Het verkeerskundig
laboratorium
voor studenten

ITS EDU LAB

The Effect of Network Structure and Signal Settings on the Macroscopic Fundamental Diagram

David de Jong

August 2012

Rijkswaterstaat
Ministerie van Infrastructuur en Milieu

TU Delft

Colophon

Author	David de Jong Delft University of Technology Faculty of Civil Engineering and Geosciences Master Student Transport & Planning Student number: 1153390
E-mail	daviddajong@hotmail.com
Graduation Committee	Prof. dr. ir. S.P. Hoogendoorn *Committee chairman* Delft University of Technology Faculty of Civil Engineering and Geosciences Department of Transport & Planning Dr. V.L. Knoop *1st daily supervisor* Delft University of Technology Faculty of Mechanical, Maritime and Materials Engineering (3mE) Delft Center for Systems and Control Dr. ir. J.W.C. van Lint *2nd daily supervisor* Delft University of Technology Faculty of Civil Engineering and Geosciences Department of Transport & Planning Dr. ir. H. Taale *External Support* Ministry of Infrastructure and the Environment Centre for Transport and Navigation Ir. P.B.L. Wiggenraad *Thesis coordinator* Delft University of Technology Faculty of Civil Engineering and Geosciences Department of Transport & Planning Ir. M. Schreuder *External Support* Ministry of Infrastructure and the Environment Centre for Transport and Navigation Ir. A. Reijneveld *External Support* City of The Hague Traffic Management and Public Lighting

Published by	ITS Edulab, Delft
Information	Dr. ir. H. Taale
E-mail	henk.taale@rws.nl
Date	August 21st, 2012
Status	Final version
Version number	1.0

ITS Edulab is a cooperation between Rijkswaterstaat Centre for Transport and Navigation and Delft University of Technology

TU Delft

Rijkswaterstaat
Ministerie van Infrastructuur en Milieu

Preface

This thesis forms the final milestone of my study at Civil Engineering at the Delft University of Technology. This thesis is a product of ITS Edulab, a cooperation between the Rijkswaterstaat Centre for Transport and Navigation and the Delft University of Technology; a bridge for students between theory and everyday practice.

Apart from being a concept, it is also physical place, where many of my days in the last year have been spent. Luckily it has been a lively place, due to the presence of my fellow ITS Edulab students. Thank you for spending this time with me. I will miss the exchange of ideas, our discussions and of course all the practical jokes.

Many thanks go out to my committee. You have been the signposts along my often troublesome route, and provided me with the insights and the occasional push towards the final destination. In particular I would like to thank Victor Knoop, my daily supervisor. Your help, feedback and perspectives helped me to structure the thought process and maintain focus on the research questions at hand.

I would also like to extend my gratitude to my friends and family, for supporting me over the last year. The most of thanks go out to my parents. Thank you for being my wailing wall and dragging me through my setbacks. Your unrelenting belief that I could finish this, even in the worst of times, made me believe it as well.

Lastly I would like to thank my colleagues. Thank you for giving me the space and time to focus on my study and for picking up the slack in my absence.

David de Jong
Delft, August 2012

"Every housewife knows that a washing machine must not be completely filled if a good result is to be achieved. The same applies to the urban road network"

- Ernst Joos, Deputy Director of Zurich Transport Authority, Zurich (2000) -

Abstract

Recently it has been proposed that the performance of a complete network can be represented graphically using only aggregated data for flow and density. The resulting graph is the macroscopic fundamental diagram (MFD), which relates the average flow and the average density (network density) of a network to each other. The resulting shape is often concave, meaning that the network output could be maximised if a fixed number of vehicles can be maintained in a network.

In order to gain a deeper insight into what factors affect the shape of the MFD, how the MFD is related to the structure of a network and whether or not it is applicable for control strategies, this thesis aims to answer the following research questions:

1. Does a relationship exist between the shape of the macroscopic fundamental diagram and the structure of the underlying network and on which factors does this depend?
2. Is the macroscopic fundamental diagram of the subnetwork and its perimeter affected by different signals settings and does this affect its applicability for control strategies?

The basis of the first question is formed by the assumption that the shape of the MFD not only depends on characteristics of the roads in the network (length, capacity, speed, signal settings) but is also influenced by average travel times in the network. Changing the physical structure of the network, should therefore result in different travel times and a differently shaped MFD.
The second question is based on the assumption that the shape of the MFD changes when signal settings are changed. If this is the case, this could impact the applicability of the MFD for control strategies aiming to maintain the optimal accumulation in a part of the network by changing signal settings (such as perimeter control), as the input for the control strategy, the optimal accumulation, changes by the control strategy itself.

In order to investigate these effects in different networks, a model has been developed, which fully automates the creation of different networks (including controlled intersection with full signal schemes) for a microscopic simulation model. With this model 7 networks, measuring 3x3 km have been created and simulated, in order to obtain multiple MFDs to answer how the network structure affects the shape of the MFD.
In each of these networks one or more subnetworks (neighbourhoods) were created, by adding additional arterials in the network. In order to assess the effect of different signal settings on the MFD of the subnetwork and its perimeter, the traffic flow is restricted by changing the timing of the traffic signals in both the perimeter and the subnetwork.

Regarding the relation between the shape of the MFD and the network structure, it is concluded that the structure of a network in itself does not have a strong influence on the shape of the MFD. Differences between MFDs are not to be caused by topological differences, but by the different characteristics of the links (length, speed, capacity) and the intersections in the underlying network. Substantial and consistent differences are found for controlled and uncontrolled parts of

the network, as signalised parts of the networks often have a maximum performance and optimal accumulation twice that of non-signalised parts.
This study has not been able to quantify how each factor influences the shape of the MFD, in order to construct MFDs based only on knowledge of the network and without the use of simulations. This because the shape of the MFD was found to differ strongly, even for similar networks. Some differences could not be explained by any of the above factors and as such it is concluded that the stochastic and dynamic nature of traffic has a strong influence on the shape of the MFD. Complex interaction of vehicles cause uneven distribution of traffic over the network, lowering the production and performance, increasing the optimal accumulation and is an important cause for scatter in the MFD. Especially the accumulation is highly sensitive to differences in the distribution of traffic over the network, making the optimal accumulation particularly hard to predict.

Regarding the effect of signal settings on the shape of the MFD of a subnetwork and its perimeter and its applicability for control strategies, it is concluded that a strong relation between the MFD of the subnetwork and its perimeter does exist, in which both react in the same way to changes in traffic demand. Changes in the signal settings can have a strong impact on the shape of the MFD of both the subnetwork and its perimeter.
The ratio between the performance of the subnetwork and perimeter is highly consistent and not directly affected by changes in the signal settings. As a result of optimising the signal timings, more traffic can be processed in the network and the performance will eventually increase as well. The optimal accumulation of the perimeter is found to be highly sensitive to changes in the signal settings. As a consequence it is concluded that the MFD is difficult to use for control strategies aiming to adapt signal timings, in order to maintain the optimal accumulation in a part of the network, because these changed signal timings result in a different optimal accumulation.

The shape of the MFD is far from consistent and can be changed by a lot of different factors. How each factor exactly influences the shape is still not fully understood. As such, it is recommended one should be careful in using the MFD for evaluation studies, as input for control strategies, or as a basis for policy-making decisions.

In order to gain more insight in the MFD and the impact of different factors on its shape, the following topics for further research are proposed:

- Quantification of the impact of stochastic network design and dynamic traffic behaviour on the shape of the MFD;
- Obtaining MFDs for highly heterogeneous networks, to investigate if it is likely that a well-defined MFD for such networks are likely to be ever obtained;
- Development of methods capable of accurately removing scatter and faulty traffic states in order to obtain well-defined and usable MFDs;
- Obtaining MFDs for existing urban networks can greatly increase our understanding of the shape of the MFD and can be used to calibrate models producing MFDs;
- Investigation of the short-term impact of changing signal timings on the optimal accumulation, in order to determine the upper bound of the control horizon.

Table of contents

Preface ..... **iii**

Abstract ..... **v**

Table of contents ..... **vii**

Nomenclature ..... **ix**

1 Introduction **1**

1.1 Managing traffic congestion at a macroscopic level 1
1.2 Important terminology 2
1.3 Problem definition and main research objectives 3
1.4 Research questions 7
1.5 Thesis outline 8

2 The macroscopic fundamental diagram **9**

2.1 Basic principles 9
2.2 Some history 10
2.3 Underlying principles and theories 11
2.4 Fields of application 25
2.5 Analysis, criticism and applicability 33
2.6 Conclusion 38

3 Assessing factors impacting the macroscopic fundamental diagram **39**

3.1 Exploring the macroscopic fundamental diagram in-depth 39
3.2 Factors influencing the shape of the MFD 51
3.3 Conclusion 63

4 Research approach **65**

4.1 Research objective 65
4.2 Choosing a test method 66
4.3 General research approach 67
4.4 Networks to be used 68
4.5 Effects to be investigated 70
4.6 A hypothesis for the different effects 71

5 Modelling framework **77**

5.1 General framework 77
5.2 Simulation environment 78
5.3 Network creation model 80
5.4 Simulation algorithms 87
5.5 Data processor 91

6 Simulation setup 95
6.1 Networks 95
6.2 Simulation scenarios 97
6.3 Simulation settings 99

7 Effect of network structure on the macroscopic fundamental diagram 103
7.1 Simulation results 103
7.2 First assessment and general observations 105
7.3 Adding additional arterials 106
7.4 Asymmetric arterial placement 112
7.5 Different traffic loading patterns 114
7.6 Stochasticity in network design 119
7.7 Scalability of differently shaped subnetworks and perimeters 122
7.8 Conclusion 135

