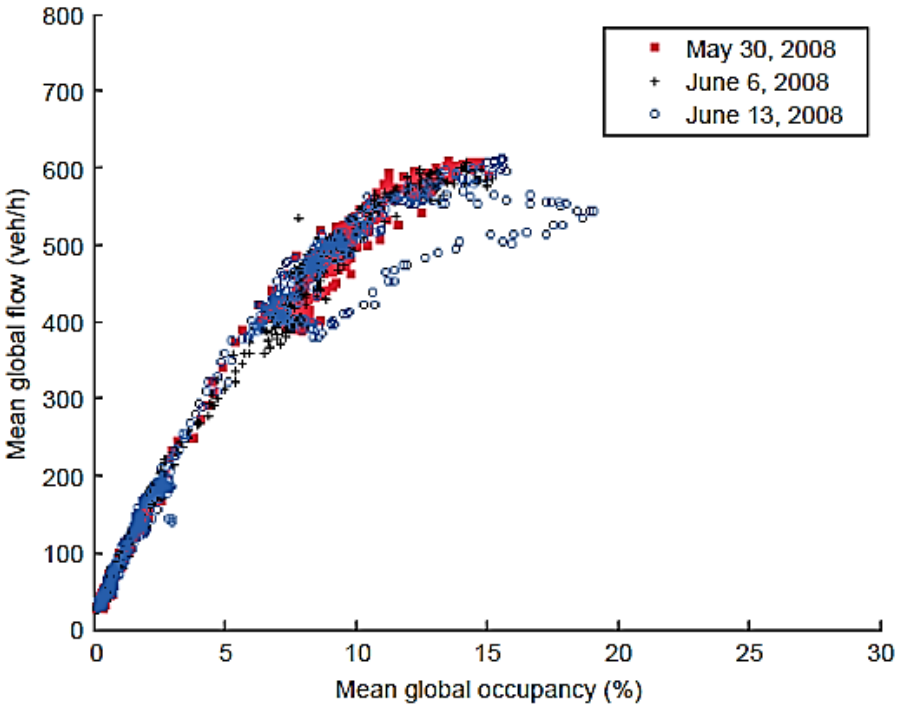


Figure 2-30: Toulouse, MFD for the 30th May, 6th June and 13th June 2009.

Source: Buisson and Ladier (2009)

The Toulouse MFD's critical occupancy was observed to be significantly lower than the Yokohama MFD's critical occupancy. Buisson and Ladier (2009)'s argued that this observation was "probably the result of the specific configuration of loops detectors".

The main difference between the two MFDs was the scatter. As discussed earlier, various types of non-homogeneity may have caused the scatter, including; (a) difference road network types, highways versus surface street, (b) the appearance and disappearance of congestion and (c) difference in data measurement location.

The set of detectors was further partitioned into sub-networks; one corresponding to the ring road (29 loop detectors) and the other to the penetrating roads (16 loops). Figure 2-32 below shows the analysis results

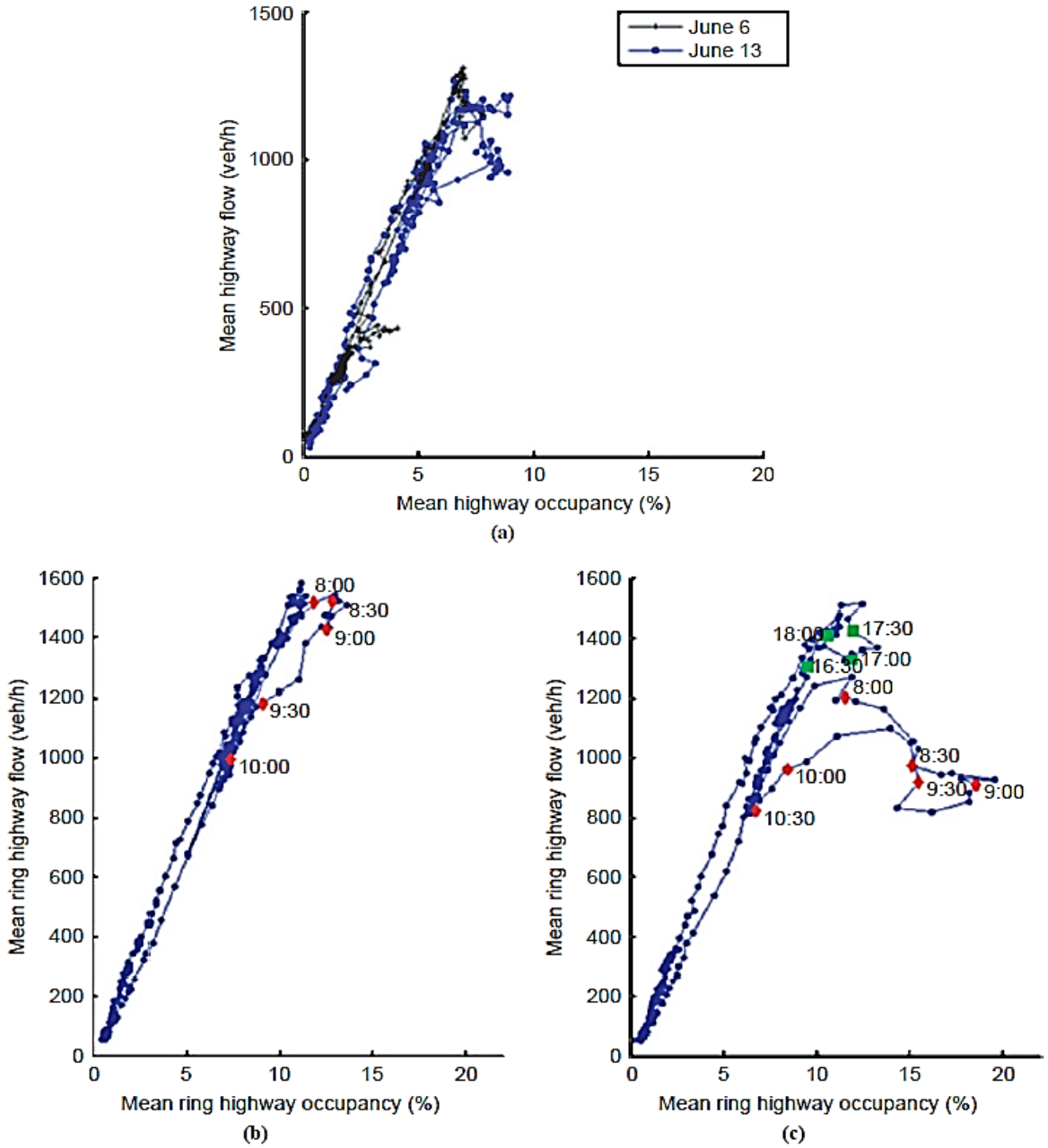


Figure 2-31: Toulouse's Macroscopic Fundamental Diagram for the (a) penetrating roads network on the 6th and 13th 2008, (b) ring road network for the 6th June 2008 and (c) ring road network for the 13th 2008

Source: Buisson and Ladier (2009)

The two subnetworks exhibited significantly different maximum average flows; 1600 veh/h/lane for the ring highways (see Figure 2-32 (b) and Figure 2-32 (c)) versus 1300 veh/h/lane for penetrating highways, shown in Figure 2-32 (a).

The results confirmed that congestion observed on the penetrating roads subnetwork was insufficiently spread over the highways forming part of the sub-network to allow observing strong congestion (see Figure 2-32 (a)); the minimal congestion flow was observed to be 75% of the maximal mean flow. The ring road subnetwork (see Figure 2-32 (b) and Figure 2-32 (c)) minimal congestion flow was observed to be half of the maximal mean flow.

Further analysis of the results for the ring road network, showed that congestion was mainly observed during the morning peak of the 13th June 2008 with a hysteresis-like shape. The ring road was heterogeneously congested at any moment during the morning of the 13th June 2008. The hysteresis-like shape could be explained by this heterogeneity in congestion.

In summary, Buisson and Ladier (2009)'s paper was devoted to relaxing Geroliminis and Daganzo (2008)'s explicit homogeneity conditions in selecting the data which resulted in the observed MFD in downtown Yokohama. They concluded that;

a) The network road type has a string impact of the observed features of the MFD

b) Given a data set that is homogeneous in terms of road type, separating the ring road from penetrating road is a critical in obtaining less scatter in the MFD.

c) The hysteresis-like shape observed in freeway networks is due to nonhomogeneous spatial evolution of congestion.

2.7. Applications of MFD for Traffic Monitoring and Control

The previous section 2.6 provided a brief review of the data provision for the estimation of MFDs, the relevant derivation formulae and record of some of the key findings observed in the cities of Yokohama, Zurich, Sendai and Toulouse. This concluding section focuses on the some of the key applications of MFD for urban-scale traffic monitoring and control.

Improved Urban Mobility and Congestion Relief

The successful implementation of MFD for urban-network traffic monitoring and control requires an accurate analysis and understanding of properties of the MFD for that urban-network. To date, there are reported studies that aimed to use the MFD for traffic control and monitoring. Daganzo (2007) first proposed the application of MFD to describe an adaptive control approach to relieve congestion and aid urban mobility. The fundamental idea consists in the control and monitoring of aggregate vehicular accumulations at an urban-network level.

Ramp Metering Strategies

Yoshii *et al.,* (2010), Geroliminis and Sun (2011a) developed and introduced ramp metering strategies with the purpose of optimizing the total number of vehicles circulating within an urban network. The MFD, even though it was originally purposed for urban-scale networks, was found suitable for the identification of control targets used in the coordination of metering control of different number of on-ramps simultaneously. In addition, Haddad *et al.,* (2013) presented a model which takes into account the freeway and the surface network.

Perimeter Control (or Gating) Strategies

In urban network controls, gating strategies are used to control traffic inflow through the modification of traffic signals at the boundaries in order to optimize urban network and possibly prevent it from reaching congested state. Haddad and Geroliminis (2012) proved that MFDs may be used to derive state-feedback control strategies capable of stabilizing and maximizing the output of the system.

They formulated control mechanisms for the two-reservoir system and demonstrated how the mechanism will function. Geroliminis *et al.,* (2013) used macroscopic fundamental diagrams to formulate predictive control measures for the two-reservoir system. This was followed by Aboudolas and Geroliminis (2013)'s introduction of a boundary flow and perimeter control in multi-reservoir non-homogenous networks.

Zone-Based Routing, Cordon Pricing and Evaluation of Network Performance

The MFDs have also been used for "zone-based routing systems" by Knoop, et al., (2012), "cordon pricing" by Geroliminis and Levinson (2009) and Zheng *et al.*, (2012) and "the evaluation of network performance with mixed traffic" by Xie *et al.,* (2013). Xie *et al.,* (2013)'s research work is particularly notable because it extended the traditional single-mode MFD to a multimodal framework.

2.8. Summary and Conclusions

This chapter commenced with a brief review of the conventional measures of traffic congestion focusing on highway capacity measures (levels of service), travel time measures and delay measures. These measures, though they serve other useful purposes, were limited to highways and arterials and could not be utilized to measure congestion for urban-scale networks.

This gave rise to macroscopic fundamental diagrams (MFDs) being recommended for urban-scale traffic monitoring and control. However, prior to the detail review of MFDs, the chapter first provides a review of the traffic flow theory including the review of Greenshields (1935)'s macroscopic stream model and shock wave theory.

The chapter provided a comprehensive background to the hysteresis phenomenon. It was established that the hysteresis is observed at both microscopic and macroscopic levels of traffic flow. It was established that the hysteresis phenomenon is observed whenever there is a disturbance in the traffic flow. Overwhelming evidence shows that urban-scale freeway networks' MFDs exhibit hysteresis patterns.

The chapter went on to provide a detailed review of macroscopic modelling for networks from earlier research work with empirical network observations, the two-fluid model and some of the recent development on the description of input-output system.

A well-defined MFD characterized the state of traffic as a function of network average density and could therefore be used for network monitoring and control. The successful application of MFD in traffic control and monitoring relies on a well-defined MFD.

Recently, Buisson and Ladier, (2009) proved that non-homogeneity strongly impact the shape of the MFD giving rise to new insights on the properties of MFDs under nonhomogeneous conditions such as network made up of different road types and heterogeneous distribution of congestion (Ji *et al.*, 2010; Daganzo, et al., 2011; Geroliminis and Sun, 2011b).

The resulting MFDs (under heterogenous conditions) were highly scattered and exhibit a hysteresis patterns and/or bifurcation points at low densities. The observation of scattered data points with hysteresis loops implied that the networks area contained different traffic states. As a result, the scattered MFD could result in inaccurate characterization of the state of the network which may potentially results in misinformed traffic monitoring and control.

Further research was conducted to gain understanding of conditions under which scatters form in the freeway MFDs. In this regard, Mazloumian, et al., (2010) and Geroliminis and Sun, (2011a) made significant findings regarding the impact of network homogeneity on the shape of MFD. They observed that the variation of link densities has significant impact on the network average flows.

The chapter also presented a review of the data provision and formulae for the estimation of MFDs in the cities of Yokohama, Zurich, Sendai and Toulouse, including a review of the some of the key features of the derived MFDs. The chapter went further to provide a review of some of the causes of heterogeneity, which resulted in less scattered MFDs. The driver's adaptivity to various states of traffic was considered a major component as discussed under 2.5.3.

Given that the drivers behaviour is key input requirement for simulation models, further research and investigations is required and recommended for reliable analysis of the simulation models. Based on the literature reviewed, it is clear that the limitation of the exiting research work is that most have been carried out using simulation models or using loop detector data obtained from freeway networks. There are number of studies from signalized arterial networks using empirical data.

Courbon and Lecrercq (2011) showed that for arterial networks, the distance between loop detectors and the downstream traffic signal has a significant impact on the shape of the macroscopic fundamental diagram. As a result, several studies have encouraged the use of vehicle's trajectory to estimate the MFD (Leclercq et al., 2014; Saberi et al., 2014). However, due to the limited access to probe vehicle data, a widely applicable method is required for the empirical estimation of MFD.

In conclusion, this research study will seek to address the following questions;

a) How to utilize the available empirical data to estimate the MFD of a freeway network?

b) What features will the freeway network MFD exhibit and,

c) How should the network be partitioned for better control and monitoring?

Chapter 3 Research Methodology

This chapter commences with the presentation of the methodology framework. Thereafter, the data requirements, data collection method and limitations are presented and discussed. It concludes with the presentation of the data analysis process including the analytical formulae used to estimate the MFD for the City of Cape Town freeway network.

3.1. Research Methodological Framework

The research methodology framework was structured to ensure that the research objectives set out earlier in Chapter 1, are achieved. It illustrates the key concepts and relations between them as illustrated in Figure 3-1.

The input data was collected using the Cape Town Freeway Management System (FMS) fixed loop detectors positioned at various location along the freeways found within the study area. Fixed loop detectors are the only source of traffic data for this research. Data collected was organized and analysed to obtain the required variables and parameters in the preferred order and arrangement. Microsoft Excel spreadsheets were used to analyse and run various analysis scenarios.

The excel spreadsheets were also used to aggregate and analyse individual loop detectors information. Two types of MFD were derived; the whole month MFD and the daily MFDs. The relationship between average traffic flow and density for the entire network for the two types of MFDs was established and analysed.

The analysis of the MFD was conducted in conjunction with the average flow and density time-series. The impact of traffic demand on the shape of the daily MFDs was explored. Furthermore, the impact of partitioning the network into two sub-networks was explored and implications thereof discussed.

The following research methodology was followed to empirically estimate the MFD for the City of Cape Town freeway network;

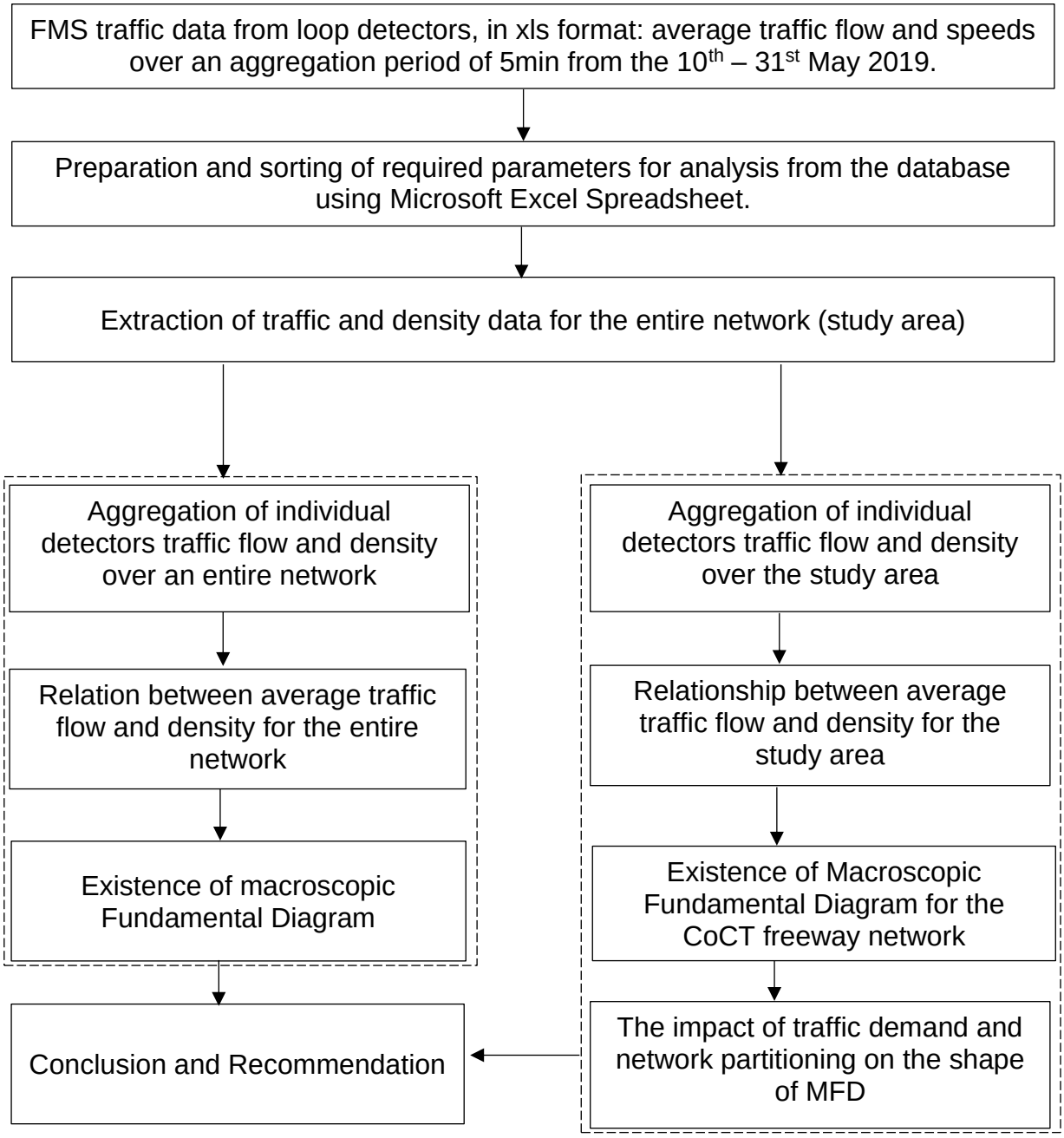


Figure 3-1: Research Methodological Framework

3.2. Data Requirement and Collection Method

Loop detectors data was utilised for the estimation of the macroscopic fundamental diagram for the City of Cape highway network. Data was collected from selected sections of the National Routes N1, N2, N7 and R300. Attempts were made to source loop detector data from major arterials such as the M5 and M3 from the City of Cape Town Municipality with no success.

In total, 39 fixed loop detectors positioned at various locations along the chosen corridors were identified, of which 7 were malfunctional due damages caused by construction works. The data was made available by SANRAL through its Freeway Management System (FMS). It was recorded every 5 minutes from the 10th to 31st May 2019, Monday to Sunday. The variables recorded included average flow q_i and vehicle average speeds over a 5 minutes interval.

3.3. Data Limitations

Data collected by fixed loop detectors faces various forms of limitations and data used for this study was no exception. As mentioned earlier, 39 functional fixed loop detectors positioned at various locations along the chosen corridors were identified, of which 7 were malfunctional due damages caused by construction works. Missing and incomplete data from malfunctioning detectors were not utilized in the estimation of the MFD.

Loop detectors which were positioned on the on/off-ramps were not considered. Considered also, was the reality that loop data system may fail at some time (power failures, errors in transmission etc.) and some periods might have found to be missing in the data. The missing data, for any period not more than 15 minutes, was replaced with data obtained from interpolation. If the data was not recorded for more than 15 minutes, the data was considered missing and excluded from the data analysis.

Data collected was analysed using Microsoft Excel. First, only required and relevant attributes were extracted from the raw dataset. The required attributes were; (a) geographical location of the loop detectors, (b) records date and time, (c) number of lanes per direction and (d) average number of vehicles and speeds over an aggregation period of 5min per loop detector.

3.4. Conversion: Volume to Flow

The raw dataset provides traffic flow and densities per direction. The first calculation step is therefore to divide the flow per direction by the total number of lanes in that direction within that particular section of the road (see Figure 3-2) using the following equation;

$$q_L = \frac{q_D}{n} \qquad (3\text{-}1)$$

Where, n is the number of lanes per direction, q_L and q_D are the flow per lane (in veh/lane) and flow per direction (in veh/direction), respectively.

To demonstrate this, an example of information recorded by loop detector 614 on the 10th May 2019 between 6:00-6:05 is used. Table 3-1 shows the conversion of flow and density per direction to flow and density per lane while Figure 3-2 shows the number of lanes per direction where loop detector 614 is located.

Table 3-1: Conversion of Flow and Density at Loop Detector 614, 2019/05/10 at 6:05

Detector No.	Date (Y/M/D)	Time (hh:mm)	Flow (veh/hr)		Density (veh/km)	
			Per Direc	Per Lane	Per Direc	Per Lane
615	2019/05/10	6:05	1884	942	20.70	10.36
614	2019/05/10	6:05	1836	918	20.40	10.20

$$q_L = \frac{q_D}{n} = \frac{1836}{2} = 918$$

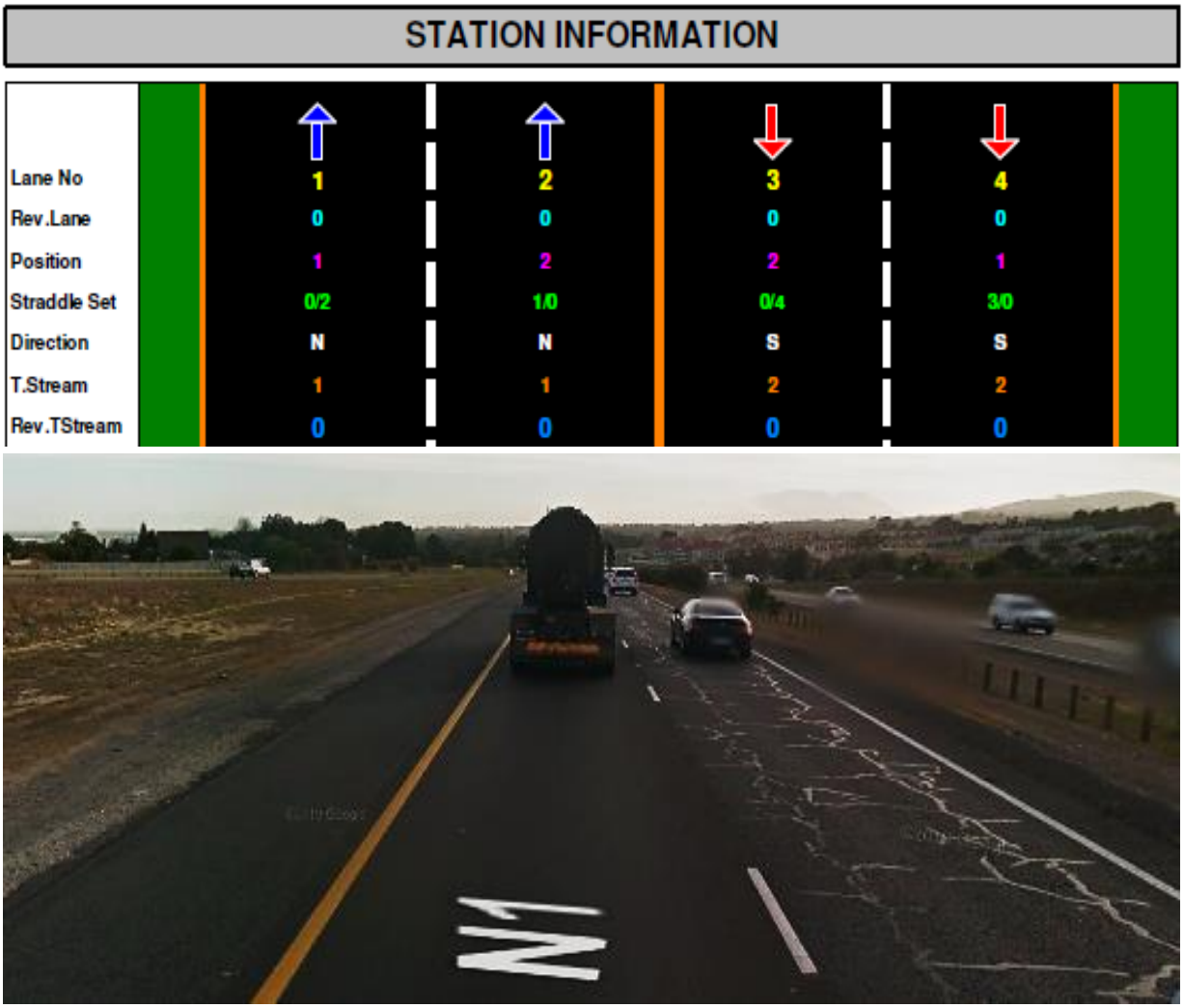


Figure 3-2: Number of Lanes Per Direction, Loop Detector 614

3.5. MFD Definition

In defining the MFD, each plotted point is the pair (average density $\bar{k}_t$, average flow $\bar{q}_t$) observed during each period t for each day. The average density and average flow are defined by Equation (3-2) and (3-3), respectively (Wang *et al.*, 2015). The analytical formulae used by Wang *et al.*, (2015) to estimate the MFD for Sendai network are consistent with Buisson and Ladier (2009)'s Equation (2-26) and (2-27) used to estimate MFD for the Toulouse network.

The average density $\bar{\mathrm{k}}_{\mathrm{t}}$ (the horizontal coordinate of MFD) is defined as the average value of the total flow during period t, which is the sum of the $\mathrm{k}_{\mathrm{t}}^{\mathrm{i}}$ observed by loop detector i for the entire network - see Equation (3-2). Whereas the average flow $\bar{\mathrm{q}}_{\mathrm{t}}$ (the vertical coordinate) is defined as the average value of the total flow during period t, which is the sum of the $\mathrm{q}_{\mathrm{t}}^{\mathrm{i}}$ observed by loop detector i for the entire network - see Equation (3-3).

$$\bar{k}_t = \frac{\sum_{i=1}^{I_0} k_t^i}{|I_0|} \qquad (3\text{-}2)$$

$$\bar{q}_t = \frac{\sum_{i=1}^{I_0} q_t^i}{|I_0|} \qquad (3\text{-}3)$$

Where $|I_0|$ in Equation (3-2) and (3-3) is the number of loop detectors in set I_0 (see Figure 4-2).

To demonstrate this, table 3-2 below shows how the average density $\bar{k}_t$, the average flow $\bar{q}_t$, the aggregation period t and the total number of loops across the network are determined.

Table 3-2: Calculation of the average density and average flow

Detector No.	Detector Name	Date Time	Duration (HH:MM)	Time (HH:MM:SS)	Vehicles (No)	Ave. Speed (km/hr)	Flow (veh/hr)		Density (veh/km)	
							Per Direc	Per Lane	Per Direc	Per Lane
614	Okavango	2019/05/10 6:05	0:05	6:05:00 AM	153	90	1836	918	20.40	10.20
616	Van Riebeeck I/C	2019/05/10 6:05	0:05	6:05:00 AM	84	100	1008	504	10.08	5.04
617	Kraaifontein	2019/05/10 6:05	0:05	6:05:00 AM	90	98	1080	540	11.02	5.51
619	Klapmuts I/C	2019/05/10 6:05	0:05	6:05:00 AM	59	108	708	354	6.56	3.28
621	Franschoek I/C	2019/05/10 6:05	0:05	6:05:00 AM	21	102	252	126	2.47	1.24
711	Spine Road	2019/05/10 6:05	0:05	6:05:00 AM	229	88	2748	1374	31.23	15.61
714	Gordons Bay	2019/05/10 6:05	0:05	6:05:00 AM	17	102	408	204	4.86	2.43
880	Akasia Park	2019/05/10 6:05	0:05	6:05:00 AM	89	89	1068	534	11.61	5.80
881	Milnerton	2019/05/10 6:05	0:05	6:05:00 AM	27	107	644	548	17.87	5.96
947	Baden Powell	2019/05/10 6:05	0:05	6:05:00 AM	52	107	248	624	12.24	6.12
982	Oostenberg	2019/05/10 6:05	0:05	6:05:00 AM	30	104	720	360	6.61	3.30
991	Paarl I/C South	2019/05/10 6:05	0:05	6:05:00 AM	25	104	600	300	5.77	2.88
1275	Winelands WIM	2019/05/10 6:05	0:05	6:05:00 AM	63	109	756	378	6.94	3.47
1399	Hoogstede	2019/05/10 6:05	0:05	6:05:00 AM	103	104	1236	412	11.88	3.96
1402	Hindle Rd New	2019/05/10 6:05	0:05	6:05:00 AM	124	107	1488	496	13.91	4.64
1403	Kuilsrivier South Ne	2019/05/10 6:05	0:05	6:05:00 AM	12	105	144	48	1.37	0.46
1442	Stellenbosch Rd	2019/05/10 6:05	0:05	6:05:00 AM	75	108	900	300	8.33	2.78
1443	Roosendal New	2019/05/10 6:05	0:05	6:05:00 AM	133	105	1596	532	15.20	5.07
1464	N7 Melkbos	2019/05/10 6:05	0:05	6:05:00 AM	49	110	588	294	5.35	2.67
1986	Joostenbergvlakte IC	2019/05/10 6:05	0:05	6:05:00 AM	62	107	744	372	6.95	3.48
2324	Huguenot Plaza 1 New	2019/05/10 6:05	0:05	6:05:00 AM	11	76	132	132	1.74	1.74
2325	Huguenot Plaza 2 New	2019/05/10 6:05	0:05	6:05:00 AM	11	75	132	132	1.76	1.76
8227	Kuilsrivier north ne	2019/05/10 6:05	0:05	6:05:00 AM	90	104	1080	360	10.38	3.46
8243	Swartklip IC WB On	2019/05/10 6:05	0:05	6:05:00 AM	252	89	3024	1512	33.98	16.99
8243	Swartklip IC WB On	2019/05/10 6:05	0:05	6:05:00 AM	89	93	1068	534	11.48	5.74
18002	ST-J_N007_01_20-1	2019/05/10 6:05	0:05	6:05:00 AM	45	104	540	270	5.19	2.60
18011	N1 Breerivier	2019/05/10 6:05	0:05	6:05:00 AM	14	97	168	168	1.73	1.73
18012	ST-J_N001_02_25-9	2019/05/10 6:05	0:05	6:05:00 AM	18	95	216	108	2.27	1.14

t
Aggregation period of 5 minutes

$|I_0|$
Total number of loop detectors across the network

$\sum_{i=1}^{I_0} q_t^i$
Sum of the q_t^i observed by loop detector i for the entire network

$\sum_{i=1}^{I_0} k_t^i$
Sum of the k_t^i observed by loop detector i for the entire network

3.6. Aggregated Lane Data

Section 2.6.2 (a) discussed how Geroliminis and Daganzo (2008) demonstrated that disaggregated lane data results in a highly scattered plot of flow and occupancy for a whole weekday with time slices of 5 minutes.

The flow and density plot of disaggregated data of two different loop detectors positioned along the same highway, the National Route N1 for a weekday with a time slice of 5 minutes will be plotted and results will be compared with Geroliminis and Daganzo (2008)'s observations.

3.7. Whole Month and Daily MFDs

Wang *et al.*, (2015)'s Equations *(3-2*) and *(3-3*) will be used to plot the monthly and daily MFDs, the relation between average flow $\bar{q}_t$ and density $\bar{k}_t$, for the City of Cape Town highway network. The results for the daily MFDs will be further analysed and discussed in detail. The analysis will entail exploring the impact of changes in traffic demand levels and network partitioning on the shape and features of the MFD.

3.8. Traffic Demand Impact on the Daily MFD.

Wang *et al.*, (2015) showed that changes in traffic demand levels had an impact on the shape and features of the MFD for Sendai road network. The impact of changes in the level of traffic demand on the shape and features of the MFD for the City of Cape Town freeway network will be analysed and discussed in detail. The analysis will involve derivation and comparison of MFDs on Saturday and Sundays during which the traffic demand levels are lower than the weekdays traffic demand levels.

3.9. Partitioning of the Highway Network

As discussed earlier in section 2.6.2 (d), Buisson and Ladier (2009)'s paper showed that; (a) Given a data set that is homogeneous in term of road-type, separating the ring roads network from penetrating roads network is a key factor in obtaining minimal scatter in the MFD and (b) the hysteresis-like shape observed in highway networks is due to spatially non-homogenous evolution of congestion.

The Cape Town freeway network will be partitioned into two sub-networks of penetrating freeways (National Routes N1, N2, N7) and ring freeway (R300), shown in blue and red colours in sed in chapter 5.