8 Effect of signal settings on the macroscopic fundamental diagram 137
8.1 Simulation results 137
8.2 The effect of controlling subnetwork inflow 139
8.3 The effect of controlling subnetwork outflow 146
8.4 Controlling inflow vs. controlling outflow 150
8.5 The effect of signal settings on the accumulation 152
8.6 Conclusion 154

9 Practical implications 157
9.1 Obtaining MFD data in urban networks 157
9.2 Interpretation of the MFD 158
9.3 Fields of application 159
9.4 Implications for using the MFD in practice 160
9.5 Usability of the MFD 163
9.6 Conclusion 164

10 Conclusion and recommendations 165
10.1 Conclusions 165
10.2 Recommendations 168

Bibliography 171

Appendix A: Network creation model
Appendix B: Simulation results
Appendix C: Digital data

Nomenclature

Abbreviations:

A	**A**ccumulation
AFD	**A**rterial **F**undamental **D**iagram
CoV	**C**oefficient **o**f **V**ariation
COM-interface	**Com**munication-interface (Virtual serial port)
DoD	**D**eviation **o**f **D**ensity
DTA	**D**ynamic **T**raffic **A**ssignment
DTM	**D**ynamic **T**raffic **M**anagement
FD	**F**undamental **D**iagram
K	Average networ**k** density
MFD	**M**acroscopic **F**undamental **D**iagram
MPC	**M**odel **P**redictive **C**ontrol
OD	**O**rigin-**D**estination
PD	**P**ro**d**uction
PF	**P**er**f**ormance
QOD	**Q**ueue **O**ver **D**etector
SCATS-L	**S**ydney **C**oordinated **A**daptive **T**raffic **S**ystem – Linking
SCATS-F	**S**ydney **C**oordinated **A**daptive **T**raffic **S**ystem – Free
SOTL	**S**elf-**O**rganizing **T**raffic **L**ights
VT	**V**ariational **T**heory

Units:

cyc	***cyc**les*
h	***h**our*
km	***k**ilo**m**eter*
m	***m**eter*
s	***s**econd*
veh	***veh**icles*

Variables used in literature study (chapter 2):

b	Fraction of time spend in extended red phases	-
C	Cycle time	s
D	Total network length	m
d	Average distance travelled	m
f	Fraction of time stopped in green phase	-
F	Number of full links	-
g	Exit flow	veh/h
$G(n)$	Exit function	veh/h

G	Green time	s
h	Fraction of time moving toward intersection	-
i	Intersection/link	-
k	Density	veh/km
k^w	Weighted average density	veh/km
K	Average network density	veh/km
l	Link length	m
L_T	Length of freeway loop	km
L	Cumulative number of vehicles that exited the system	veh
$\bar{L}$	Average trip length	km
M	Number of links in network	-
n	Number of vehicles	veh
n_{lanes}	Number of lanes	-
N	Number of exit opportunities	-
O	System output / Trip completion rate	veh/h
p	Probability	-
P	Production	veh-km/h
q	Flow	veh/h
q_B	Bottleneck capacity	veh/h
q_r	Average passing rate	veh/h
q^u	Unweighted average flow	veh/h
q^w	Weighted average flow	veh/h
Q	Total flow	veh/h
$\hat{Q}$	Maximum flow	veh/h
r	Maximum passing rate	veh/s
R	Red time	s
s	Saturation flow	veh/h
S	Standard deviation of number of vehicles	veh/km
t	Time	s
T^0	Travel time under free flow conditions	
$T_i(k_i)$	Average local level (link) upper bound for flow	veh/h
$T(k)$	Average global level (network) upper bound for flow (MFD function)	veh/h
T_{cyc}	Cycle time	s
T_{los}	Lost service time	s
u	Speed	km/h
u_f	Free flow speed	km/h
V^0	Average speed under free flow conditions	km/h
V^{av}	Average vehicle speed	km/h

w	Backward traveling speed	km/h
x	Vehicle location	m
y	Vehicle destination	m
δ	Signal offset	-
μ	Average speed of moving observer	km/h
$\lambda(x)$	Density distribution of exit opportunities	-
π	Probability that an opportunity satisfies a driver's need	-
κ	Jam density	veh/km

Variables used besides literature study:

acc	Maximum acceleration	m^2/s
A	Accumulation	veh
B	Set containing path costs	-
C	Capacity	veh/h
d	Distance (with dash between subscripts: distance between)	m
D	Set containing path travel times	-
dec	Average deceleration	m^2/s
DG	Duality Gap	-
f	Frequency	1/h
k	Density	veh/km
K	Average network density	veh/km
l	Length	km
M	Costs	h
n	Number	-
N	Maximum number of vehicles arriving during red phase	-
OD_{fact}	OD-matrix scaling factor	-
OD_{incr}	Increment factor used to scale OD-factor to obtain gridlock point	-
OD_{prec}	Stop criterion for precision of OD increment factor	-
p	Probability	-
P	Percentage	%
PD	Production	veh-km/h
PF	Performance	veh/h
q	Flow	veh/h
q_I	Slop parameter	veh/h
R	Route set	-
s	Spacing	
S	Deviation of density	veh/km
t	Time	s

t_{OD}	Period in which vehicles are loaded onto the network	s
tt	Travel time	h
T	Period / Lap time	s
u	Speed	km/h
U	Average network speed	km/h
UT	Utility	1/h
w	Width	m
X	X-coordinate	m
Y	Y-coordinate	m
θ	Angle	-
α_{BPR}	Scaling factor for BPR function	-
α_{det}	Detour factor	-
α_{SLOP}	Slop parameter	-
β_{BPR}	Scaling factor for BPR function	-
β_{SLOP}	Slop parameter	-
β_{adj}	Saturation flow adjustment factor	-
γ_b	Distribution ratio of area types	-
γ_{demand}	Fraction of total traffic demand to be simulated	-
δ_{red}	Reduction factor for trips not related to investigated subnetwork	-
σ	Standard deviation	-
ρ_{gl}	Gridlock threshold (percentage of non-moving vehicle)	%
ρ_f	Free flow threshold (percentage of non-moving vehicle)	%
φ	Radius	m
ν	Kinematic viscosity	km^2/h
μ	Sensitivity parameter	-
Ω	Average network backward wave speed	km/h

Subscripts

arr	Arriving	*h*	2nd turn in conflict	*right*	Right turn
art	Arterial	*high*	Highly placed	*runs*	Simulation runs
aux	Auxiliary value	*i*	Link	*s*	Destination
A	Area	*in*	Inflow	*sat*	Saturation
b	Type of area	*ind*	indentation	*safe*	Safety
cf	Conflict group	*init*	Initial	*seed*	Random seeds
cfz	Conflict zone	*itr*	Iteration	*sig*	Signal
c	Capacity	*j*	Jam	*sim*	Simulation
clear	Clearance	*int*	Intersection	*start*	Start
crit	Critical	*isl*	Island	*stop*	Stop bar
ctrl	Controlled outflow rate	*j*	Route	*sn*	Subnetwork
cyc	Cycles	*lanes*	Lanes	*sup*	Supplement
data	Interval for obtaining data	*left*	Left turn	*td*	Travel distance
DUE	Deterministic User Equilibrium	*lost*	Internal lost	*th*	Threshold
edge	Part of network in VISSIM	*low*	Lowly placed	*thr*	Through turn
entry	Entry	*max*	Maximum	*tn*	Turn
ett	Expected travel time	*min*	Minimum	*trips*	Trips
exit	Exit	*pt*	path	*upd*	Update
f	Free flow	*out*	Outflow	*veh*	Vehicles
feed	Feeder	*tot*	Total	*VIS*	VISSIM
fin	Financial costs	*node*	Node	*w*	Width
fn	New feeder	*ovl*	Overlap	*x*	x-direction
fo	Original feeder	*port*	Portal	*y*	y-direction
fw	Freeway	*prec*	Stop criterion	*wait*	Waiting
grid	Grid cell	*r*	Origin	*yellow*	Yellow phase
green	Green phase	*red*	Red phase	*Z*	Network
g	1st turn in conflict	*rest*	Rest	Δt	Time interval

1 Introduction

As welfare increases in a region or a country, it is common for the mobility of its inhabitants to increase as well, causing a rise in vehicle ownership. At increasing welfare cities are growing and the number of activities employed in these cities becomes ever larger, resulting in more urban traffic and harder to reach city centres.
To alleviate these newfound traffic problems, improvements in infrastructure and traffic control are made, such as extra roads, additional lanes and traffic lights. Although urban traffic flow improves by these measures, it is mostly just a matter of time before new traffic problems arise, as a result of the ever increasing number of vehicles.
As the space for construction of new infrastructure becomes more and more limited, it is the traffic that needs to be managed in order to increase city mobility. Using network-wide traffic management strategies, traffic can be distributed more evenly over the network or over time, generally increasing the overall network performance.

1.1 Managing traffic congestion at a macroscopic level

It has recently been proposed that the performance of a complete network, or a part thereof (subnetwork) can be represented graphically using aggregated data for flow and density. The resulting graph is the so-called macroscopic fundamental diagram, or MFD for short.
In this diagram aggregated data for the flow and densities are used, creating a relation between the total vehicle distance travelled in the network per unit of time (production) and the number of vehicles (accumulation) in a network[1]. In theory the resulting relation is often concave, which means that for any network there is an accumulation at which the production is at a maximum. When this point is transgressed, spill-back occurs and finally the network gets into a state of gridlock.[2] A schematised version of a macroscopic fundamental diagram is shown in Figure 1.1.