Figure 3-3, respectively. The results of the analysis for two selected Fridays will presented and discussed in chapter 5.

Figure 3-3: Cape Town freeway network, red (ring freeway) and blue (penetrating freeway).

Source: (Google Earth)

Chapter 4 Research Study Area

This chapter presents a brief overview of the study area, the City of Cape Town in South Africa. It provides a brief description of the geographical location, demographic dynamics and road network infrastructure.

4.1. Geographic and Demographic Information

The study area, City of Cape Town, is the port city on South Africa's southwest coast. Its geographical coordinates are 33°49'31.2" South 18°36'33.2" East with an average elevation of 30m above mean sea level. The study area (see Figure 4-11), has an estimated surface area of 750 km^2 with a population density of 1530 per km^2. From 2001 to 2011, Cape Town's population increased by 2.6%, reaching 3.75 million inhabitants with the current population being estimated to be 4.5 million (City of Cape Town, 2019).

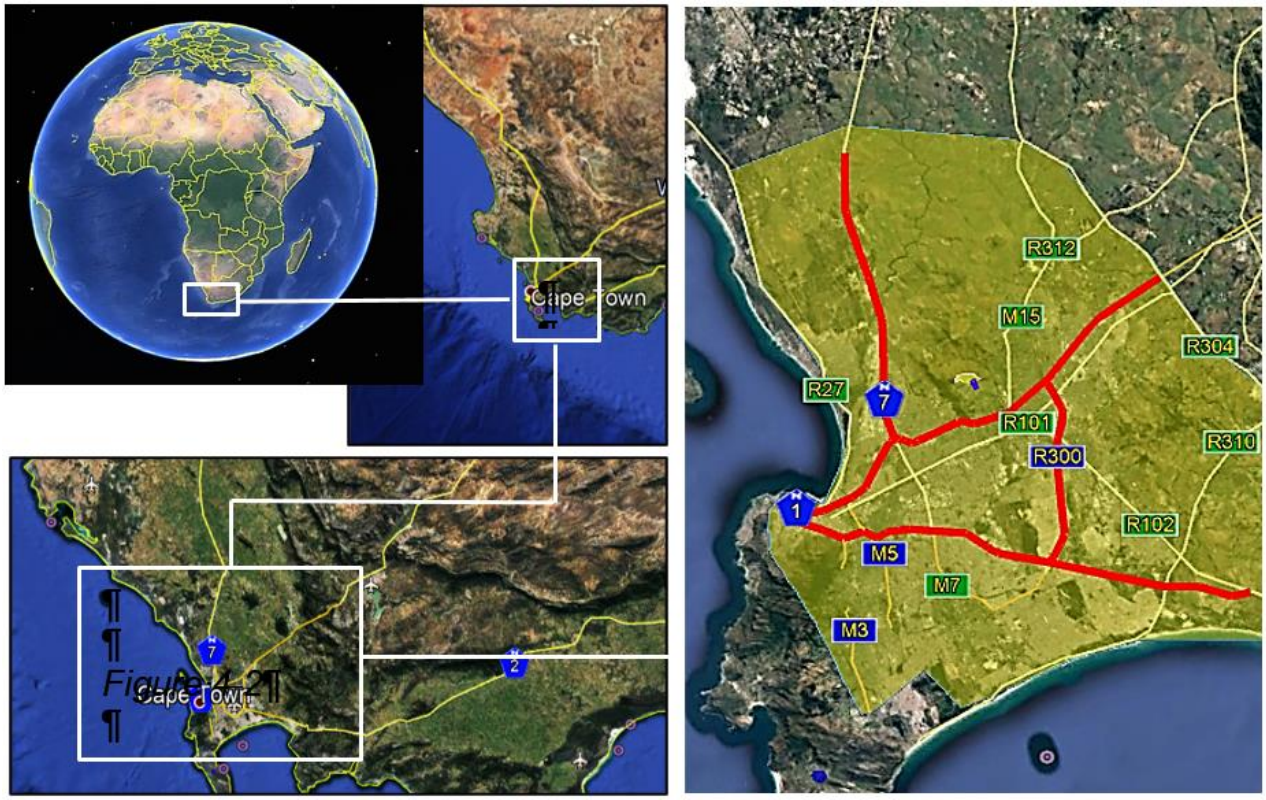


Figure 4-1: Study Area, Cape Town.

Source: (Google Earth)

4.2. Road Network Infrastructure

The City of Cape Town's roads infrastructure is the city's largest assets, currently valued at R92 billion. About 155 km of freeways are concurrently owned and managed by the CoCT, SANRAL and the Western Cape Provincial Roads Department (City of Cape Town, 2019). The Freeway Management System (FMS) is used to monitor the strategic and economic road infrastructure with more than 250 cameras and Variable Message Signs (VMSs) for live communication with road users.

On daily basis, the Transport Management Centre monitors the FMS. Traffic management in the city is made possible by over 1 500 signalized intersections and 400 signalized pedestrian crossing. As indicated in section 3.2, the road network infrastructure includes the identified 39 loop detectors positioned along the freeways (see Table 4-1) of which 7 were found to be malfunctional.

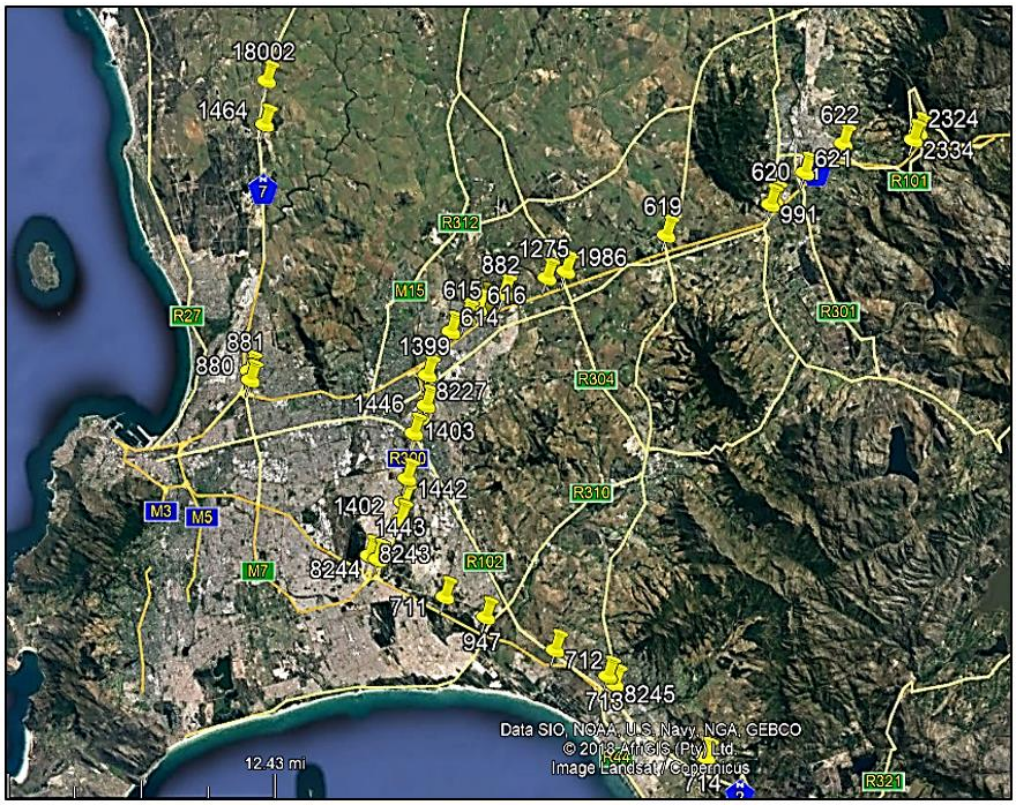


Figure 4-2: Position of 39 fixed loop detectors used for the study.

Source: (Google Earth)

The study area covers only the portions of the N1, N2, N7 and R300 highways as shown in Figure 18. Table 4-1 below describes the highways forming part of the network.

Table 4-1: Description of the freeways making up the study area

Freeway	Description
N1	The first section of the N1 is shared with the start point of the N2 in Cape Town CBD. It's a four-lane elevated freeway that runs along a strip of the City of Cape Town and the Port of Cape Town. The corridor is highly congested during the morning peak (going towards CBD) and evening peak (leaving CBD). It heads through Goodwood and Bellville suburbs where the R300 terminates at it. At Paarl, the freeway ends and the N1 is tolled as it passes through the Huguenot Tunnel running under the Du Toitskloof Mountains.
N2	The N2 corridor starts in Cape Town CBD at the northern end of Buitengracht Street. The first section of N2 is shared with the N1. It's also a four-lane elevated freeway. At the Hospital Band interchange, the N2 and the M3 merge to form a 10-lane freeway before diverging again.
N7	The N7 it's a freeway and exits City of Cape Town limits at M12 Interchange. From there is remains a dual carriageway until just after Malmesbury at R45 intersection. Beyond M19 interchange, the corridor remains a single carriageway
R300	The R300 is 22km long freeway connecting Mitchells Plain with the N2, Kuils River and the N1 at Bellville. It is a three-lane carriageway.

Chapter 5 Analysis, Results and Discussions

This chapter demonstrates the City of Cape Town highway network MFD from empirical data set using the method outlined in Section 2.6.2 (c) (Wada & Hara, 2015). The results are consistent with the outcomes of similar studies conducted in Toulouse (Buisson & Ladier, 2009), Chicago, Portland and Irvine (Saberi and Mahmassani, 2013).

The flow and density time series are presented to provide an overview of the network average flow and density dynamics during peak and off-peak periods. The flow and density plot of disaggregated lane data is demonstrated and discussed. The whole month MFD for the City of Cape Town's highway network is presented and briefly analyzed.

In order to better understand the MFD characteristics, daily (Fridays) MFDS are presented and analysed. The impact of the changes in traffic demand levels on the shape of the daily MFDs is investigated. Lastly, in order to understand the dynamics of congestion spread during both morning and evening peak periods, the network is later partitioned into penetrating highways subnetwork and ring highway subnetwork

5.1. Disaggregated Lane Data

As discussed in section 2.6.2(a) and shown in Figure 2-19(a), Geroliminis and Daganzo (2008) demonstrated that disaggregated lane data results in a scatter plot of flow and occupancy for a weekday with time slices of 5minutes.

Figure 5-1 shows flow and density plot of disaggregated data of two different loop detectors positioned along the same link, the National Route N1 for a weekday with time slices of 5minutes. The observed scatter here is consistent with the one observed in Yokohama. The disorder corresponds to that of a typical single traffic lane as observed in Yokohama.

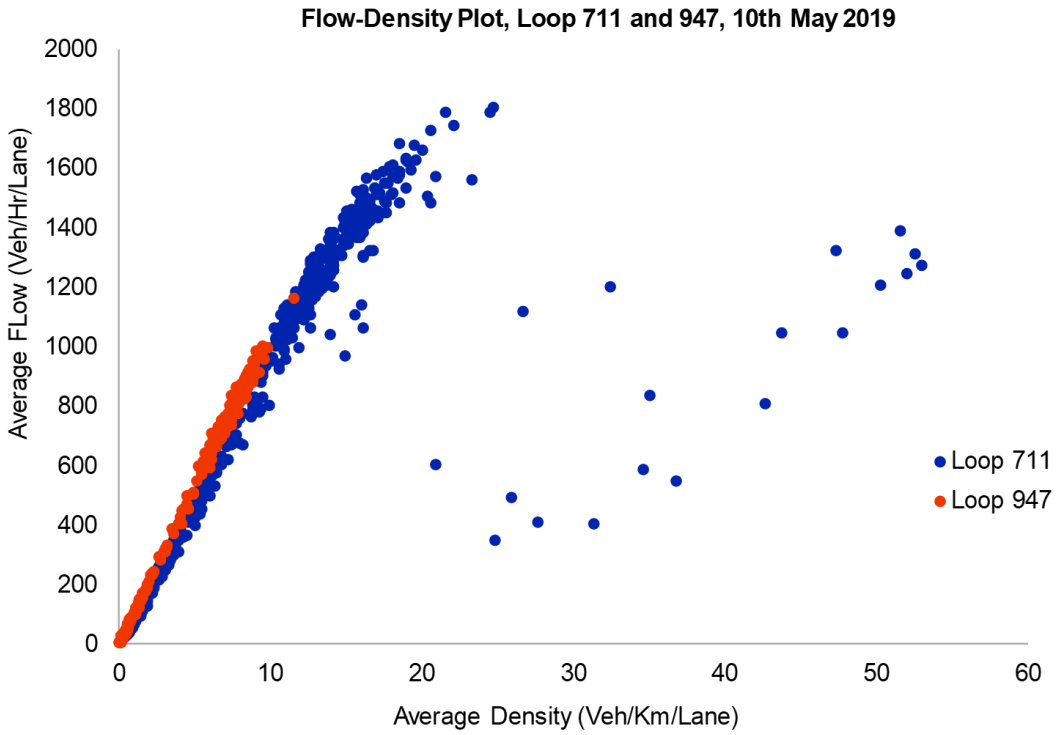


Figure 5-1: Flow-Density Plot, Loop 711 and 714, 10th May 2019

5.2. Flow and Density Time Series

Based on time series of average flow for all weekdays (Mon – Fri), it was established that Fridays exhibit more pronounced morning and evening peak periods than any other weekdays, especially the evening peak. The average flow and density time series below, Figure 5-2 (a) and Figure 5-2 (b) respectively, illustrates the traffic flow behaviour during peak and non-peak periods for the 10th, 17th, 24th and 31st May 2019 between 4h00 and 20h00.

An apparent observation is the consistent behaviour of traffic flow for the four Fridays of the month of May, shown by similar flow and density trends. It is also observed that both the evening and morning peaks are well pronounced and consistent with the city's congestion patterns. The evening peak (1026 veh/hr/lane) is more pronounced and observed to be slightly higher than the morning peak (874 veh/hr/lane). This effect is due to Fridays' evening specific origin-destination patterns which has two main motivations; leaving the city for the weekend (leisure) and returning to the main residence (commuters).

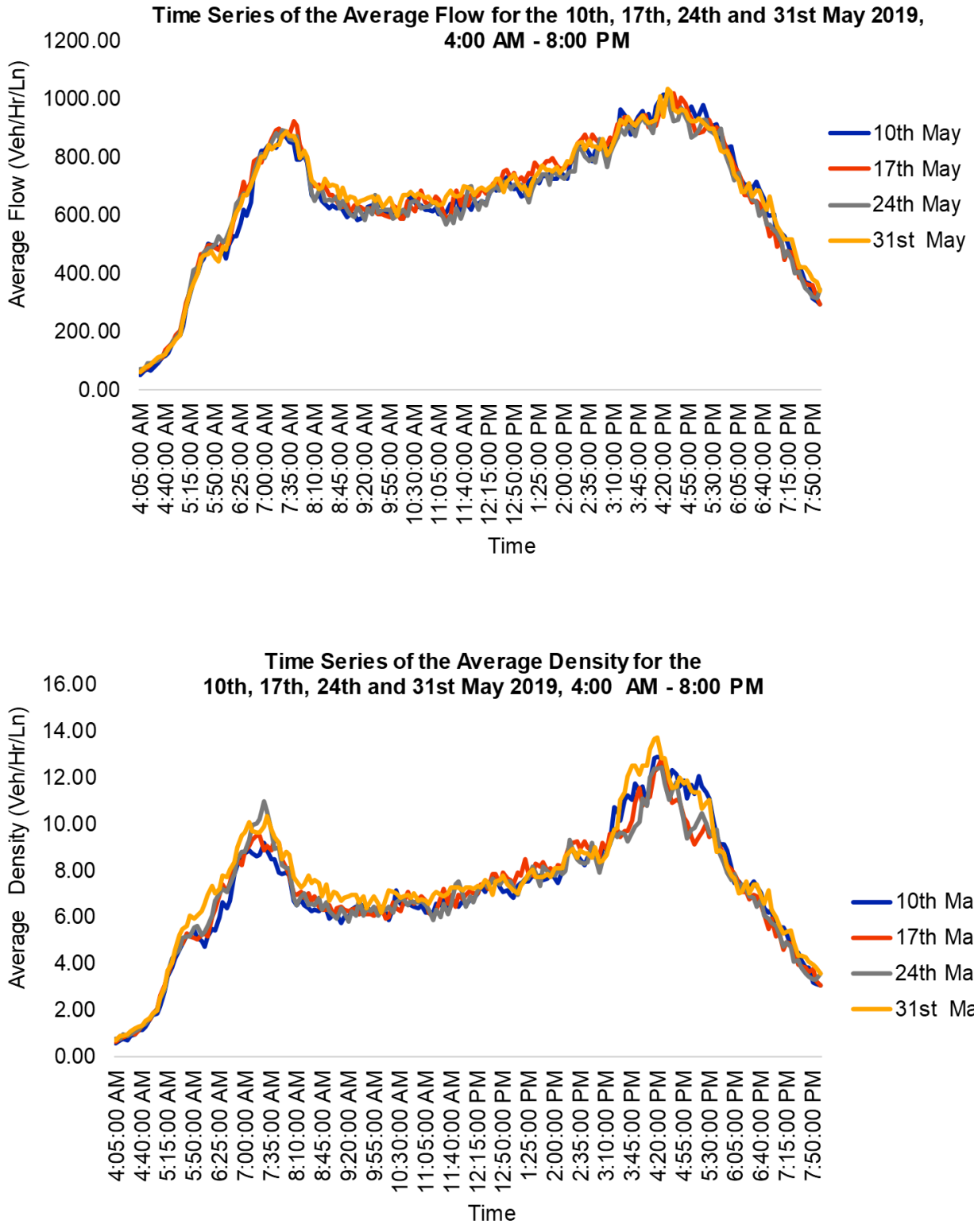


Figure 5-2: (a) Average flow time-series and (b) average density time series for the 10th, 17th, 24th and 31st 2019

5.3. CoCT Freeway Network: Whole Month MFD

Figure 5-3 below illustrates the whole month MFD for the City of Cape Town freeway network, the relationship between the network average flow $\bar{q}_t$ and density $\bar{k}_t$ as defined by Equations (3-2) and (3-3), respectively.

The results confirm the existence of the whole network MFD for the Cape Town freeway network and are consistent with outcomes of similar studies in France (Toulouse) (Buisson & Ladier, 2009) and Japan (Sendai) (Wada & Hara, 2015).

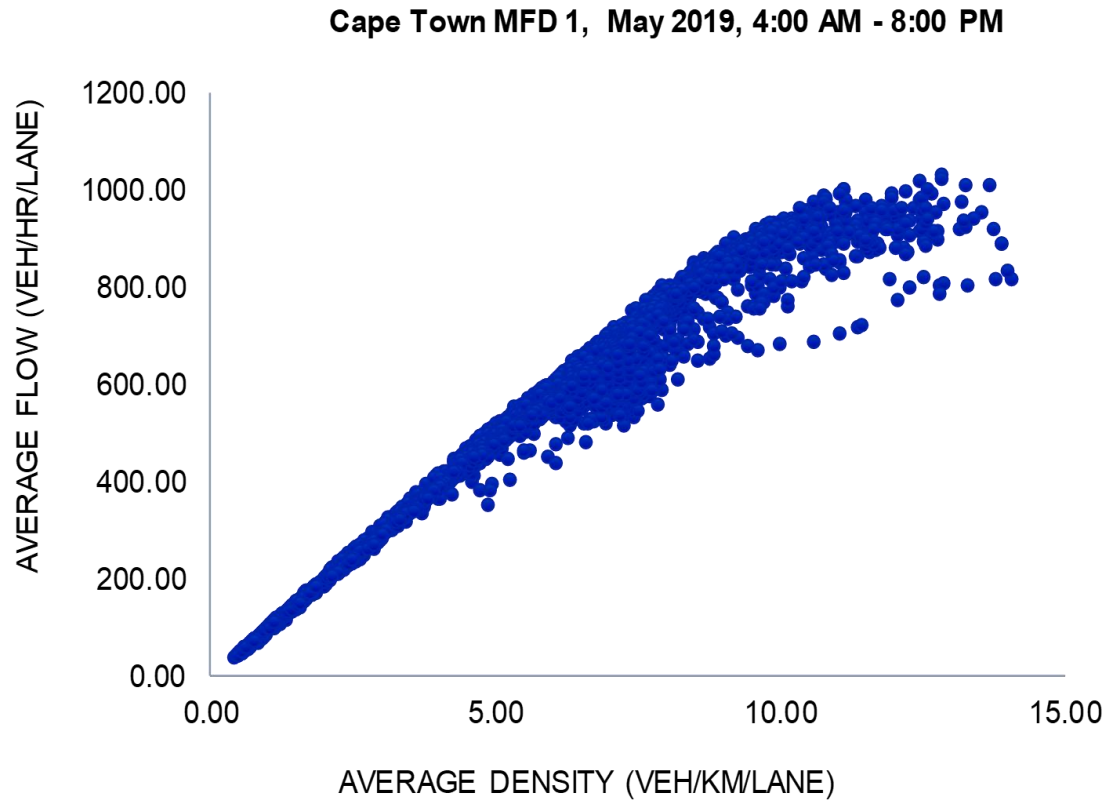


Figure 5-3: MFD for the month of May 2019, 4:00 AM – 8:00 PM

Comparing Figure 5-3 with Daganzo and Geroliminis (2008)'s Figure 2-19 (b), one can observe that the maximum flow for the City of Cape Town is almost double (1026 veh/hr/lane for the City of Cape Town and 540 veh/hr/lane in Yokohama). The critical density, corresponding with the observed maximum average flow, is much lower here.

This may be as a result of the specific arrangement and setting of loops and may have no particular importance. The main difference observed here is the scatter. Geroliminis and Daganzo figure, Figure 2-19 (b) (Daganzo & Geroliminis, 2008) exhibited no scatter whereas here a slightly higher scatter is observed. Various types of heterogeneity may have caused the observed scatter.

Comparing the Cape Town whole month MFD with the Toulouse's MFD (see Figure 2-30), one can observe that the critical density, corresponding to the maximum flow is in same order of magnitude (12.85 veh/km/lane here whereas it was about 15 veh/km/lane in Toulouse. The maximum flow is however slightly higher here i.e. 1026 veh/hr/lane in Cape Town and 600 veh/hr/lane in Toulouse.

Cape Town's whole month MFD and Toulouse exhibit the same scatter pattern and the formation of hysteresis loop. In order to better understand these observations, section 5.4. below looks closer at the daily MFDs for the four Fridays of the month of May 2019 between 4h00 and 20h00.

5.4. CoCT Freeway Network: Daily MFD

Four daily MFDs for the month of May 2019 are drawn and analysed here. Figure 5-4 (a) present the MFDs derived using data obtained from 39 loop detectors positioned along the penetrating (N1, N2 & N7) and ring (R300) freeways of the city of Cape Town. The MFD shapes observed on the 10th, 17th and 24th were very similar.

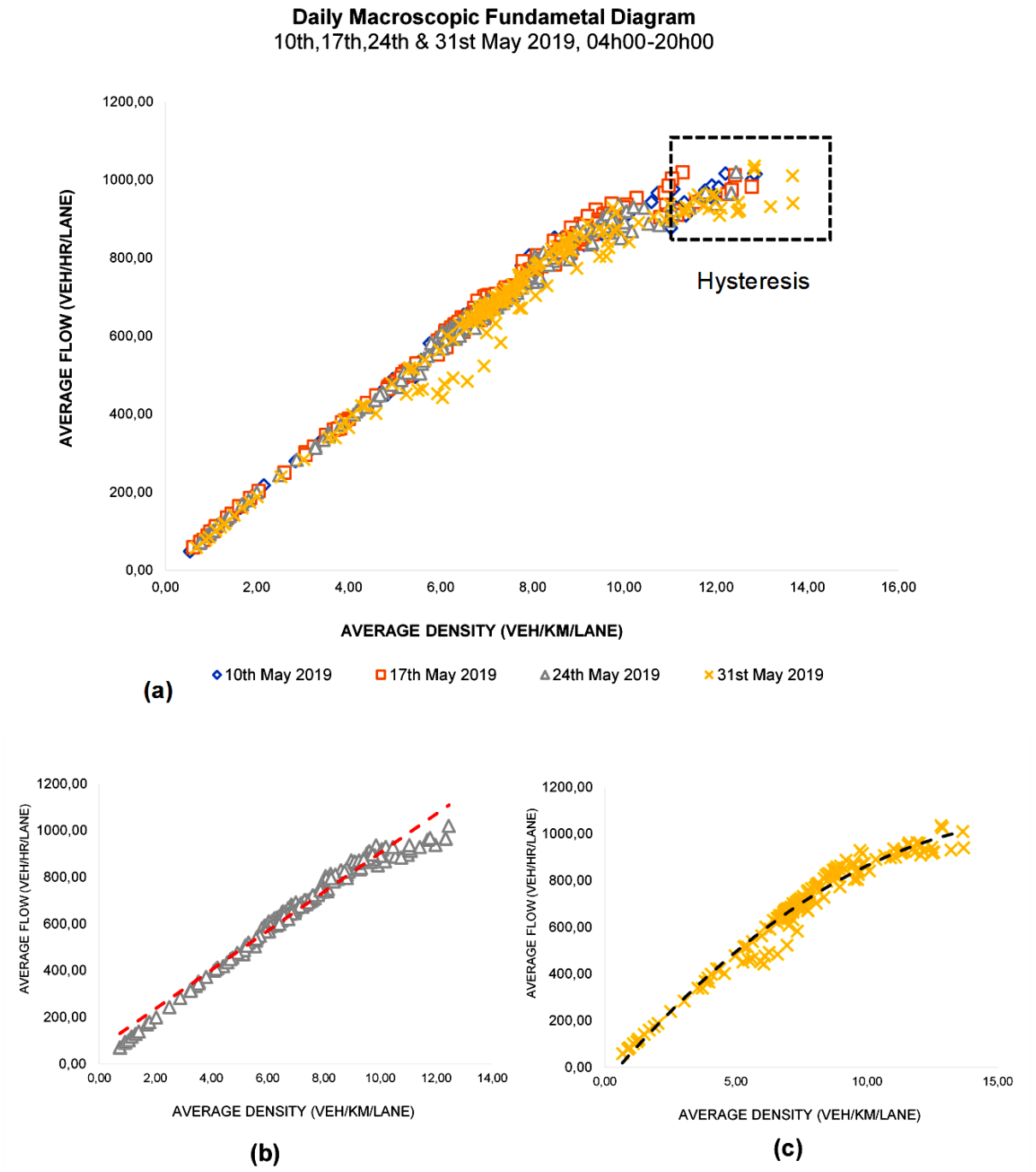


Figure 5-4: (a) Daily MFDs for the 10th, 17th, 24th and 31st May (b) Daily MFD for the 24th May and (c) Daily MFD for the 31st May 2019.

In order to simplify the representation of results, further analysis will be restricted to the two Fridays, 17th and 31st May 2019. Figure 5-4(b) and Figure 5-4(c) represents the evolutions observed for 24th May and 31st May 2019, respectively. On the first Friday (24th May), the network does not seem experience much congestion, whereas, the following Friday (31st May) shows congestion with what appears to be a hysteresis loop as shown in Figure 5-5.

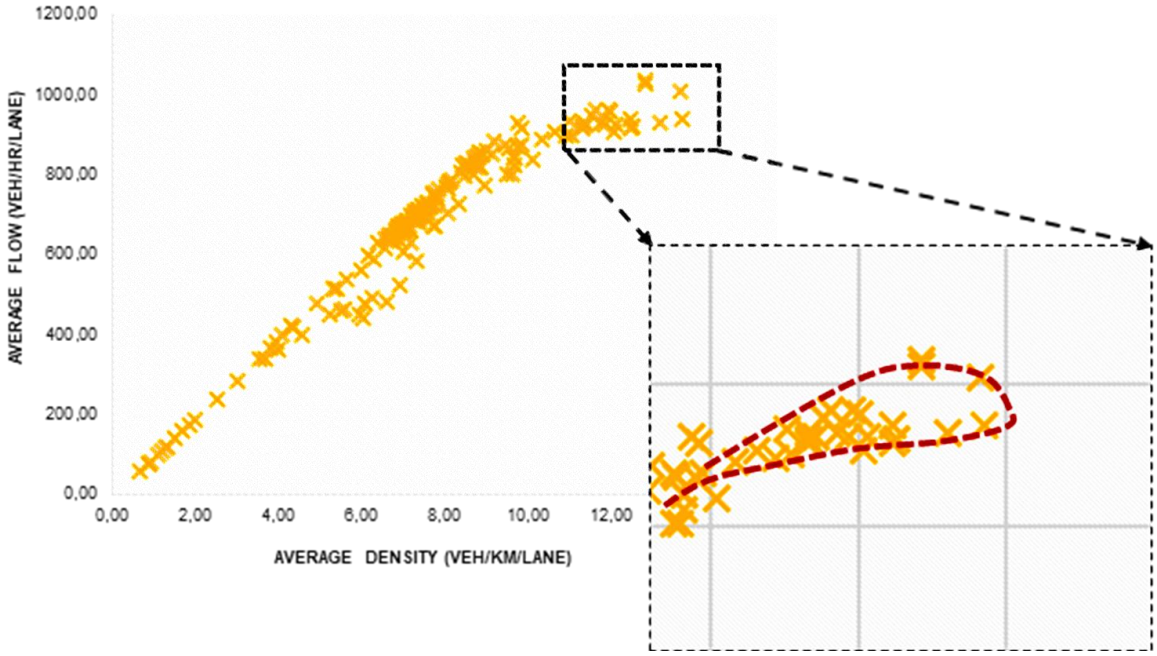


Figure 5-5: Daily MFD for the 31st May 2019

The general shape of the MFD observed on the 31st May appears to be an incomplete parabolic shape followed by the formation of what appears to be a hysteresis loop. The observed formation of the hysteresis loop past the critical density is consistent with the results of other similar studies involving freeway networks (Buisson & Ladier, 2009) (Saberi & Mahmassani, 2012) (Saberi & Mahmassani, 2013) (Knoop & Hoogendoorn, 2013) (Geroliminis & Sun, 2011).

Comparing Figure 5-5 with Buisson and Ladier (2009)'s freeway MFD shown in Figure 2-31 (c), one can observe that maximum flow is of the same magnitude (1026 veh/h/lane here whereas it is about 1250 veh/h/lane in Toulouse). The critical density here is slightly higher than the equivalent critical occupancy observed in Toulouse (Buisson and Ladier, 2009).

This observation might be because the traffic observed on the whole Cape Town freeway network does not necessarily correspond simultaneously in every site to a capacity situation. The formation of the hysteresis loop is considered a significant feature of the MFD. This feature is exhibited by both the whole month and daily MFDs as shown in Figure 5-4 and Figure 5-5 respectively.