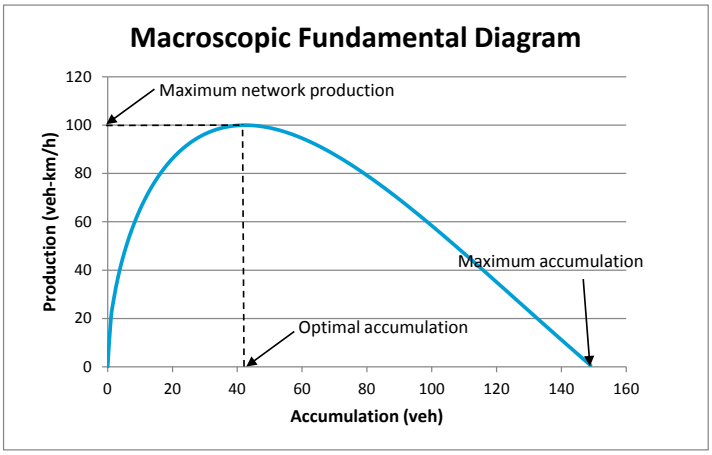


Figure 1.1: *Schematised version of a macroscopic fundamental diagram*

[1] Instead of the production and accumulation, the average flow (performance) respectively the average density (network density) or commonly used as well.
[2] Practically speaking this is not always necessarily true, as the possibility exists that certain networks do not reach this point of maximum production.

It is than most evident that when one wants to optimise traffic flow in a network, one should strive to keep the amount of vehicles in the network, or any subnetwork thereof at this optimal level. To this end, the principal of perimeter control has been proposed, which aims to regulate the inflow and outflow of vehicles at the boundary of a (sub)network in such a way that the number of vehicles within the (sub)network is at or under this optimum. One of the best ways to achieve this, is to use the available traffic signals and prolong the red times for inbound traffic and increase green time for outbound traffic, or vice-versa, depending on the actual traffic state in the (sub)network, as well as on the perimeter.

1.2 Important terminology

Within this report different terms are used for network elements and units of measurement for the MFD. Below the most important terms regarding the network and the MFD are given.

1.2.1 Network elements

The network shown in Figure 1.2 contains two subnetworks consisting of streets with a heterogeneous pattern. Around these subnetwork a road network of a higher order is present, connecting the subnetwork to the rest of the network, the complete set of these roads are called the arterial. For each of the subnetworks, a part of the arterial forms a perimeter around the subnetwork and as such is called the 'subnetwork perimeter', or 'perimeter' for short. Access from the subnetwork to the perimeter and vice-versa is governed by traffic signals, which are placed at intersections connecting the subnetwork to the perimeter.

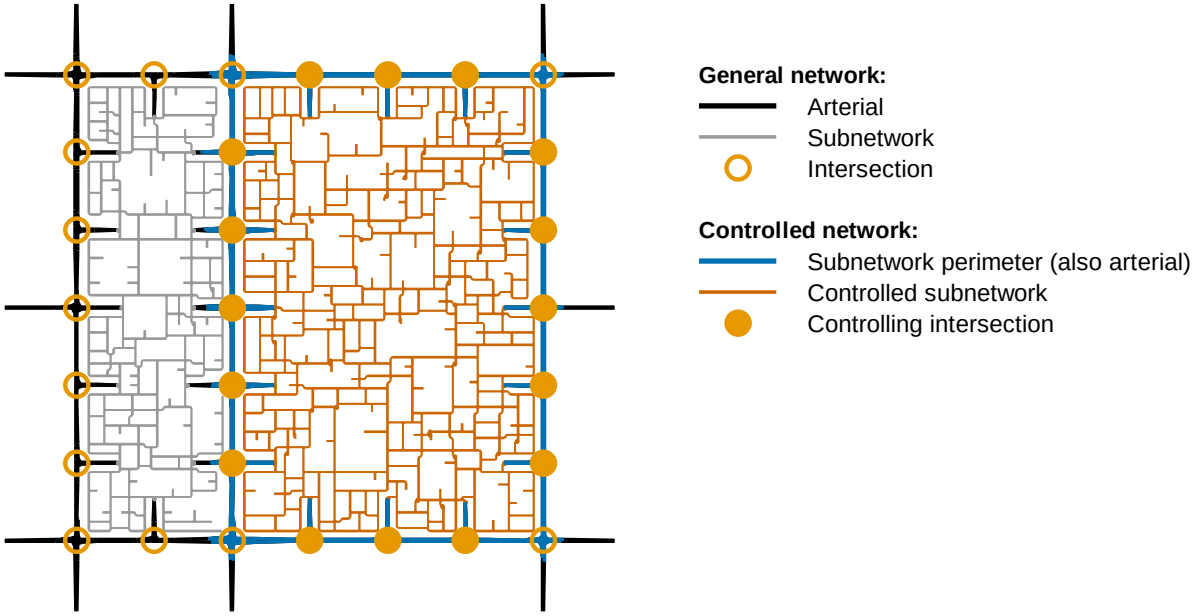


Figure 1.2: *Example of a network, showing its different elements*

1.2.2 MFD parameters

When diving into the world of the MFD, one encounters the terms 'production' and 'performance' quite often. These units are often used on the y-axis of the diagram and express the quality of a network and the way in which it is controlled. These values are related to the 'accumulation', or the 'network density'. In this report too, these terms are often used and as such, it is important to point out the difference between each of them, all the more because these terms are not always used consistently throughout literature. Below, a general description and the unit of measurement of these terms is given, as they are used within this thesis.

- **Production (*veh-km/h*)**: Distance travelled by all vehicles in the network over a unit of time. This value is obtained by multiplying flow measurements over the length of the representative part of the link or network. The production can also be obtained by tracking vehicles throughout the network.
  When calculating representative link or network lengths, the length of each link is multiplied by its number of lanes. This means that network length is equal to the amount of lane-kilometres in the network.
- **Performance (*veh/h*)**: Average flow in the network.
  This value is obtained by dividing the production by the total network length (also called *weighted average flow*), or simply by averaging all flow measurements (also called *unweighted average flow*).[3]
- **Accumulation (*veh*)**: Number of vehicles in the network.
  This value is obtained by multiplying density measurements (or occupancy) over the representative part of the link of network.
- **Network density (*veh/km*)**: Average density in the network.
  This value obtained by dividing the accumulation by the total network length (also called *weighted average density*), or simply by averaging all density measurements (also called *unweighted average density*).

1.3 Problem definition and main research objectives

1.3.1 The effect of network structure on the shape of the MFD

It is assumed in current literature that the shape of the MFD should be directly related to the physical layout of the network and the way that network is controlled, implying that it is independent of driver characteristics and traffic demand patterns. The production therefore should not be a viewed as a property of traffic, but as a property of the network itself. Using a numerical example, this is illustrated below.

[3] In Dutch the term 'performance' translates into 'prestatie', which is expressed in veh-km/h. It should be pointed out that this unit is actually equal to the unit in which 'production' is expressed.

In Figure 1.3, two networks are portrayed, both having an origin A and a destination B, which are 8 km apart. In both networks, two roads are present connecting A and B. The travel distance in the first network is 15 km and in the second network it is 10 km. At a speed of 50 km/h, the travel time is 18, respectively 12 minutes. The networks are assumed to be fully loaded and in free flow, with 2.000 vehicles entering and exiting each network every hour, divided equally over both links. Now, as 2.000 vehicles each have travelled 15 km through network 1, the total production of network 1 is 2.000 · 15 = 30.000 veh-km/h. The production of network 2 is only 2.000 · 10 = 20.000 veh-km/h, which is 50% lower than network 1. At first glance this seems to imply that network 2 performs worse than network 1. However, the travel time in network 2 is 50% less than network 1, meaning that the traffic performance of network 2 is better.

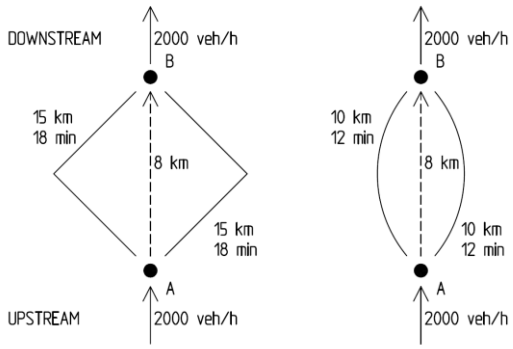


Figure 1.3: *Network production ≠ Network quality*[4]

The example illustrates that the network production does not simply measure how well and how fast traffic is processed, but is a combination of how well traffic is processed and how much traffic can be stored in the network. Although in this case network 1 is the worst performing network (with respect to the travel time), this could change when congestion sets in. For example, if the outflow would become lower than the inflow, spill-back will occur. However, this spill-back will take longer to reach A in network 1, than in network 2, resulting in less congestion upstream of A.
The second thing that can be found from the example is that the network production lowers when the distance between origin and destination is shortened. However, when the average travel distance decreases, the average travel time decreases as well, and as a consequence, the accumulation also decreases (if flow is kept constant), e.g. we see that the accumulation of network 1 is 2.000 · (18 / 60) = 600 vehicles, while the accumulation in network 2 is only 400 vehicles.