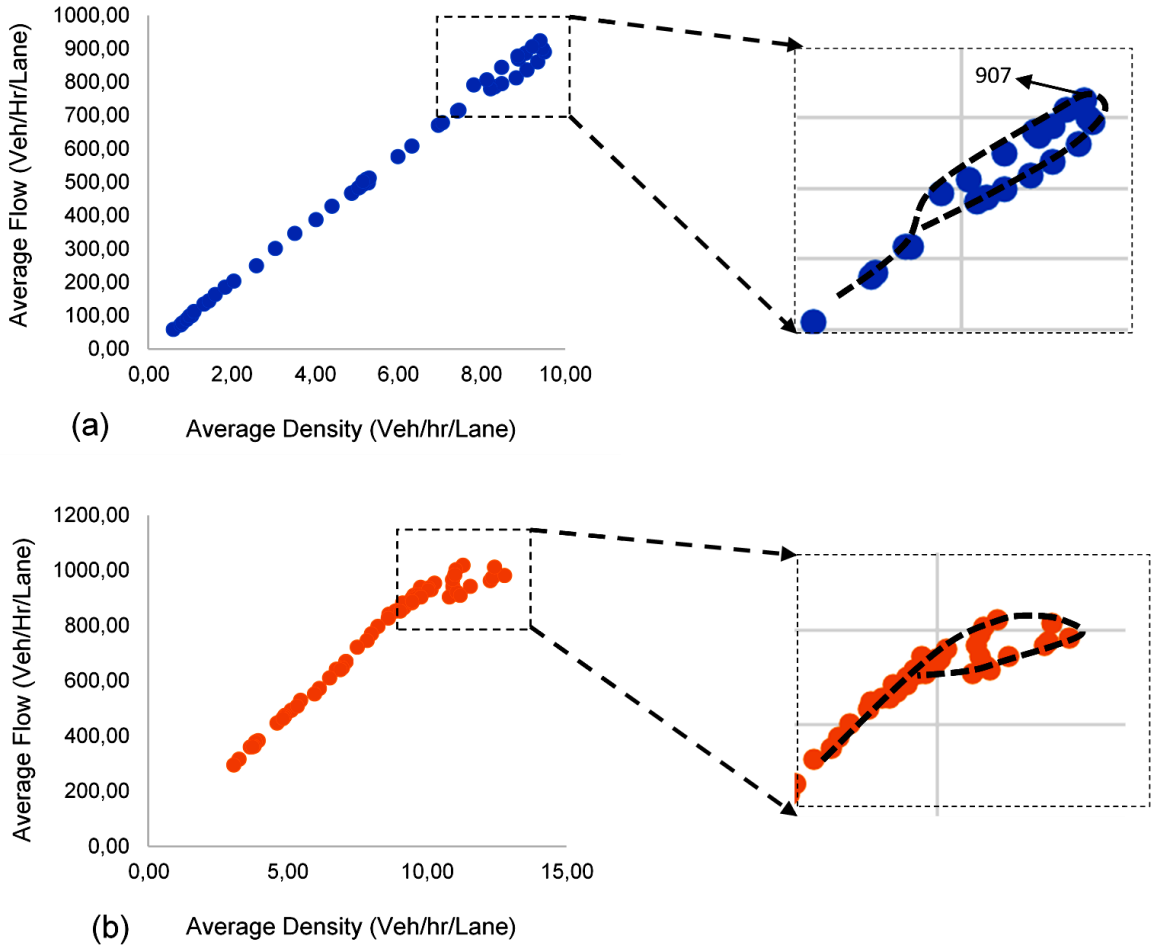
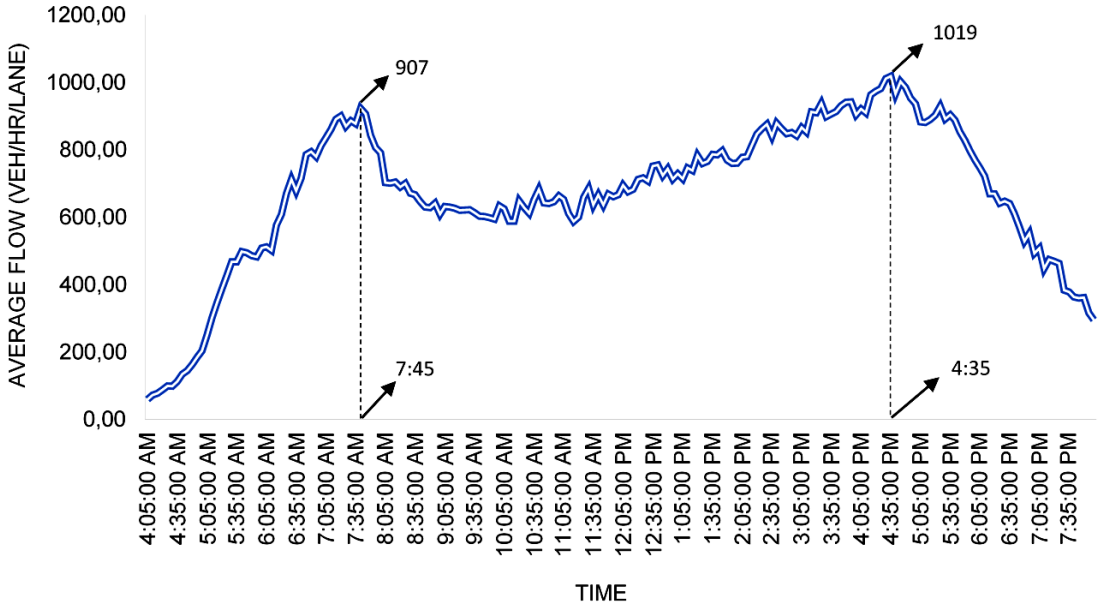
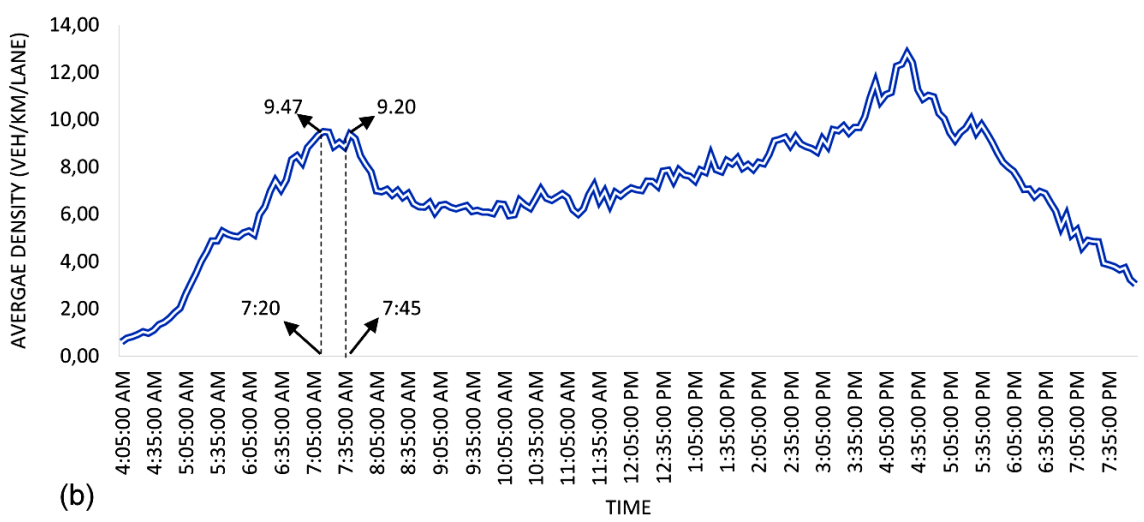


Figure 5-6: (a) MFD in the morning (4h00-12h00) and (b) evening (12h00-20h00) of 17th May 2019.



(a)



(b)

Figure 5-7: Time-series of the (a) average flow and (b) average density of 17th May 2019

The feature is expressed more clearly by diving the 16 hours of data into two parts which are the morning (04h00-8h00) and the evening (15h00-20h00) of the 17th and 31st May 2019 (Fridays). It is observed from Figure 5-6 and Figure 5-7 that;

- In the morning (4:00 – 12:00), a single loop may form (See Figure 5-6 (a)) or it may not (see 5-6(a) & 5-8(a)).
- In the evening (13:00 - 24:00), a single loop always forms (See Figures 5-6(b) & 5-8(b))

The observed weekday MFD loops can be explained in conjunction with the time-series pattern of the average flow and density (see Figures 5-7 & 5-8). First, Figures 5-6(a), 5-6(b) and Figures 5-7(a), 5-7(b) show that the coordinates of plots of the loop on the MFD correspond to the pair measurements in peak hours. It can be seen that the loop forms on the MFD in the morning from 6:45 AM – 7:45 AM and in the evening from 4:00PM – 5:00PM.

Furthermore, Figures 5-7(a), 5-7(b) show that during peak period, the average flow begins to decrease immediately after 7:45 AM, when the average flow is at its maximum value (907 veh/h/lane) and the average density remains near its maximum (9.47 – 9.20 veh/km/lane) for about 20 minutes (7:20 AM – 7:45 AM). This observation shows that the average density does not decrease significantly near the right side of the upper curve of the loop, only average flow transit to the to the lower curve, forming hysteresis loop.

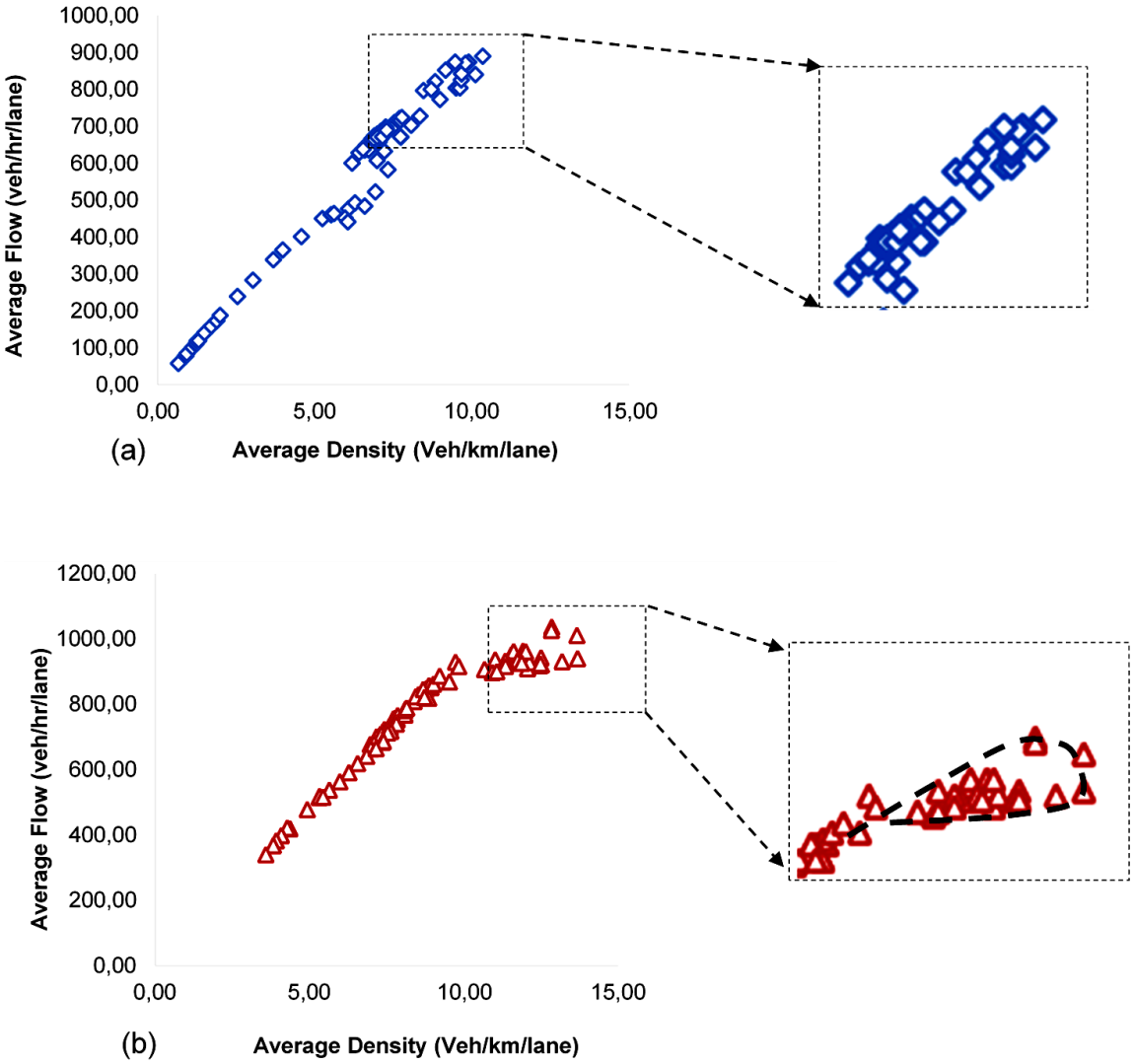


Figure 5-8: MFD in the (a) morning and (b) evening of 31st Dec 2019

The following section 5.5 explores cases where the loop does form due to changes in traffic demand condition.

5.5. Daily MFD: Traffic Demand Impact

Wang *et al.*, (2015) showed that changes in traffic demand conditions changed the shape and features of the MFD for Sendai (Japan) network. The results here show that changes in the level of traffic demand changes the shape and features of the MFD for the City of Cape Town. By observing the MFDs for the weekends throughout the month of May 2019 for the City of Cape Town freeway network, the following common features were observed;

- An explicit loop seldom appears in a Saturday MFD (see Figure 5-9(a))
- A single loop form during one day in a Sunday MFD (see Figure 5-19(a))

The MFD features observed on the 18th May (Saturday) and 19th May (Sunday) 2019 can be explained in conjunction with the features of their traffic demand. In the following, the relationship between the MFD and time-series pattern of traffic demand are analysed in detail.

Figure 5-10(b) shows that the average traffic flow on a Saturday does not exhibit any prominent peak in the morning (as it is found in weekday (Friday) MFD), it instead increases very slowly with slope S_{1f} between 1:55 PM and 5:15 PM and then swiftly decreases with slope S_{2f}. On the other hand, Figure 5-10(c), shows that the average density increases gradually with slope S_{1d} between 2:00 PM and 5:10 PM. This increase until 5:10 corresponds to the state transition on the upper curve with slope S_{1f} / S_{1d} of the MFD (see Figure 5-0(a)).

The average density remains near its maximum between 5:10 PM and 6:00 PM, and begins to decrease with slope S_{2d} at 6:00 PM, while the average traffic flow begins to decrease with S_{2f} at 5:15 PM. This decrease corresponds to the state of transition on the lower curve with S_{2f} / S_{2d} of the MFD (see Figure 5-10(a)).

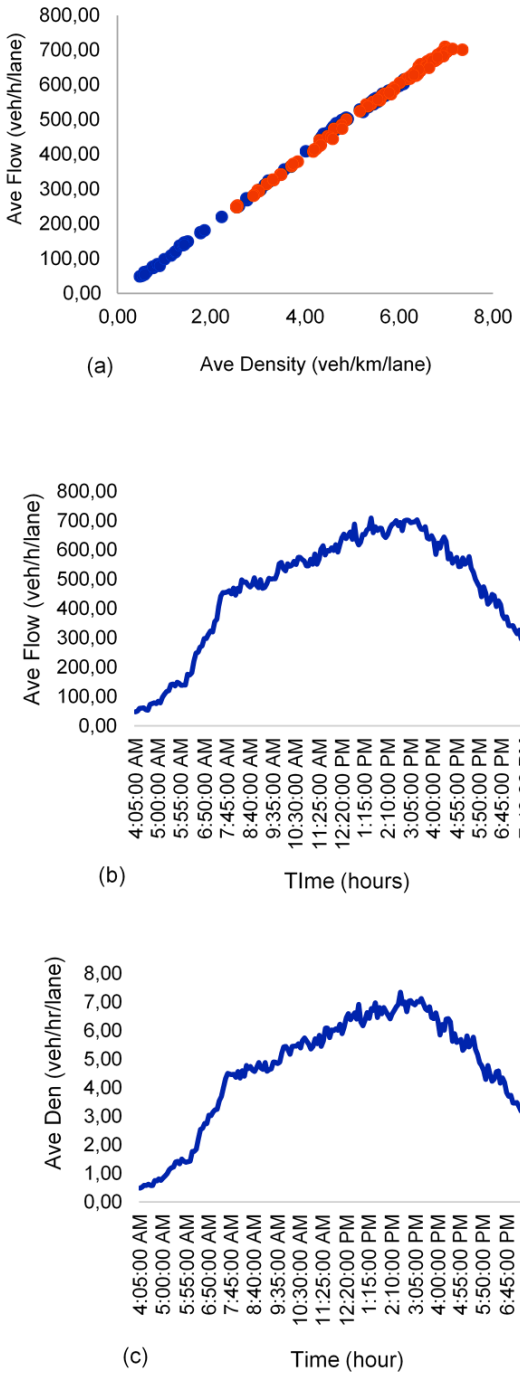


Figure 5-9: (a) MFD, (b) average flow time series and (c) average density time series on the 18th May (Saturday).

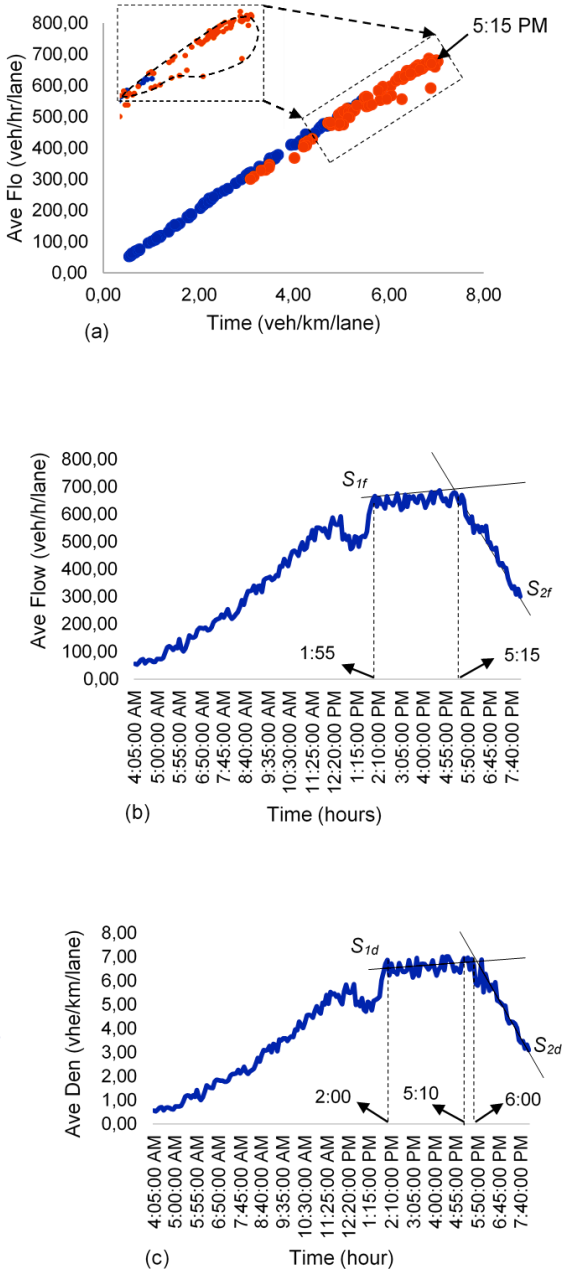


Figure 5-10: (a) MFD and (b) average flow time series and (c) average density time series on the 19th May (Sunday)

Given that the average traffic flow decreases with increasing congestion spread, the decreasing slope S_{2f} / S_{2d} of the MFD is larger than the increasing slope S_{1f} / S_{1d} of the MFD. Hence, it is observed that the state changes clockwise along the upper curve when increasing and along the lower curve when decreasing (forming a loop). The hysteresis loop does not form on Saturday (18th May 2019) because the traffic demand on that particular day was low throughout the day.