From the above it can be concluded that the shape of a MFD that relates production to accumulation is determined on the production side by the amount of infrastructure offered and on the accumulation side by the average travel times. The average travel times in turn depend on the distance that has to be travelled and the average travelling speed. The latter again depends on maximum allowed speed, signal settings and average following distance.
This then implies that when the network structure is changed, i.e. by adding higher-level roads, resulting in shorter travel distance and shorter travel times (also due to higher travel speeds), the shape of the MFD should change as well.

[4] In order to avoid error of speech, the direction of upstream and downstream is given as well

Although the above parameters have been acknowledged to influence the shape of the MFD, no study so far has showed how the shape of the MFD changes when the structure of a network is changed. Apart from that, it also has not been proven too what degree these relations hold, e.g. does the maximum production indeed double if the amount of infrastructure is twice as much and does a linear relationship between the average travel time and accumulation exist.

1.3.2 The effect of signal settings on the subnetwork and perimeter MFD

As is stated earlier, the settings of traffic signals can be adapted, in order to change the flow between the subnetwork and the perimeter. This strategy is called 'perimeter control'. Within such a strategy, the aim is to adapt the signals in such a way that the production is maximised, which in turn is realised by keeping the number of vehicles in the subnetwork and/or the perimeter at or near the optimal accumulation, which is derived from the macroscopic fundamental diagram. Finding this optimal accumulation however is not fully straight-forward.

Consider the network shown in Figure 1.2. If for example every controlling signal only releases a single vehicle per hour from the subnetwork to the perimeter, the subnetwork will fill up quite rapidly. As the subnetwork becomes more congested, the total production rapidly decreases, because the number of vehicle-kilometres that can be travelled per unit of time becomes ever less. As the subnetwork can only remain in free flow when the number of vehicles in the network is equal to or less than the number of vehicles that can leave the subnetwork per unit of time (16 in this case), the optimal accumulation should theoretically be equal to this number of vehicles, as any more vehicles would automatically mean that the subnetwork would fill up.
If however, the number of vehicles that is 'released' to the perimeter is equal to the number of vehicles within the subnetwork, none of the links should become congested and the total production should always be at a higher level than in the first case. As a consequence, the optimal accumulation should also be much higher. As with the previous case, this should theoretically become equal to the total number of vehicles released by the signals each hour.

Now consider the above example again, but looking from the perimeter instead. In the first case (only a single vehicle is released to the perimeter each hour), the perimeter is fully flowing and is almost unaffected by traffic originating from the subnetwork. However, as time increases and the number of vehicles in the subnetwork grows, it becomes ever harder for vehicles from the perimeter to enter the subnetwork, resulting in spill-back and finally causing the perimeter to get in a state of gridlock.
In the second case vehicles can easily flow out of the subnetwork, creating space for vehicles from the perimeter, keeping the perimeter flowing.
If however there would be a very high traffic demand from the subnetwork to the perimeter and all of this traffic would be loaded onto the perimeter, this could also cause congestion on the perimeter, as it becomes overloaded. This in turn limits the number of vehicles that can travel from the subnetwork to the perimeter, again resulting in gridlock of the total

system. This results in a lower overall production and a higher optimal accumulation, due to an increase of the average network density.

The above example illustrates that one should not expect to find a shape for the MFD of a subnetwork and the perimeter that is independent of the way both systems are operated. In fact, it should be expected that the shape of these MFDs are highly dependent on each other, in which the signal settings affect the shape of both the MFDs.
It can also be derived from the above, that depending on the demand pattern, there should be a signal state for which the combined production of the perimeter and subnetwork is maximised.

The key behind the perimeter control strategy is to make changes to the signal timings to improve or reduce the inflow to or outflow from the subnetwork to the perimeter, in order to maintain the number of vehicles in one of these systems (or both) at or near the optimal accumulation.
Changing the signal timings will most probably affect the shape of the MFD of both these systems, as they are assumed to be dependent on one another. Whether one improves while the other deteriorates, or both react in the same way to changes in the signal settings is not known.

As the MFDs change, it should be expected that the optimal accumulation of both systems changes as well, meaning that the signals should actually be set in a different way, as the control target itself has changed.
Using model predictive control (MPC), the effect and the optimal setting of the traffic signals can be determined. This however is a highly time-consuming task, which in most cases takes more time than the control horizon itself.
It is therefore more interesting to know the effect on the MFD of different signal timings before–hand. If the shape of the MFD can be predicted based on the network layout and the way it is controlled, the optimal accumulation can be determined mathematically and new signal settings can be determined without relying on a prediction model.
Apart from the obvious advantage shown above, this relation can also give more insight in the effect of different signal timings on the production of the network. This insight could prove useful in determining more 'static' control tactics for coordinated signal control in a network.

Although some research of the effect of signal timings on the MFD has been done, this effect is never tested using a microscopic model, in which the signals of every intersection are tailored to that intersections specific traffic demand. Instead, macroscopic models with less complex node models have been used to investigate this relation. As the traffic process at intersections is highly complex, this could result in different outcomes.
Furthermore the relation between the MFD for a subnetwork and its perimeter and the way both are affected by changes in signal timings has not been investigated, even though this is important to know, as changes to the signal timings could very well have a significant impact on the shape of the MFD and therefore the way in which the network is to be controlled.

1.4 Research questions

1.4.1 The effect of network structure on the shape of the MFD

As discussed in the previous section, it is currently not clear to what degree a relationship between the structure of a network (subnetwork size, perimeter length, number of intersections, signal settings) and the shape of the MFD exists.
To get a thorough insight in these relations, this thesis aims to answer the following question:

1. *Does a relationship exist between the shape of the macroscopic fundamental diagram and the structure of the underlying network and on which factors does this depend?*

This research question is broken down into the following sub questions/research aspects:

- *1.1. Does a general shape for the macroscopic fundamental diagram exist or is this network specific?*
- *1.2. Which factors influence the shape of the macroscopic fundamental diagram, and is it possible to quantify the effect of each of these factors?*
- *1.3. How is the shape of the macroscopic fundamental diagram affected by changes in the network structure?*

1.4.2 The effect of signal settings on the subnetwork and perimeter MFD

Apart from investigating the relationship between network structure and the MFD, the effect of signal timings on the shape of the MFD of the controlled subnetwork and its perimeter has not been investigated in-depth as well. Understanding this relation however, is important when one wants to utilise the MFD as the main input for a control strategy based on (sub)network perimeter control.
To gain a deeper understanding of these effects, the following research question is defined:

2. *Is the macroscopic fundamental diagram of the subnetwork and its perimeter affected by different signals settings and does this affect its applicability for control strategies?*

This research question is broken down into the following sub questions/research aspects:

- *2.1. Does a relationship between the macroscopic fundamental diagram of the subnetwork and its perimeter exist?*
- *2.2. How are the macroscopic fundamental diagram of the subnetwork and its perimeter affected by changes in the signal settings?*
- *2.3. Can the macroscopic fundamental diagram be used as an input for control strategies?*

1.5 Thesis outline

As a first step towards answering the research questions set forth in the previous section, chapter 2 will start by presenting the research that has been done with respect to the MFD. Goal of this chapter is to gain decent insight in the subject and to investigate what knowledge is available, in order to aid our research questions.
In chapter 3, a thorough investigation of the MFD will be made, to find out how the MFD works, which factors make up the MFD and how the MFD is affected when changes to any of these factors are made. From the insights obtained from this assessment, chapter 4 will present a research approach, capable of giving a substantial answer to the research questions. To this end, a proper test method is chosen and the effects that will be investigated are discussed.

As this thesis will rely on a simulation model for which multiple networks have to be designed, chapter 5 presents a newly developed network creation model and two simulation algorithms to produce the data needed to answer our research questions. Chapter 6 will then discuss how the traffic simulator is set up and what parameters for these models and algorithms are used in the simulation.

After obtaining the results from the simulation, a thorough analysis of the effect that the network structure has on the MFD will be made in chapter 1. In chapter 8 the effects of different signal settings on the MFD are discussed, in order to answer the second research question.

After the analyses have been carried out, chapter 9 will discuss the practical implications of these findings. Apart from that it will also be discussed whether or not the MFD is usable in practice.

The report will be ended with a conclusion and various recommendations in chapter 10.

2 The macroscopic fundamental diagram

In this chapter studies that have been done regarding the macroscopic fundamental diagram are discussed. The aim of this chapter is to investigate to what degree current literature is able to answer the research questions set out in this thesis and which elements are important to take into account when evaluating the effect of different signal timings on the MFD. In section 2.1 the basic principles of the MFD are explained and the information that can be derived from it. In section 2.2 a brief historical overview of how the MFD came to be will be given, after which section 2.3 explains the development of the MFD in more detail, giving an in-depth description of the work done so far and the underlying principles and theories of this diagram. In section 2.4 an outline of the field of application of the MFD are presented.
In section 2.5 a summary and analysis of the literature will be made and the applicability of the literature will be discussed, after which section 2.7 will end this chapter with a conclusion.

2.1 Basic principles

If asked to describe the basic idea of the macroscopic fundamental diagram (MFD) in one sentence, one could state that a MFD shows how well vehicles can travel through a network, depending on the number of vehicles in that network.