5.6. Daily MFD: Partitioning the Freeway Network

This section presents the results of partitioning the Cape Town freeway network into sub-networks of penetrating freeways (National Routes N1, N2, N7) consisting of 23 loop detectors and ring freeway (R300) consisting of 6 loop detectors, shown in blue and red colours in Figure 3-3.

The results of the analysis are presented in Figure 5-11 and Figure 5-12 for the two selected Fridays. The maximum average flow for the two sub-networks are observed to be significantly different from each other; 943 veh/h/lane for penetrating freeways and 1539 veh/h/lane for the ring freeway as shown in Figures 5-11(a) and 5-12(a), respectively. The congestion observed on the penetrating freeways is not sufficiently spread in the set of 23 loop detectors located on those freeways; the minimum congestion flow is of 697 veh/h/lane, that is 75% of the maximal observed flow (see Figure 5-11(a))

Further analysis of the penetrating freeway network reveals that on the 24th May, congestion is observed during both the morning and evening peak periods (see Figure 5-11(b)) whereas on the 31st May congestion is observed mainly during the evening peak period (see Figure 5-11 (c)), with hysteresis-like shape. Thus, the penetrating freeway network is not homogeneously congested during the evening peak period of the 31st May 2019. This heterogeneity in the congestion level is probably the reason for the hysteresis-like shape.

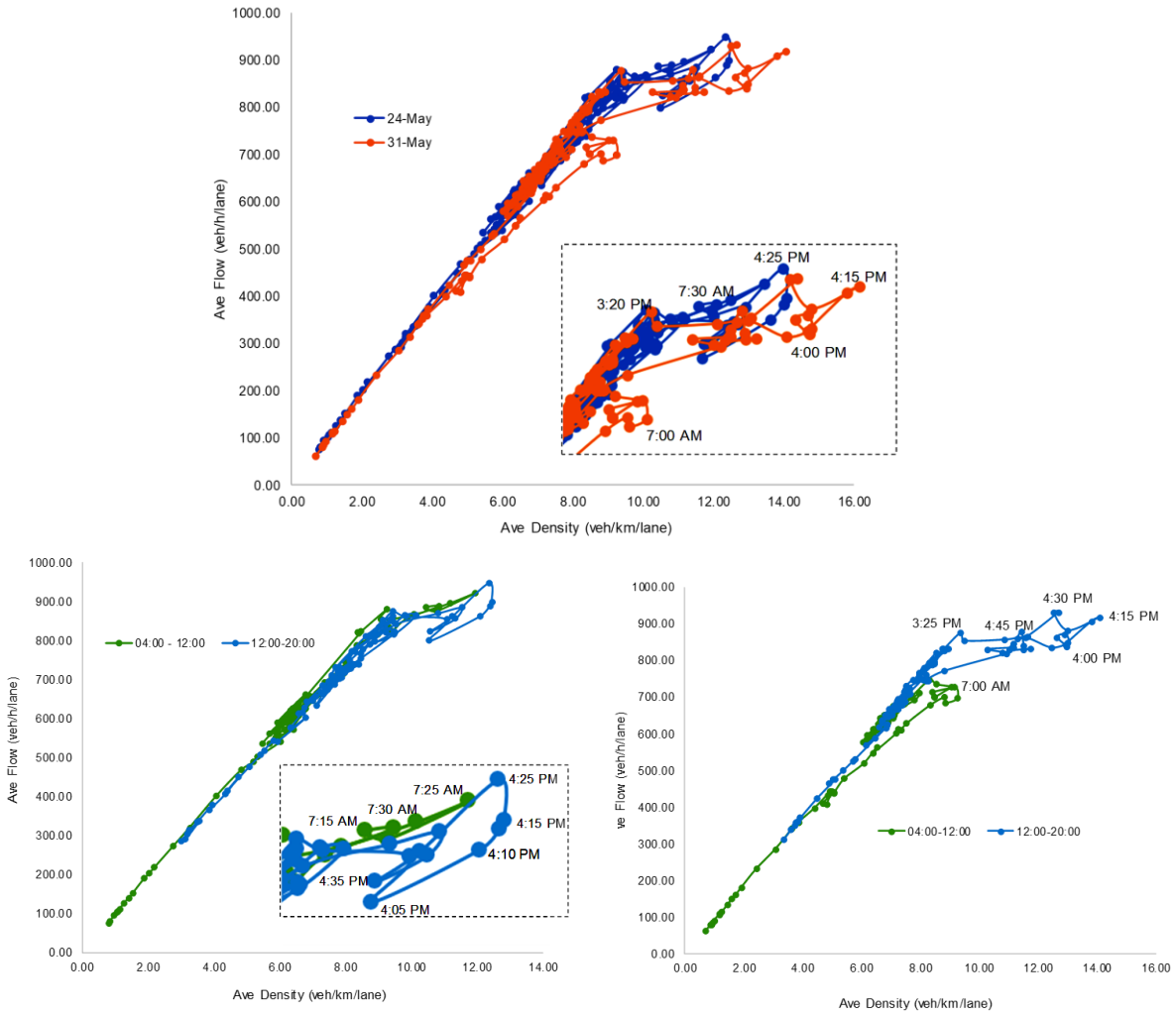


Figure 5-11: Penetrating freeways network (a) MFD for the 24th and 31st May, (b) MFD for the 24th May and (c) MFD for the 31st May

The ring highway maintains a free-flow traffic state during both the morning and peak periods of the 24th and 31st May 2019 (see Figure 5-12(a)), with relatively high average traffic flow of 1522 veh/h/lane observed during the evening peak period of both the 24th and 31st May 2019. Comparing the City of Cape Town's ring freeway MFD (see Figure 5-12(a)) with the Toulouse's ring freeway MFD (See Figure 2-32(b)) , it can be seen that the MFDs for the two ring freeways indicates mainly the free-flow regime with observed maximum average flows being of the same order of magnitude (1522 veh/hr/lane here and 1600 veh/hr/lane in Toulouse). An interesting observation was the significant difference in the average traffic flows observed during the morning peak of the 24th and 31st May 2019. The average traffic flow observed during the morning peak period of the 31st May 2019 was 72% higher than the average traffic flow observed on the 24th May 2019 (See Figure 5-12(b)).

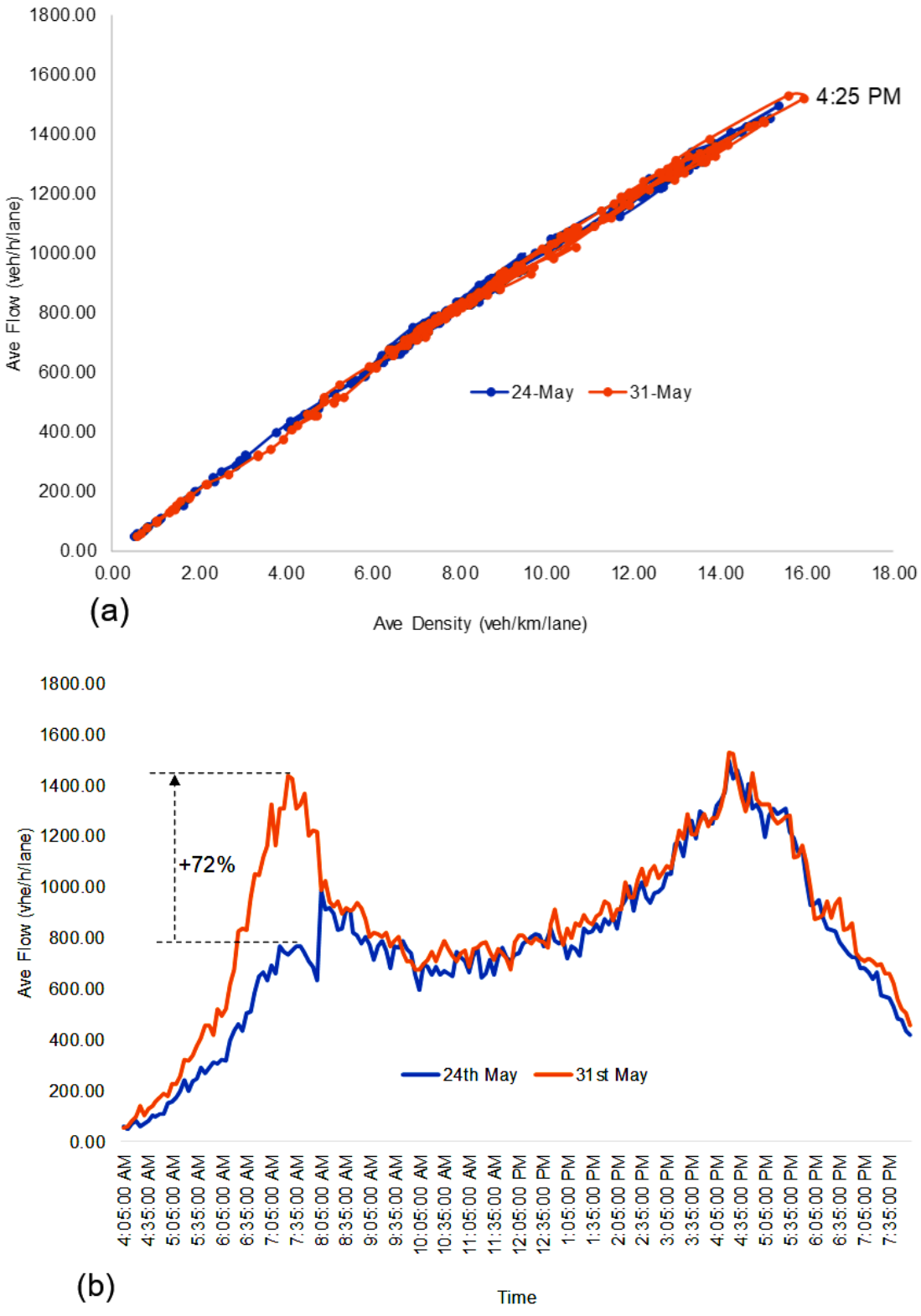


Figure 5-12: (a) Ring freeway MFD and (b) average flow time-series for the 24th and 31st May 2019

5.7. Further Characterization and Interpretation of the Congested Penetrating Freeways Network

Partitioning the network into penetrating freeways network and ring high network has shown that the ring freeway network maintains a free-flow state (no congestion observed on network level) during both the morning and evening peak periods whereas the penetrating freeways network was found to be congested with the formation of hysteresis loops. This section therefore seeks to further characterize and interpret the hysteresis phenomenon observed in the penetrating freeway network. The hysteresis loops observed will be characterized in terms of its size and shape as discussed in sub-section 2.3.3 (Saberi and Mahmassani, 2013).

5.7.1. Penetrating Freeway Network MFDs: Comparison Across Different Days

This sub-section first investigate whether some of the key findings of Buisson and Ladier (2009), Saberi and Mahmassani (2013), Saberi and Mahmassani (2012) and Geroliminis and Sun (2011a) for previously studied freeway networks are consistent with the findings of the Cape Town's penetrating freeways network.

Figure 5-13 and Figure 5-14 illustrates that the 24th and 31st May MFDs for the penetrating freeways network shows that due to the presence of the hysteresis effect, the MFD is not well defined in the same way as the Geroliminis and Daganzo (2008)'s MFD as shown in 2-19 (b).

In addition, the observed scatter and hysteresis pattern here differ slightly in size and shape across different days, similar to Saberi and Mahmassani (2013)'s Chicago freeway network MFD shown in Figure 2-9. The slight difference in the pattern of the hysteresis observed across the two different days could be associated with the varying spatial distribution of congestion over the network for each of the days.

Furthermore, as earlier demonstrated in Geroliminis and Sun (2011a) and Saberi and Mahmassani (2012), the results here confirm that, for the same average density, the average flow is higher during the loading period (morning) than the recovering period (evening) as shown in Figure 5-13.

Moreover, the hysteresis loops observed in the MFD for the 24th May 2019 is made up of two consecutive hysteresis loops of the morning and evening loading-recovery cycles and each cycle of network loading-recovery that resulted in the formation of a hysteresis loop as shown in Figure 5-13 and consistent with Saberi and Mahmassani (2013) findings (see Figure 2-10). The same findings were also observed by Geroliminis and Sun (2011a) and confirm their statement that the "evolutions of traffic states follow different paths during the onset and offset of the morning and evening peaks".

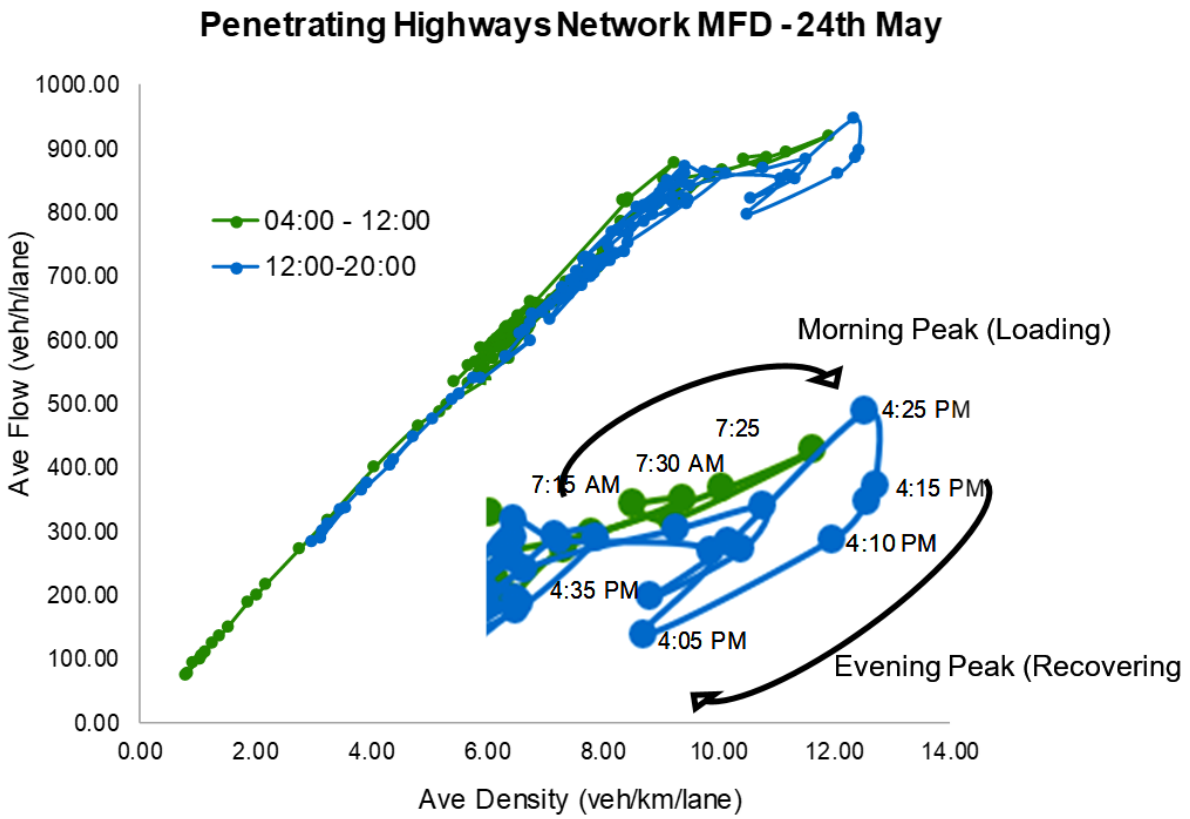


Figure 5-13: Penetrating Freeways Network MFD for the 24th May 2019

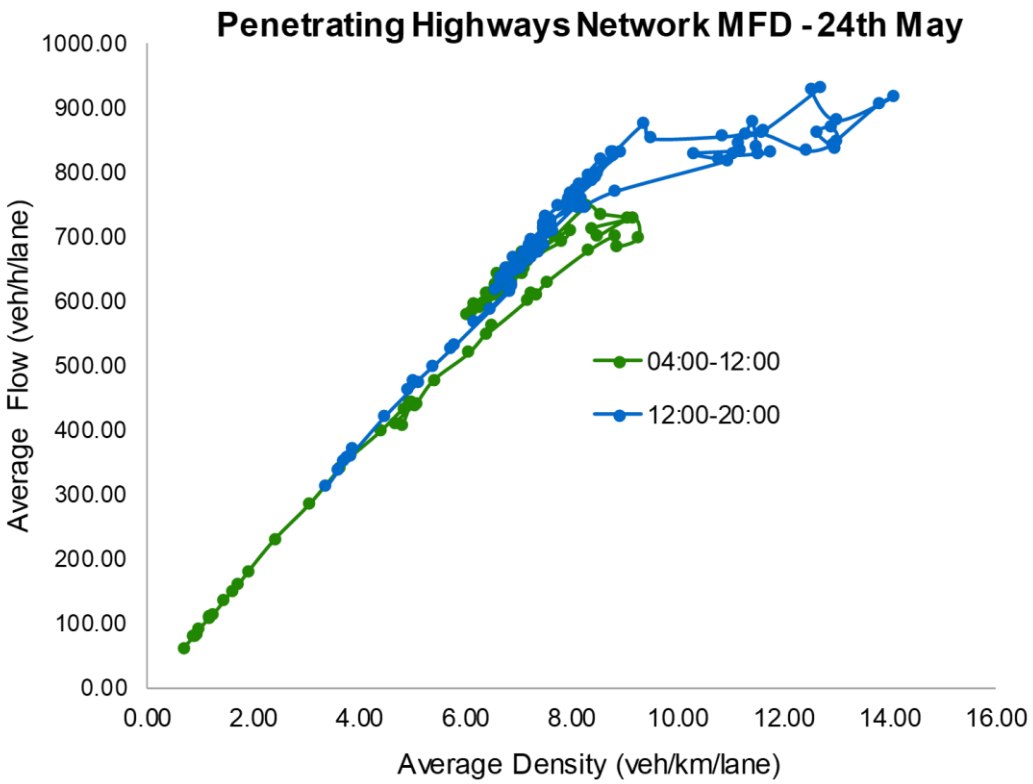


Figure 5-14: Penetrating Freeways Network MFD for the 31st May 2019

5.7.2. Shapes of Hysteresis Loop

As discussed in sub-section 2.3.3 (b) and on the basis of empirical observations, two types of hysteresis are characterized as type H1 and H2 and shown in Figure 2-11 (Saberi and Mahmassani, 2013). The results here confirm that the hysteresis loops observed in both the 24th and 31st May MFDs are type H2 hysteresis loops. It is observed that the network is recovering with the average network flow remaining roughly unchanged when the average network density decreases as shown in Figure 5-13 9a) and Figure 5-13(b).