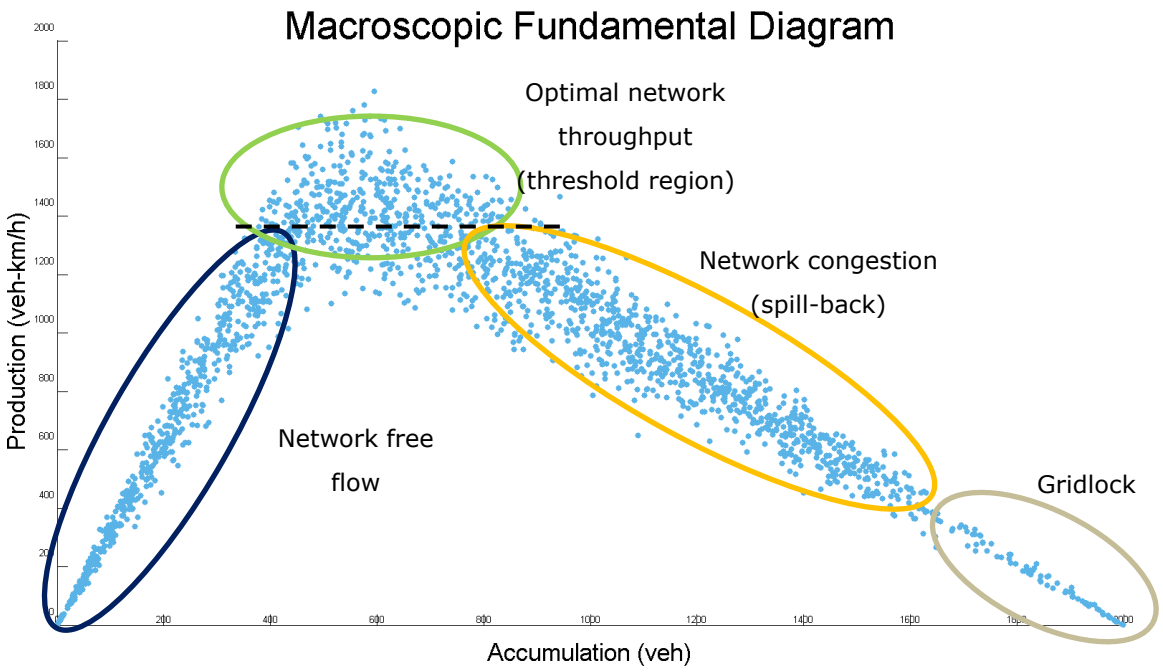


Figure 2.1: *Example of a macroscopic fundamental diagram, indicating different network regimes*

In Figure 2.1 an example of a MFD is given, relating the accumulation to the production of the network. The production is the total flow of all links in the network combined. The accumulation represents the total number of vehicles in that network. Figure 2.1 clearly illustrates the different traffic states in a network and shows a decline in network production after a certain threshold for the accumulation has been reached.
Before this 'threshold region' is reached the network is in free flow conditions, which shows a linear relationship. In this state traffic on almost all links of the network is in free flow and conditions are homogeneous. When increasing the amount of traffic, the network becomes less homogeneously loaded (certain links are congested and others are not) and scatter arises. Also spill-back causes congestion to spread through the network as queues block the outflow of upstream links. When increasing the number of vehicles in the network even further, the network becomes heavily congested, and gets into a state of total gridlock, in which no vehicle is able to move anymore.
In Figure 2.1 the accumulation at which spill-back starts, is chosen as the point where the production is roughly equal to the transition point from network free flow to the 'threshold region', represented by the dashed line. In this region most links of the network are in a congested state.

2.2 Some history

Although numerous variants of the MFD exist today, i.e. using different aggregated data to relate the production to the accumulation, the essence of every MFD does not deviate much from the fundamental diagram (FD). The fundamental diagram describes a relationship between the flow, density and speed on a link and has been around for over 75 years. It has been introduced by Greenshields (1935) and numerous mathematical formulations for its shape have been made throughout the years.

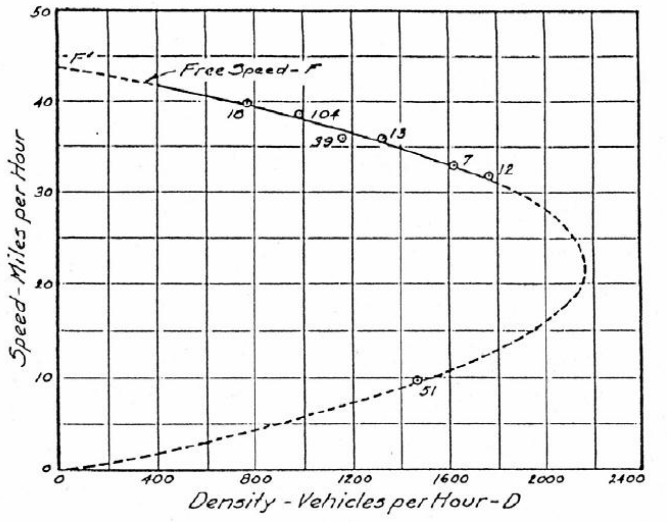


Figure 2.2: *The first Fundamental Diagram as q-v diagram (Greenshields, 1935)*

The first time a macroscopic relationship between average flow and density has been proposed, was in Godfrey (1969). Although a number of researchers have tried to improve on the proposed theory using simulation (Mahmassani, 1987, Mahmassani and Peeta, 1993) or real data (Ardekani

and Herman, 1987, Olszeweski et al., 1995), their data has been found to be to sparse or not investigated deeply enough to demonstrate the existence of an invariant macroscopic relation for real urban networks, as stated in Geroliminis and Sun (2011b).

This subject has been reintroduced recently in Daganzo (2005a, 2007) which used it as a part of an urban traffic dynamics model. In Daganzo (2007) a theoretical relationship between the accumulation and the exit flow (number of vehicles leaving the network) was formed. The resulting relationship was named the exit function, denoted as $G(n)$.
The term 'production' was subsequently introduced in Geroliminis and Daganzo (2007), in which a diagram relating the accumulation to the production was shown for the first time. Although Geroliminis and Daganzo (2008) states that Geroliminis and Daganzo (2007) was the first to demonstrate that a MFD exists, it was actually Geroliminis and Daganzo (2008) which was the first article that coined the term 'macroscopic fundamental diagram'. Interestingly enough, when not counting the title of this article and the abstract as main body of the text, this principle was even introduced first by its now commonly known abbreviation of MFD. Apart from the title and the abstract, the full term was never used in the main article. Moreover, this article was the first to actually state the definition of the MFD as being a relationship between accumulation and production, which it credited as being one of the two postulates of a theory on perimeter control from Geroliminis and Daganzo (2007), stating that *"…(i) that homogeneously congested 'neighbourhoods' exhibit an MFD relating 'production' … and 'accumulation'…"*. These postulates however, were never made in Geroliminis and Daganzo (2007) and the term MFD had not been used either. As such, an actual proposition and definition for the term of 'macroscopic fundamental diagram' has still never been made, as all articles refer to articles that never explicitly defined what an MFD is.

After different articles by Daganzo and Geroliminis, it was since 2009 that the MFD finally seemed to stick in the scientific world, and spurred researchers into investigating this relationship in more detail (as can be seen by the number of articles related to the MFD, produced since then). It should however be noted that most of the published work so far is still (co-)authored by Daganzo or Geroliminis.

2.3 Underlying principles and theories[5]

2.3.1 Input data and axes properties

As stated earlier the MFD is a relationship between the so called 'production', and 'accumulation'. Throughout the different studies done so far, there are multiple ways in which this relationship can be shown. They do however, still portray the same relation and shapes can roughly be compared, even if the axes properties are different.

[5] People who are familiar with the research done in respect to the macroscopic fundamental diagram can skip section 2.3 and 2.4, as this mainly gives an overview and summary of the different studies done.

On the y-axis, almost all studies use average flow, as either the average number of vehicles per second (veh/s) or per hour (veh/hr) and is also often referred to as 'Network flow' or 'Performance'. These values can be either weighted or unweighted. The weighted average is $q^w = \sum_i q_i l_i / \sum_i l_i$ and the unweighted average is $q^u = \sum_i q_i / \sum_i 1$. When all links have the same length, both of these values are equal. In more complex (real) networks, lengths of roads, and or detector spacing can vary, making the weighted average a more accurate value. Determining the network length associated with each detector point can be a tedious task in real networks and as such the unweighted average is easier to use. The term 'Production' is the total number of kilometres travelled within a certain period of time (veh-km/h).

On the 'Accumulation' axis a wide variety of notations is used. They include among other: veh, veh/m, veh/lane-km, detector occupancy (%) and veh-s. Like the production, these values can be either weighted or unweighted.

A more elaborate discussion on the different data and parameters involved in creating the MFD is given in section 3.1.

2.3.2 Genesis: Daganzo's theory

In Daganzo (2007), the author argues that in theory, the most current models can predict almost anything on a multi-modal transport network in minute detail, but not in practice.
The most important reasons for this in his opinion are:

1. Dynamic models require too many inputs, such as dynamic OD-matrices;
2. Driver navigation is an unpredictable gaming activity;
3. Oversaturated networks behave chaotically.

Daganzo argues that by modelling city traffic at an aggregate level problem 3 could be alleviated. The model presented in his paper should alleviate the other two.

As stated earlier, the macroscopic relation studied by Daganzo (2007) was called an exit function. Using the example of a homogeneous looping road, in which (i) only endogenous traffic is present, (ii) origin flows are uniformly distributed along the link and have priority, (iii) average trip length is equal for all origins and (iv) the flows and densities that can be steadily sustained on the road are related by a unimodal fundamental diagram, it should follow that steady states with local density and internal flow should be uniform everywhere.
As the total distance travelled per unit time can be expressed as (i) the product of exit flow g and average distance travelled d, and (ii) as the sum of the distance travelled by the vehicles in the system at any given time $nu = lku = lq = lQ(n/l)$, it should hold that $gd = lQ(n/l) \rightarrow g = (l/d)Q(n/l) \equiv G(n)$. This shows that exit flows are a fixed multiple of the circulating flow. Under the condition that input changes slowly, the exit function can roughly predict the output in the dynamic case and we can write

$$dL(t)/dt \equiv g(t) \approx G\big(n(t)\big). \tag{2.1}$$

After this the theory was extended to inhomogeneous conditions. Again the loop sample is used, but (i) has a variable number of lanes and (ii) input demands are different.
Using the intervening opportunities model as set out in Schneider (1959), which states that trip making is not related to the distance between origin and destination (as in the gravity model), but to the relative accessibility of opportunities of satisfying the objective of the trip, the following was proposed:

> Assume that drivers look for "opportunities" at the various exits along the loop and take the first exit that satisfies their need. These opportunities are distributed along the road with density $\lambda(x)$ and all opportunities are equally likely to satisfy a driver's need (with probability $\pi \ll 1$) independently of where that driver comes from. The total number of opportunities N is so large that $N\pi \gg 1$.
> Let $N(x)$ be the cumulative number of opportunities along the road from a reference point $x = 0$. Then the probability $p(x,y)$ that a driver finds its opportunity before reaching y, given that it did not find it before x, is a function of the intervening opportunities $N(x,y)$, between x and y. Clearly, $N(x,y) = N(y) - N(x)$ if $y > x$ and $N(x,y) = N(y) - N(x) + N$ if $y < x$. The probability formula is $p(x,y) = 1 - (1-\pi)^{N(x,y)} \approx 1 - exp\big(-\pi N(x,y)\big)$. If (x,y) is the segment containing all opportunities associated with a certain exit, $p(x,y)$ will approximate the fraction of vehicles taking the exit.

Based on this theory and the assumption that opportunities are equally likely to satisfy a driver's need independently where he or she comes from, it should hold that the probability of a driver taking a certain exit is independent of its origin and can thus be viewed as a property of the road. The fraction of vehicles taking an exit should then be a fixed fraction of circulating traffic. The grand total of vehicles exiting during a given interval is

$$g = \int_0^{L_t} Q(k(x),x)p(x)dx. \qquad (2.2)$$

This alleviates problems 1 and 2 as this makes the aggregated traffic flow invariant of origin and individual driver antics are removed.
Although the principle of the MFD was not defined in this paper, it was this theory that provided the base for it, as it stated that a fixed relation between circulating traffic and the number of exiting vehicles (production) should theoretically exist.

2.3.3 The Yokohama experiment: Empirically providing evidence for the existence of the MFD

Using detector data and taxi-paths, from a field experiment in Yokohama (Japan), Geroliminis and Daganzo (2008) showed that a relationship between the average density and flow indeed seems to exist. When using data from different time periods and days, which suggested different origin-destination tables, the shape of the MFD did not seem to be affected. This in turn leads to believe that as conjectured earlier, the MFD is indeed invariant of origin and destinations.

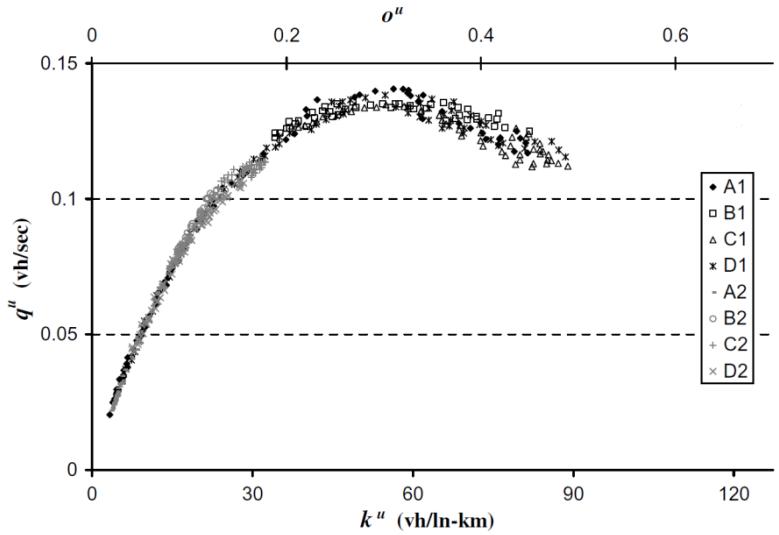


Figure 2.3: *MFD from the Yokohama experiment, over 8 different time periods (Geroliminis and Daganzo, 2008)*

Using data from taxi's fitted with GPS and a data logger, a solid relation between the occupancy of the detectors at the outbound roads and the number of vehicles exiting the network (trip completion rate) was formed, as an almost linear relationship between the two was present and remaining scatter could be properly explained and fell within the error margin of the experiment. Using this relationship, they went on to show that the total network production and trip completion rate were linearly related.

2.3.4 Introducing Variational Theory: Theoretically providing evidence for the existence of the MFD

Based on the work done in Daganzo (2007) and the encouraging results from the Yokohama experiment in Geroliminis and Daganzo (2008) a theoretical approach to the MFDs shape was taken in Daganzo and Geroliminis (2008).
Using the tenets of Variational Theory (VT) as described in Daganzo (2005b, 2005c), it was shown in Daganzo and Geroliminis (2008) that a number of theoretical 'cuts' could be defined, which define the outer boundaries of the MFD.

In VT, a capacity function gives the flow on a homogeneous portion of a street, based on the maximum rate at which vehicles can pass an observer moving with any given speed u. In this case, 3 families of 'cuts' are proposed.
The first family is the 'stationary cut', which is the bottleneck capacity at the most constraining intersection and

$$q = q_B = min_i\{s_i G_i / C_i\}. \quad (2.3)$$

The second family is the 'forward cut' with observers moving at free flow speed u_f. These observers always depart at the end of a red phase. Assuming that all red phases R_i have been

extended at the front end by an amount εG_i, with $\varepsilon \in [0,1]$. If $\mu(\varepsilon)$ is the average speed of this observer and $f_i(\varepsilon)$ is the fraction of time it spends stopped in green phases because of extended red periods, then traffic can pass at a rate of $q_r \leq \sum_i s_i f_i(\varepsilon)$. It then holds that

$$q \leq ku(\varepsilon) + \sum_i s_i f_i(\varepsilon). \tag{2.4}$$

The third family is the 'backward cut' with observers travelling in the opposite direction at speed w and also stopping for red phases. With $w(\varepsilon) > 0$, $b_i(\varepsilon)$ for the fraction of time spend in extended red phases and $h_i(\varepsilon)$ for the fraction of time the observer spends moving towards the intersection. With r_i as the maximum passing rate when moving, the observer can be passed at most at an average rate of $\sum_i [s_i b_i(\varepsilon) + r_i h_i(\varepsilon)]$. The resulting cuts are then

$$q \leq -kw(\varepsilon) + \sum_i [s_i b_i(\varepsilon) + r_i h_i(\varepsilon)]. \tag{2.5}$$

Based on the number of intersections an observer crosses, before having to wait for a red phase, different 'cuts' can be defined. An example of this is given in Figure 2.4, where the data from the Yokohama was used to estimate the different parameters. These results do not completely match the real MFD, with the stationary cut being off the most. The results are still quite encouraging though.

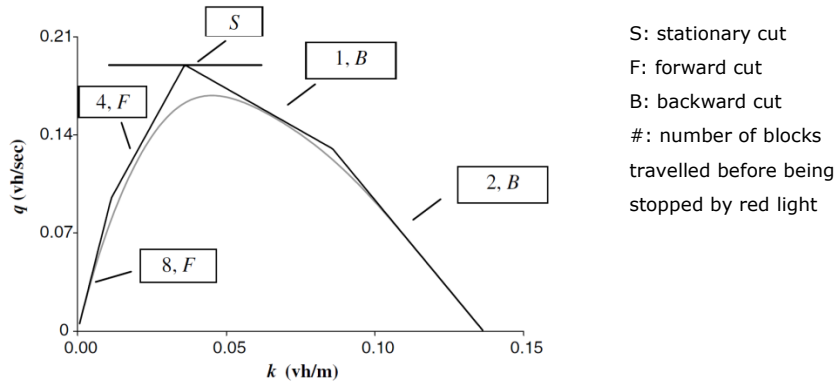


Figure 2.4: *Theoretical MFD of Yokohama with and without stochastic variations (Daganzo and Geroliminis, 2008)*

Conditional on the average network density, the upper bound for the average network flow can now be defined. These averages are defined in the sense of Edie (1963), as $q = \sum_i q_i l_i / D$ and $k = \sum_i k_i l_i / D$, in which D is the total network length. With T_i as the upper bound of possible flow-density states on a link it follows that $T_i(k_i) \geq q_i$. The average network flow then becomes $\sum_i T_i(k_i) l_i / D \geq q$ and the upper bound for the MFD can be written as

$$T(k) = max\{\sum_i T_i(k_i) l_i / D \mid \sum_i k_i l_i / D = k; 0 \leq k_i \leq \kappa_i\} \geq q. \tag{2.6}$$

Based on (2.6) it was conjectured in Daganzo and Geroliminis (2008) that a well-defined relationship with low scatter between production and accumulation should arise if the following 'regularity conditions' are met:

1. A slow-varying and distributed demand;
2. A redundant network ensuring that drivers have many route choices and that most links are on many desirable routes;
3. A homogeneous network with similar links;
4. Links with an approximate FD that is not significantly affected by turning movements when flow is steady.[6]

In turn this should result in a MFD that:

- Accurately describes the average flow in a network depending on the number of vehicles in that network;
- Is invariant of driver's origins and destinations;
- Is invariant of driver's individual route choice.