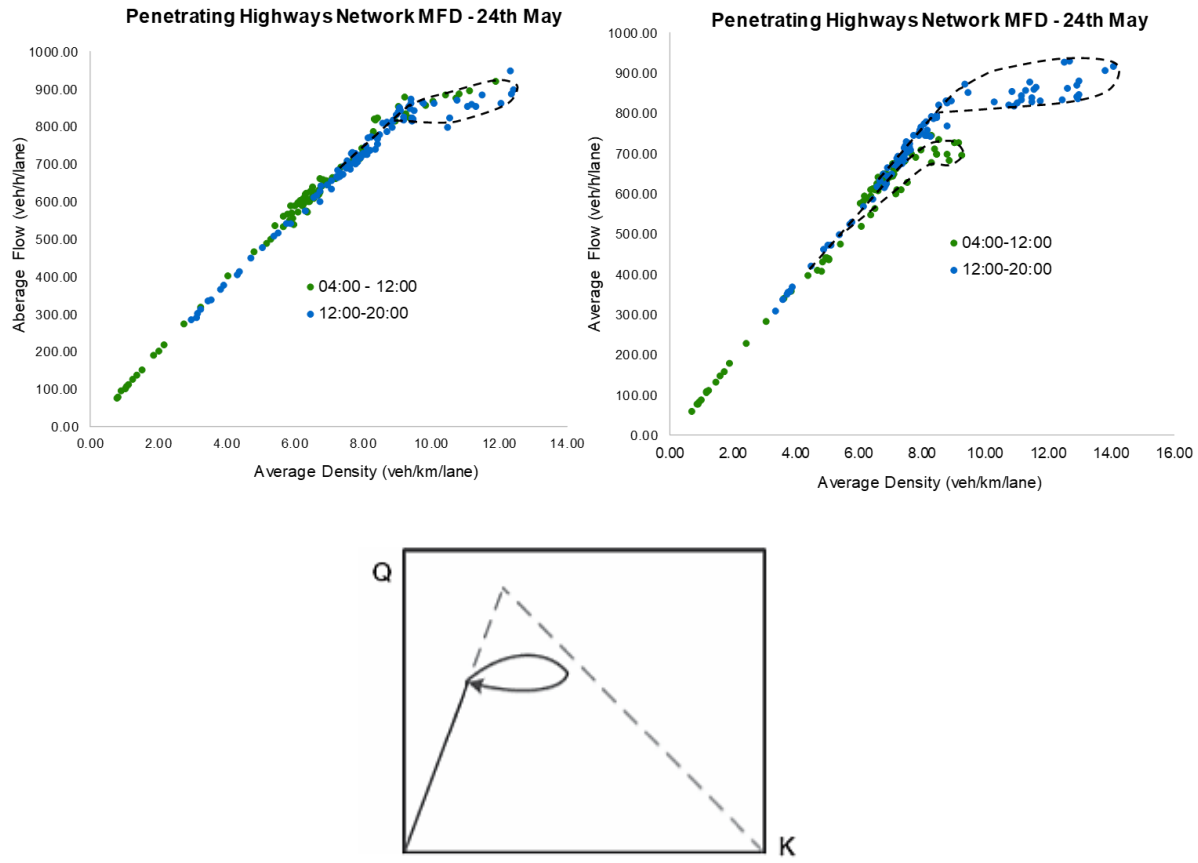


Figure 5-15**:** The penetrating freeways network MFD's hysteresis pattern for the (a) 24th May 2019, (b) 31st May 2019 and (c) schematic representation of type H2 hysteresis loop.

Table 5-1 below provides a summary of the results of the shape of the hysteresis loops observed in the morning and evening of both the 24th and 31st May 2019. Results show that the observed loops are all of the type H2 shape except for the morning (04:00-12:00) of the 24th May 2019 which exhibited no hysteresis loop as shown in Figure 5-15(a).

Table 5-1: Classification of Observed Hysteresis Loops in terms of shape

Network Type	**Date**	**Shape of Hysteresis Loop**	
		Morning	**Evening**
Penetrating Freeways Network	24th May 2019	-	H2
	31st May 2019	H2	H2

5.7.3. The Size of Hysteresis Loops

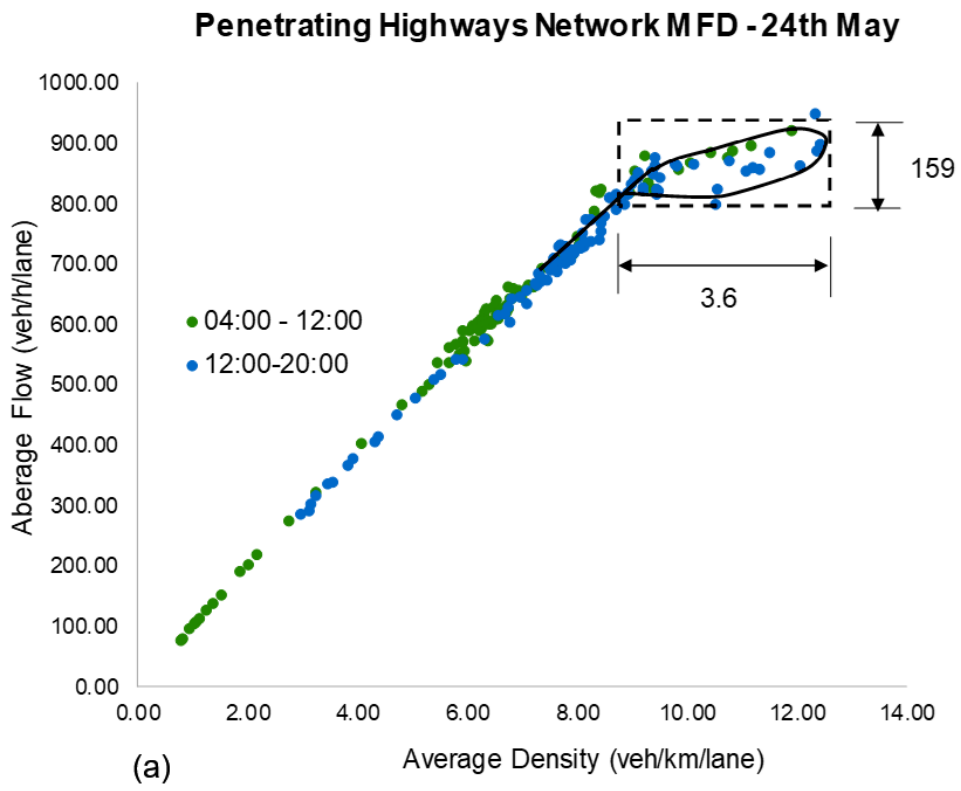
As discussed in sub-section 2.3.3 (c), another key characteristic of the hysteresis loops observed in freeway networks is the size of the loop characterized by its width and height as shown by Equation *(2-17)* and hypothetically represented in Figure 2-12.

Figure 5-17*(a)* and Figure 5-17*(b)* provides an illustration of the size of the hysteresis loops observed in the MFD of the 24th and 31st May 2019, respectively. Table 5-2 summarises the size of the hysteresis loops observed in the morning and evenings of the 24th and 31st May 2019.

Table 5-2: Classification of observed hysteresis loops in terms of size

Network Type	Date	Size of Hysteresis Loop	
		Morning	Evening
Penetrating Freeways Network	24th May 2019	-	(3.6,159)
	31st May 2019	(1.3, 70)	(5.6, 145)

By comparing the size of the loop observed here on the evening of the 24th May 2019 with the size of the loop observed in the MFD for the Portland freeway network on the 25th April 2011, it can be seen that both ΔK and ΔQ are lower here, 3.6 veh/km/lane and 159 veh/hr/lane respectively, whereas it was 7.7 and 178 in Portland. Lower value of ΔQ observed here signifies a more relatively homogeneous and stable recovery of congestion.



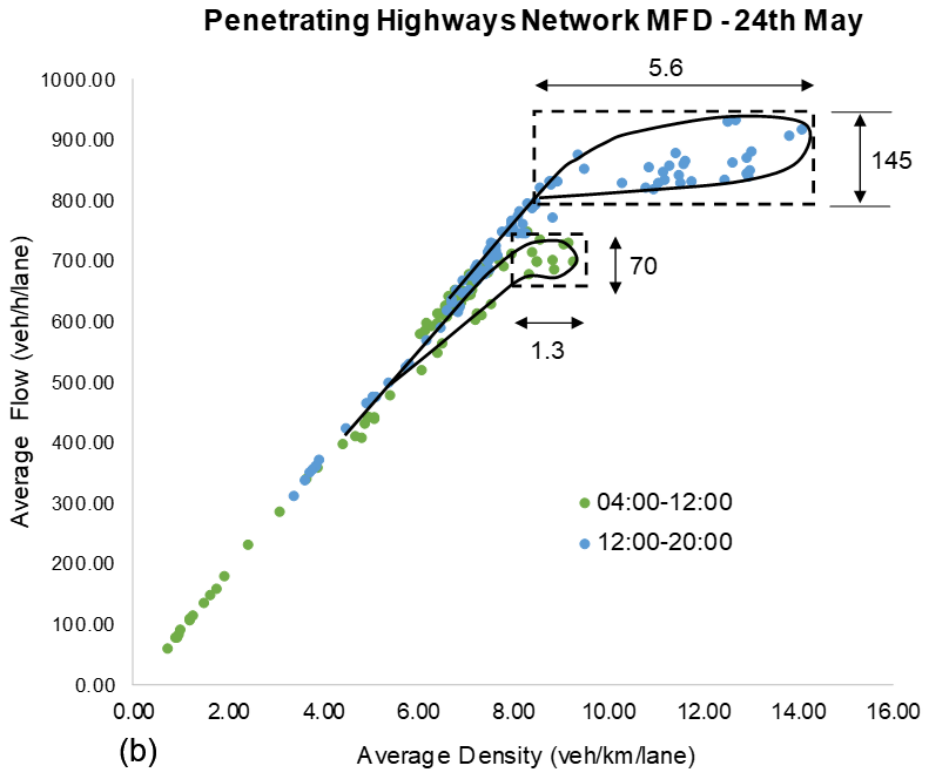


Figure 5-16: Quantification of the size of the hysteresis loop observed in the morning and evening of the 24th and 31st May 2019

5.8. Summary of Analysis & Results

This chapter's aim was to analyse the whole monthly and daily MFDs for the CoCT's highway network and explore the impact of traffic demand and network partitioning on the shape of the MFD. The analyses commenced with the analysis of disaggregated lane data which confirmed that when a high scatter-plot of flow and density from individual loop detectors were aggregated, the scatter faded, and points grouped into a half parabolic shape characterized by the formation of hysteresis loop past the critical density.

This was followed by a comparative analysis of the average flow and density time-series of the four Fridays for the month of May 2019 over an entire freeway network. Results confirmed that both the morning and evening peak flows and densities for the four Fridays were well pronounced and followed similar trend. In addition, it was observed that the evening peaks (1026 veh/hr/lane) were more pronounced than the morning peaks (874 veh/hr/lane).

The whole month MFD for the city's freeway network was derived and analysed. The results confirmed the existence of MFD with clear defined free-flow regime, critical regime and the formation of hysteresis loop and slight scatter (see Figure 5-3). However, in order to better understand the characteristics of the MFD and implications on congestion dynamics, it was important to derive and analyse the daily (Fridays) MFDs.

Daily MFDs for the 10th, 17th, 24th and 31st May 2019, were drawn and analysed. The results confirmed the existence of MFD with well-defined free flow regime, critical regime and the formation of hysteresis loop past the critical density. The results were found to be consistent with results of similar studies involving freeway networks. To simplify the representation of results, the analysis was limited to the 17th and 31st May MFDs.

Further analysis of the 17th and 31st May 2019 MFD showed that; (i) in the morning (4:00 AM – 12:00 PM), a single loop may form, or it may not and (ii) in the evening (12:00 PM – 20:00 PM), a single loop always forms. The MFDs for the two Fridays were explained in conjunction with the time-series of the average flow and density. This further analysis revealed that the coordinates of the plots of the loop on the MFD correspond to the pair measurement in peak hours.

The analysis of the daily MFDs was followed by an exploration and analysis of the impact of traffic demand on the shape of the daily MFD. By looking at the MFDs for the weekends, it was observed that; (i) an explicit loop seldom appears in a Saturday (18th May) MFD and (ii) a single loop form during one day in a Sunday (19th May) MFD. The features observed were further explained in conjunction with the features of their traffic demand.

Lastly, the impact of partitioning the freeway network into two subnetworks of penetrating freeways and ring freeway, on the shape of the MFD was analysed. The results first revealed that the maximum average flow for the two sub-networks were observed to be significantly different from each other i.e. 943 veh/h/lane for penetrating freeways and 1539 veh/hr/lane. The congestion seen on the penetrating freeways was not sufficiently spread on those freeways; the minimum congestion was 697, which was 75% of the maximum observed flow.

Analyzing the penetrating freeway network further revealed that on the 24th May, congestion was observed during both the morning and evening peak periods, whereas on the 31st May congestion was observed mainly during the evening peak period, with hysteresis-like shape. This observation confirmed that the penetrating freeway network was not homogenously congested during the evening peak period of the 31st May 2019. This heterogeneity in the congestion level is probably the cause of the observed hysteresis-like shape.

This ringway freeway was however observed to be exhibiting only the free-flow regime during both morning and evening peak periods, no congestion was observed. These observations show that the overall network was not homogenous congested during peak periods, hence observed formation of hysteresis loops and slight scatters.

The observed hysteresis loops in the MFD of the penetrating freeways network were further characterized according to their shape and size. To explore the effect of varying traffic conditions across various days, a comparison of the penetrating freeway network MFD observed on both the 24th and 31st May 2019 were compared. The results confirmed that the observed MFDs were not well-defined due to the presence of hysteresis phenomenon and slight scatter. In addition, the observed scatter and hysteresis patterns for the two days differed slightly in size and shape.

The observed hysteresis loops were further characterized in terms of their shape. The results confirmed that the hysteresis loops observed in both the 24th and 31st May MFDs are type H2 hysteresis loops, signifying stable recovery of the network during peak periods. It was observed that the network is recovering with the average network flow remaining roughly unchanged when the average network density decreases.

Lastly, the observed hysteresis loops were further characterized in terms of their size (width and height). Comparing the size of the hysteresis loops observed here with the ones observed in MFDs of freeway networks of Portland and Toulouse, it could be seen that the width and size of the hysteresis loops observed in CoCT freeway network were relatively smaller. According to Saberi and Mahmassani (2013), the observed smaller hysteresis loop sizes signified a more relatively homogeneous and stable recovery of congestion during peak periods.

Chapter 6 Conclusions

Finally, this chapter presents a summary of the whole research by first reviewing the aims and objectives of the research as set out earlier in Chapter 1. It then provides the research key findings and concluding remarks. Based on key findings, the chapter concludes with the presentation of recommendations and future research needs.

6.1. Review of Aims and Objectives

The main aim of this research was to establish the existence and implications of the macroscopic fundamental diagram for the City of Cape Town freeway network. To achieve this, empirical data sourced from fixed loop detectors of 5 minutes periods over one month of May 2019 was used to estimate the MFDs for the CoCT's freeway network. The following key research objectives set out earlier in Chapter 1 were achieved.