2.3.5 Toulouse and Amsterdam: The MFD catching on in the scientific world

After the MFDs introduction by Daganzo and Geroliminis, Buisson and Ladier (2009) were the first to comment on this new development. Using detector data from the city of Toulouse, the shape of the MFD was investigated for three different weekdays. As their network consisted of different types of urban roads and freeways, i.e. was inhomogeneous, the four regularity conditions were relaxed.

Although not reaching the congested branch of the MFD, their results showed little scatter apart from one hysteresis loop (see Figure 2.5). This hysteresis loop was credited to spatially heterogeneous evolution of congestion, most likely caused by a slow-moving platoon of trucks (strike by truck drivers).

Apart from constructing the MFD, the influence of detector placement on the MFD-shape and scatter was investigated. Dividing the detector data between urban roads and freeways and distance to traffic signals, they found that the location of the detectors has a strong impact on the slope of the MFD. They also found that freeways did not exhibit the trapezoidal shape, as suggested in Daganzo (2007).

Buisson and Ladier conclude that their findings are mainly in agreement with the regularity conditions set out in Daganzo and Geroliminis (2008). They advise to (i) construct the MFD with loops having their distance from the traffic signal within a short range of variation and (ii) to divide a city into zones on a geographic basis and taking the type of road into account.

[6] This condition should apply if links are sufficiently long because turns only affect the relationship between flow an density locally (mostly at links' ends) and not so much in their middles. It should also apply if one knows a priori that the fraction of turns is small.

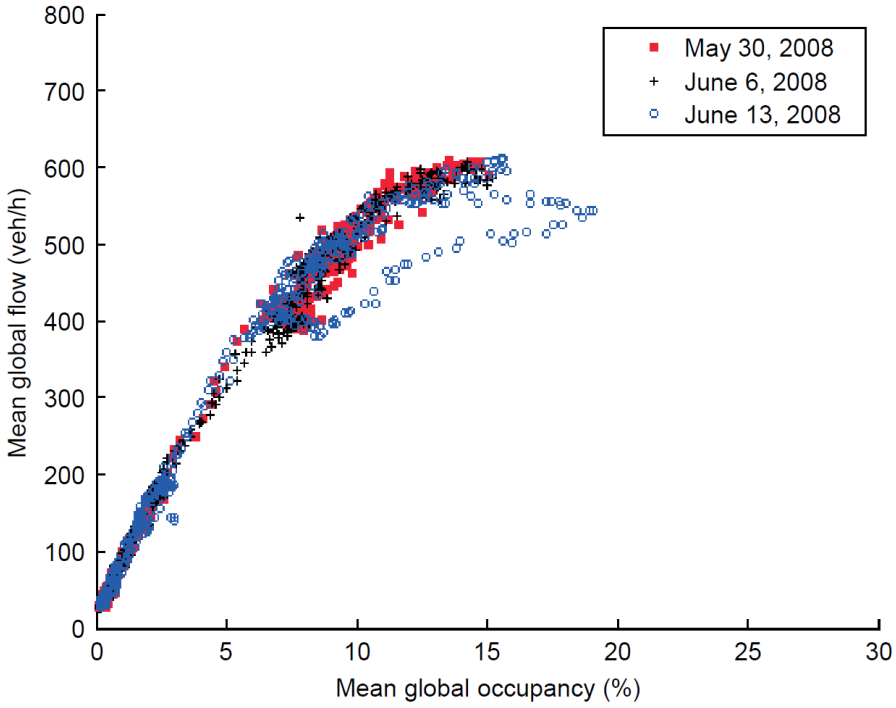


Figure 2.5: *Urban scale MFD for the city of Toulouse on three different days (Buisson and Ladier, 2009)*

This first advise however is contradicted in Courbon and Leclercq (2011), who argue that having all detectors in a short range of variation, causes the detectors to capture a similar traffic situation, reproducing the FD instead of the expected MFD. A uniform or normal distribution of loop detector positions, gives a much better approximation of the MFD, as is captures different traffic states.

Based on modelled data (VISSIM) Ji et al. (2010) investigated the shape of the MFD for a part of the network of Amsterdam, also consisting of urban roads and freeways. They found a sharp transition in the MFD and most of the measurements were in the congested region, as shown in Figure 2.6.

The research of Ji et al. (2010) further focused on the freeways and the effect of ramp metering schemes on the MFD were tested. After implementing ramp metering, the flow on the congested branch seems to be higher at the same level of accumulation. However, as these results are obtained from different directions of this freeway, it cannot be derived from the MFD that ramp metering improved traffic flow. On this they conclude that the shape of the MFD is not only a property of the network and demand, but also of the applied traffic control measures.
The article also investigated a hysteresis phenomenon on the freeway. This is contributed to the fact that weighted flow during congestion resolution is lower than during congestion onset at the same accumulation, because when congestion resolves, demand is, by definition lower than capacity.

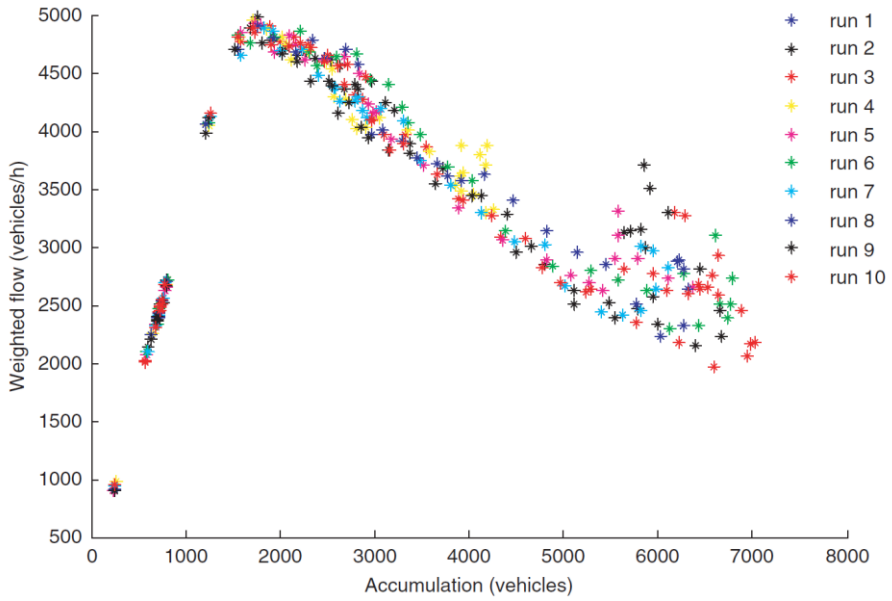


Figure 2.6: *MFD on whole network (Ji et al., 2010)*

They conclude in conjunction with Buisson and Ladier (2009) that before MFDs can be drawn, the network should be divided into homogeneous subnetworks, which can account for impact factors that result in uneven onset and inconsistent offset of congestion, such as traffic demand control, congestion charging policies and non-recurrent bottlenecks. Furthermore they found that rapidly changing traffic demand drastically affects the shape of the MFD and conclude that the shape of the MFD is not independent of demand.

2.3.6 Hysteresis and heterogeneity: Explaining the scatter

The effect of the hysteresis was further investigated in Gayah and Daganzo (2011) and Geroliminis and Sun (2011b).

Using a two-ring system[7], Gayah and Daganzo (2010) showed that at higher densities and congestion, the system becomes unevenly loaded and flow-density patterns become scattered. In Daganzo et al. (2011) these observations are described analytically. Using a two-bin system, it was shown that when small disturbances occur and drivers do not change their routes adaptively, the system becomes unstable and can even go into gridlock. The main reason for these disturbances are turns made at intersections. It was theorised that when having a network with only enough traffic to fill four links, then after a finite amount of time the network would be in gridlock, as all traffic would be wrapped around a block. As such it is not a question if a network can ever jam, but how long this takes on average. It was shown that at higher vehicle densities the network collapses quite rapidly. When increasing the turning rate, the time until collapse is even shorter.

Gayah and Daganzo (2011) expanded on this work, explaining the presence of hysteresis loops. They showed that clockwise hysteresis loops are to be expected even in the most symmetrical of networks. Important factor in this is that congestion recovers more slowly than traffic in free flow. When congestion is unevenly distributed among the network, lower network flows should arise during recovery, resulting in these loops.

[7] http://www.ce.berkeley.edu/~daganzo/Simulations/two_ring_sim.html

In Geroliminis and Sun (2011b) the shape of the MFD using real freeway network data is examined. In the resulting MFD multiple hysteresis loops are present. In analysing the data they identified two causes for this phenomenon. The first cause is the different degree of spatial heterogeneity in vehicle density in the onset and offset of the peak period. A significant second cause is the synchronised occurrence of transient periods and capacity drop in the offset of congestion.
The conclusion of both these articles does seem to match and/or seems to be in line with the hypothesis of Buisson and Ladier (2009) and Ji et al. (2010). It then seems fairly reasonable to assume that hysteresis is indeed caused by the uneven distribution of traffic over the network.