The whole month MFD for the CoCT's freeway network was derived and analysed. This was followed by the derivation and analysis of the daily MFDs for the four Fridays of the month of May 2019. The study also explored and analysed the impact of changes of traffic demand levels on the shape of the MFD. Lastly, the whole network was partitioned into two sub-networks of penetrating freeway network and ring freeway network, and the impacts on the shape of the daily MFD were analysed.

6.2. Key Findings and Conclusions

The research key findings and conclusions are summarised here in response to the aims and objectives set out in Chapter 1 and reviewed in section 6.1 above.

The whole month MFD for the city's freeway network was derived and analysed. The results confirmed the existence of MFD with clearly defined free-flow regime, critical regime and the formation of hysteresis loop and slight scatter.

However, in order to better understand the characteristics of the MFD and implications on congestion dynamics, it was important to derive and analyse the daily (Fridays) MFDs.

The daily MFDs for the 10th, 17th, 24th and 31st May 2019, were drawn and analysed. In order to achieve this, observed data of 5-minutes periods throughout the month of May 2019 was used to draw the MFD for the CoCT freeway network. The following key findings were observed;

a) a single hysteresis loop always forms past the critical density during the evening peak periods in a weekday (Friday) MFDs.

b) a single hysteresis loop may form or may not form past the critical density during the morning peak periods in weekday (Friday) MFDs.

c) an explicit hysteresis loop seldom appears in a Saturday MFD

d) a single hysteresis loop form during one day in a Sunday MFD.

The study showed that the MFD features could be explained and understood in conjunction with the increases and decrease in average flow and average density time-series patterns. Specifically, the results revealed that the coordinates of the plots of the loop on the MFD correspond to the pair measurement in peak hours.

Partitioning the freeway network into two subnetworks of penetrating freeways and ring freeway showed that;

a) Within a road-type homogeneous data set, separating penetrating freeways from the ring road is a key factor in obtaining low scatter in the MFD.

b) the maximum average flow for the two sub-networks were significantly different from each other.

c) The congestion seen on the penetrating freeways was not sufficiently spread on those freeways.

d) On the 24th May, congestion on the penetrating freeway network was observed during both the morning and evening peak periods, whereas on the 31st May congestion was observed mainly during the evening peak period, with hysteresis-like shape. This heterogeneity in the congestion level is probably the cause of the observed hysteresis-like shape.

e) The ringway freeway was however observed to be exhibiting only the free-flow regime during both morning and evening peak periods, no congestion was observed.

f) These observations show that the overall network was not homogenous congested during peak periods, hence observed formation of hysteresis loops and slight scatters.

Further characterization of the observed hysteresis loops in terms of their shape and size revealed the following dynamics;

a) Penetrating freeways network MFD's observed scatter and hysteresis patterns for across different days vary in size and shape

b) The shape of the hysteresis loops observed during both morning and evening peak periods, are type H2 hysteresis loops, signifying stable recovery of the network during peak periods.

c) The size (width and height) of the observed hysteresis loops were found to be smaller, signifying a relatively homogeneous and stable recovery of congestion during peak periods

The conclusions here are mainly in agreement with one of the "regularity condition" defined in (Daganzo and Geroliminis, 2008): it was established here on an experimental basis that the congestion spreading must be homogeneous to obtain a less scattered and well defined MFD. The use of the MFD to establish control and monitoring strategies at an urban-wide level is feasible provided that the definition of homogeneous zones/sub-networks is made carefully.

The keys findings observed here, particularly regarding the shape and size of the observed hysteresis loops, are consistent with the findings observed in the MFDs of the freeway network the cities of Toulouse (Buisson and Ladier, 2009), Chicago, Portland and Irvine (Saberi and Mahmassani, 2013).

6.2. Recommendations and Future Needs

In addition to meeting the objectives of this research, recommendations for future research are outlined here. First, there is a need for verification and explication of the analysis mechanisms that uses the decrease and increases in average flow and average density time-series patterns to explain and better understand the observed features of the MFD, as shown in this research.

Second, it is recommended that future MFD research studies for the CoCT should consider the data from the arterial network as well. In which case, it will be recommended that efforts should be made to utilize both probe vehicles and loop detector data sources. Studies have shown that a combination of probe vehicle data and loop detector produces the most accurate representation of the MFD.

Lastly, the definition of congestion threshold could be further explored. As overviewed in section 2.1, there are several number of ways of measuring congestion levels for the purpose of traffic management efficiencies. However, there is still no common consensus on how to define congestion threshold, a baseline for measuring congestion levels.

References

Aboudolas, K. and Geroliminis, N. (2013) ‘Perimeter and boundary flow control in multi-reservoir heterogeneous networks’, *Transportation Research Part B: Methodological*, 55, pp. 260–281.

Adams, W. . (1936) ‘Road Traffic Considered as a Random Series’, *Inst. Civil Engineers*, 4(121–130).

Ambühl, L. *et al.* (2016) ‘Empirical Macroscopic Fundamental Diagrams Insights from loop detector and floating car data’, (August). doi: https://doi.org/10.3929/ethz-a-010180262 Rights.

Ampountolas, K. and Kouvelas, A. (2015) ‘Real-Time Estimation Of Critical Values Of The Macroscopic Fundamental Diagram For Maximum Network Throughput’, *American Control Conference*.

Buisson, C. and Ladier, C. (2009) ‘Exploring the impact of homogeneity of traffic measurements on the existence of macroscopic fundamental diagrams’, *Transportation Research Record*.

Cassidy, M. J., Jang, K. and Daganzo, C. F. (2011) ‘Macroscopic fundamental diagrams for freeway networks: Theory and observation’, *Transportation Research Record*.

Courbon, T. and Lecrercq, L. (2011) ‘Cross-comparison of Macroscopic Fundamental Diagram Estimation Methods’, *Procedia - Social and Behavioral Sciences*, 20, pp. 415–430.

Daganzo, C. F. (2005a) ‘A variational formulation of kinematic waves: basic theory and complex boundary conditions’, *Transportation Research Part B: Methodological*, 39, pp. 180–189.

Daganzo, C. F. (2005b) ‘A variational formulation of kinematic waves: Solution methods’, *Transportation Research Part B: Methodological*, 39, pp. 930–955.

Daganzo, C. F. (2007) ‘Urban gridlock: Macroscopic modeling and mitigation approaches’, *Transportation Research Part B: Methodological*.

Daganzo, C. F., Gayah, V. and Gonzales, E. . (2011) 'Macroscopic relations of urban traffic variables: Bifurcations, multivaluedness and instability', *Transportation Research Part B: Methodological*, 45.

Daganzo, C. F. and Geroliminis, N. (2008) 'An analytical approximation for the macroscopic fundamental diagram of urban traffic', *Transportation Research Part B: Methodological*.

Daganzo, C., Gayah, V. and Gonzales, E. . (2011) 'Macroscopic relations of urban traffic variables: Bifurcations, multivaluedness and instability', *Transportation Research Part B: Methodological*, 45, pp. 276–289.

Edie, L. . (1961) 'Car following and steady-state theory for non congested traffic', *Operations Research*, 9, pp. 66–75.

Edie, L. . (1963) 'Discussion of Traffic Stream Measurements and Definitions', in *2nd International Symposium on the Theory of Traffic Flow*.

Federal Freeway Administration (1986) *The 1985 Freeway capacity manual*. Washington, D.C: United States Department of Transportation.

Gartner, N., Messer, C. and Rathi, A. (1992) *Traffic Flow Theory: A State of the Art Report*, *Transportation Research Board*.

Gayah, V. and Daganzo, C. (2011) 'Clockwise hysteresis loops in the Macroscopic Fundamental Diagram', *Transportation Research Part B: Methodological*, 2249.

Gayah, V. and Daganzo, C. . (2011) 'Effects of Turning Maneuvers and Route Choice on a Simple Network', *Transportation Research Record: Journal of the Transportation Research Board*, 2249.

Geroliminis, N. and Daganzo, C. F. (2007) 'Macroscopic Modeling of Traffic in Cities', in. Washington, D.C: 86th Annual Meeting of the Transportation Research Board.

Geroliminis, N. and Daganzo, C. F. (2008) 'Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings', *Transportation Research Part B: Methodological*.

Geroliminis, N., Haddad, J. and Ramezani, M. (2013) 'Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: A model predictive approach', *IEEE Transactions on Intelligent Transportation Systems*, 14(348–359).

Geroliminis, N. and Levinson, D. (2009) ‘Cordon Pricing Consistent with the Physics of Overcrowding’, *Transportation and Traffic Theory 2009: Golden Jubilee*.

Geroliminis, N. and Sun, J. (2011a) ‘Hysteresis phenomena of a Macroscopic Fundamental Diagram in freeway networks’, *Transportation Research Part A: Policy and Practice*.

Geroliminis, N. and Sun, J. (2011b) ‘Properties of a well-defined macroscopic fundamental diagram for urban traffic’, *Transportation Research PartB: Methodological*. Elsevier Ltd, 45(3), pp. 605–617.

Godfrey, J. W. (1969) ‘The mechanism of a road network’, *Traffic Engineering and Control*.

Greenshields, B. . (1935) ‘A Study in Highway Capacity’, *Highway Research Board*, 14, p. 458.

Haddad, J. and Geroliminis, N. (2012) ‘On the stability of traffic perimeter control in two-region urban cities’, *Transportation Research Part B: Methodological*.Haddad, J., Ramezani, M. and Geroliminis, N. (2013) ‘Cooperative traffic control of a mixed network with two urban regions and a freeway’, *Transportation Research Part B: Methodological*, 54, pp. 15–40.

Herman, R. (1992) ‘Technology, Human Interaction, and Complexity: Reflections on Vehicular Traffic Science’, *Operations Research*, 40(2), pp. 199–212.

Herman, R., Malakhoff, L. . and Aardekani, S. . (1988) ‘Trip time-stop time studies of extreme driver behaviors’, *Transportation Research Part A: General*, 22.

Herman, R. and Potts, R. . (1961) ‘Single-lane traffic theory and experiment’, in *Proceedings of Symposium on the Theory of Traffic Flow*. Amsterdam: Elsevier publishing Co, pp. 120–140.

Herman, R. and Prigogine, I. (1979) ‘A two-fluid approach to town traffic’, *Science*.

Ji, Y. *et al.* (2010) ‘Investigating the Shape of the Macroscopic Fundamental Diagram Using Simulation Data’, *Transportation Research Record*, 2161, pp. 35–50.

Ji, Y. and Geroliminis, N. (2012) ‘On the spatial partitioning of urban transportation networks’, *Transportation Research Part B: Methodological*, 46.

Keyvan-Ekbatani, M., Kouvelas, A. and Papamichail, I. (2012) 'Exploiting the fundamental diagram of urban networks for feedback-based gating', *Transportation Research Part B: Methodological*.

Keyvan-Ekbatani, M., Papamichail, I. and Papageorgiou, M. (2013) 'Urban congestion gating control based on reduced operational network fundamental diagrams', *Transportation Research Part C: Emerging Technologies*.

Knoop, V. . and Hoogendoorn, S. . (2011) 'Two-Variable Macroscopic Fundamental Diagrams for Traffic Networks', in *9th Conference on Traffic and Granular Flow*. Moscow, Russia.

Knoop, V. and Hoogendoorn, S. (2013) 'Empirics of a generalized macroscopic fundamental diagram for urban freeways', *Transportation Research Record*, (2391), pp. 133–141.

Knoop, V., Hoogendoorn, S. and Van Lint, J. (2012) 'Routing strategies based on macroscopic fundamental diagram', *Transportation Research Record*.

Laval, A. . (2009) 'Hysteresis in the fundamental diagram: impact of measurement methods', in *89th Annual Meeting of the Transportation Research Board*. Washington D.C.

Laval, A. . (2010) 'Hysteresis in Traffic Flow revisited: An improved measurement method', *Transportation Research, Part B*, 45(2), pp. 385–391.

Leclercq, L., Chiabaut, N. and Trinquier, B. (2014) 'Macroscopic Fundamental Diagrams: A cross-comparison of estimation methods', *Transportation Research Part B: Methodological*.

Leclercq, L. and Geroliminis, N. (2013) 'Estimating MFDs in simple networks with route choice', *Transportation Research Part B: Methodological*.

Lecrercq, L. *et al.* (2015) 'Macroscopic traffic dynamics with heterogeneous route patterns', *Transportation Research Part C: Emerging Technologies*, 59.

Lindley, J. A. (1987) 'Urban freeway congestion: quantification of the problem and effectiveness of potential solutions.', *ITE Journal*.

Lomax, T. J. *et al.* (1997) 'Quantifying Congestion. Volume 1: Final Report', *NCHRP Report*.

Maes, W. (1979) ‘Traffic data collection system for the Belgian motorway network-measures of effectiveness aspects’, *Proceedings of the International Symposium on Traffic Control Systems.2D-Analysis and Evaluation*.

Mahmassani, H. ., Saberi, M. and Zockaie, A. (2013) ‘Urban network gridlock: Theory, characteristics, and dynamics’, *Transportation Research Part C: Emerging Technologies*.

Mahmassani, H., Williams, J. and Herman, R. (1987) ‘Performance of urban traffic networks’, in *Proceedings of the 10th International Symposium on Transportation and Traffic Theory*.

Mazloumian, A., Geroliminis, N. and Helbing, D. (2010) ‘The spatial variability of vehicle densities as determinant of urban network capacity’, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*.

Newell, G. . (1965) ‘Instability in dense highway traffic’, in *Proceedings of the Second International Symposium on the Theory of Traffic Flow*. London, pp. 70–80.

Saberi, M. *et al.* (2014) ‘Estimating Network Fundamental Diagram Using Three-Dimensional Vehicle Trajectories’, *Transportation Research Record: Journal of the Transportation Research Board*.

Saberi, M. and Mahmassani, H. (2012) ‘Exploring Properties of Networkwide Flow-Density Relations in a Freeway Network’, *Transportation Research Record: Journal of the Transportation Research Board*, 2315, pp. 150–165.

Saberi, M. and Mahmassani, H. . (2013) ‘Hysteresis and capacity drop phenomena freeway networks: Empirical characterization and interpretation’, *Transportation Research Record: Journal of the Transportation Research Board*, 2391(5).

Smeed, R. J. (1967) ‘The road capacity of city centers’, *Highway Research Record*.

Thomson, J. M. (1967) ‘Speeds and Flows of Traffic in Central London: 1. Sunday Traffic Survey’, *Traffic Engineering and Control*, 8.

TomTom (2020) *Cape Town traffic report | TomTom Traffic Index*. Available at: https://www.tomtom.com/en_gb/traffic-index/cape-town-traffic/ (Accessed: 12 April 2020).

Treiterer, J. and Myer, J. . (1974) 'The hysteresis phenomenon in traffic flow', in *Proceedings of the Sixth Symposium on Transportation and Traffic Flow Theory*, pp. 213–219.

Tsubota, T. *et al.* (2013) 'Real Time Information Provision Benefit Measured by Macroscopic Fundamental Diagram', in *World Conference on Transport Research (WCTR 2013)*. Rio de Janeiro, Brazil.

Tsubota, T. (2014) *Exploring the properties of macroscopic fundamental diagram: Analysing Brisbane Urban Network*. Queensland.

Venables, A. J. (2007) 'Evaluating Urban Transport Improvements Agglomeration and Income Taxation', *Journal of Transport Economics and Policy*.

WANG, P. fei *et al.* (2015) 'An Empirical Analysis of Macroscopic Fundamental Diagrams for Sendai Road Networks', *Interdisciplinary Information Sciences*, 21(1), pp. 49–61.

Wardrop, J. G. (1952) 'ROAD PAPER. SOME THEORETICAL ASPECTS OF ROAD TRAFFIC RESEARCH.', *Proceedings of the Institution of Civil Engineers*.

Xie, X., Chibaut, N. and Leclercq, L. (2013) 'Macroscopic Fundamental Diagram for Urban Streets and Mixed Traffic: Cross Comparison of Estimation Methods', *Transportation Research Record: Journal of the Transportation Research Board*.

Yoshii, T., Yonezawa, Y. and Kitamura, R. (2010) 'Evaluation of an Area Metering Control Method Using the Macroscopic Fundamental Diagram', in *World Conference on Transport Research*. Lisbon, Portugal.

Zahavi, Y. (1972) 'Traffic Performance Evaluation of Road Networks by the α-Relationship. Parts I & II', *Traffic Engineering & Control*, 14.

Zhang, H. . (1999) 'A mathematical theory of traffic hysteresis', *Transportation Research Part B*, 333, pp. 1–20.

Zhang, H. . and Kim, T. (2005) 'A car-following theory for multiphase vehicular traffic flow', *Transportation Research Part B*, 39, pp. 386–398.

Zheng, N. *et al.* (2012) 'A dynamic cordon pricing scheme combining the Macroscopic Fundamental Diagram and an agent-based traffic model', *Transportation Research Part A: Policy and Practice*, 46, pp. 1291–1303.

Appendices

Appendix 1: Loop Detectors Information

(Attached)

Appendix 2: Processed Input Data

(Attached)