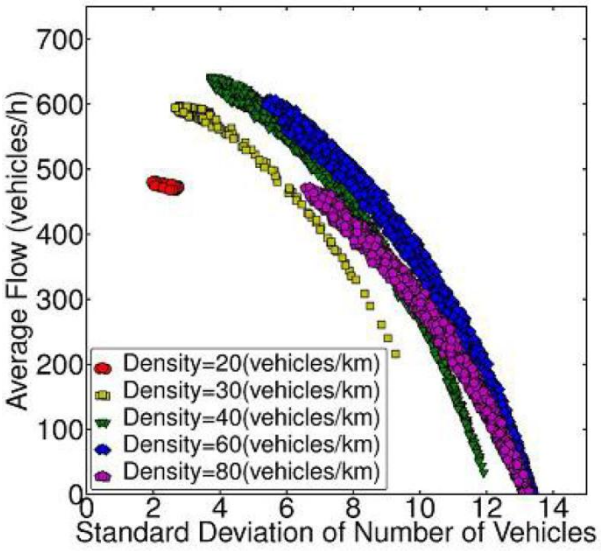


Figure 2.7: *Average vehicle flow as a function of the variability of the number of vehicles (Mazloumian et al., 2010)*

The idea that the uneven distribution of traffic over the network causes scatter in the MFD was first introduced by Mazloumian et al. (2010). In this paper the effect of inhomogeneity in the spatial distribution of densities on the MFD is investigated. For this a new modelling technique, called macroscopic flow quantization is introduced. Using a lattice-like unidirectional road network with periodic boundary conditions, with vehicle destinations concentrated in the centre of the network, they showed that the average flow is determined by the average density and the standard deviation of vehicles, throughout the different links.
When plotting the standard deviation of vehicles S against the average flow, a quite pronounced relationship between the two shows, as can be seen in Figure 2.7. This relationship can be approximated as

$$S^{max} \approx \sqrt{F^{max}[(K-\kappa)l]^2 + (M - F^{max})(Kl)^2}. \tag{2.7}$$

When plotting the average flow $\hat{Q}$ against the number of full links M, an almost linear relationship is obtained, defined as

$$\frac{dQ}{dF} \approx 2\left(\frac{G}{C}\right)\frac{\hat{Q}}{M}. \tag{2.8}$$

Mazloumian et al. (2010) states that while these results are very encouraging, they cannot be directly utilised to develop control strategies because (i) the variability is a time-dependent quantity and (ii) the critical density that maximises flow varies with the variability.
This paper showed that the deviation of density is a key variable which is required (i) for the existence of an invariant MFD, (ii) to explain the wide variation of average network flows at equal levels of accumulation and (iii) to provide a robust and well-defined macroscopic functional relationship even in cases where origin-destination flows significantly vary.
In the article it is stated that further investigations are needed to identify (i) whether these functional relationships also hold for more complex road networks and turning relations, (ii) how traffic congestion spreads with time as a function of topological and demand characteristics, and (iii) how these outcomes can be applied to real cities in order to avoid high levels of congestion.
It is concluded that in order to enhance traffic performance, one should focus on reducing the variability of the vehicle densities. As gridlock is caused by full links creating spill-back queues on upstream links, it follows that the number of full links within a network should be decreased. This in turn could be achieved by prioritizing critical vehicle queues or restricting access to subnetworks exceeding a certain density threshold. The latter is theoretically proposed in Daganzo (2007) and demonstrated by Geroliminis and Daganzo (2007) and Yoshii et al. (2010), which are discussed in paragraph 2.4.1.

In Geroliminis and Sun (2011a) the spatial variability is the subject of study as well. Using real data from the Yokohama network, they showed (using various statistical tests) that evenly-distributed congestion is not a necessary condition for a well-defined MFD. With these outcomes they concluded that the occupancy distribution for different time intervals with similar average occupancy is similar in a well-defined MFD. When investigating the Yokohama data further it was found that the Coefficient of Variation, COV, (standard deviation divided by the mean) is almost constant at detector occupancy values of >10% (see Figure 2.8). The authors concluded from this that one can derive a relationship between the level of spatial heterogeneity and the variance of an MFD.

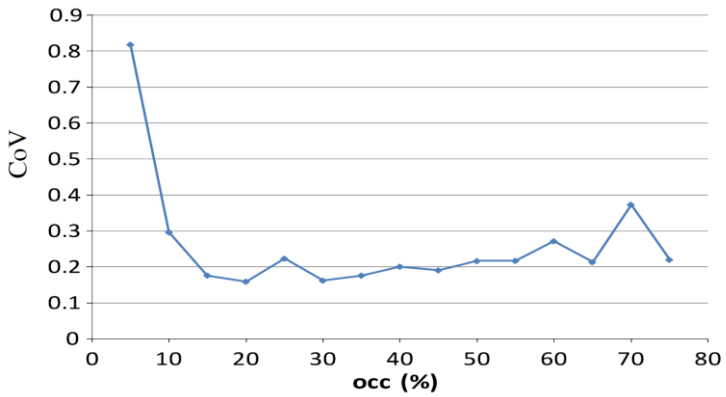


Figure 2.8: *Coefficient of Variation of individual detector flow, q, for different occupancy for a sample of detectors (Geroliminis and Sun, 2011b)*

These results seem to match those of Mazloumian et al. (2010), which mathematically derived this relationship, as expressed in (2.7). It is however interesting to see that this principle does seem to hold for real networks. If the equation itself is correct is not clear though.
When using the same tests for the Minneapolis, Twin Cities freeway network, it was found that different network occupancies were found under congested and uncongested regimes, suggesting that variability in traffic states does influence the shape of the MFD (hysteresis loops). This again seems in line with earlier findings regarding the shape of the MFD for freeway networks.

In order to verify Daganzo's proposition that a MFD with low scatter should be observed when all lanes on all links throughout a network are either in the congested or uncongested regime, Cassidy et al. (2011) looked at vehicle trajectories for two freeway stretches. After filtering the data, to meet these regimes, a somewhat linear relation was obtained, roughly resembling the individual FDs. However, the proposed method removed about 75% of the samples, meaning that the above conditions of a complete congested or uncongested network are not often obtained for freeway networks. As such Cassidy et al. (2011) conclude that freeway MFDs may have a smaller domain of application than their counterparts for surface street networks.
They also found that the network did seem to perform best, when the above conditions were met, suggesting that policies to evenly spread out congestion could in some cases have real merit for serving the most trips. This implicitly suggests that variation of vehicle densities affects the performance of a network and the shape of the MFD.

2.3.7 Signal timings: A force to be reckoned with

The effect of signal settings on the MFD has been investigated by a number of researchers. The effect of signal settings was first shown in Daganzo and Geroliminis (2008), when Variational Theory (VT) was used to mathematically predict the shape of the MFD. Within the framework of VT, a 'stationary cut' (see Figure 2.4) was introduced, creating the upper bound of the MFD. This 'stationary cut' was based on the bottleneck capacity at the most constraining intersection. Assuming the available infrastructure to be sufficient for the traffic demand, causing no spill-back, this bottleneck capacity should therefore be equal to the flow of the most restricted signal.

Using VT Boyacı and Geroliminis (2010) extended this work by incorporating different signal cycles, signal offsets and street lengths and determined their effect on the shape of the MFD. They also showed that every combination of green-ratio and street length (=offset) has an optimal capacity and range, which translates in an optimal signal timing of an intersection, as to keep traffic flow in the 'threshold region'.
Using a series of graphs, the effect of accurate and inaccurate timings and offsets for signal on an arterial was demonstrated. One of these graphs is shown in
Figure ***2.9***.

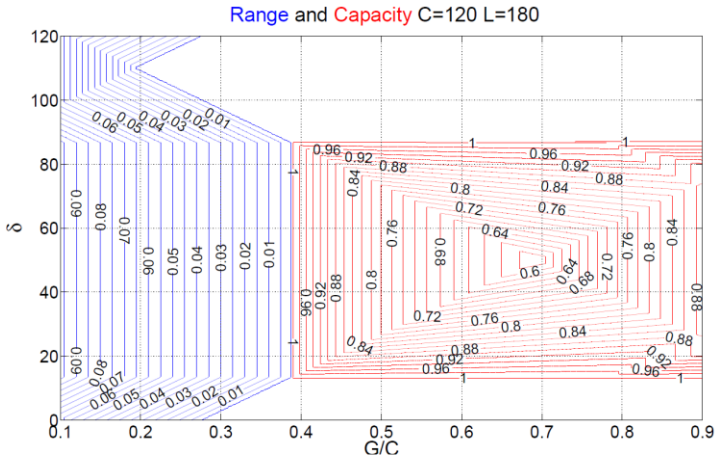


Figure 2.9: *Effect of different signal timings on network capacity (Boyacı and Geroliminis, 2010)*

C: cycle time (s)
L: length of street between intersections (m)
G/C: Permeability (s) = Green time / cycle time
δ: Signal offset (s)

The figures in the blue graph represent the 'range', e.g. the difference between maximum and minimum density yielding the maximum flow.

The figures in the red graph show the ratio between the observed flow and the maximum flow.

Although the above results can be used for arterials with equally sized links, this theory is hard to convert to more heterogeneously spaced networks, with intersecting arterials.

In Laval (2010) a Cellular Automaton model was used to investigate the effects of signal timing and offsets on the shape of the MFD. In accordance with Boyacı and Geroliminis (2010) it was found that an optimal range for signal timing and offset exists and that 'bad' offsets can have a serious impact on the performance of a network.

Helbing (2009) presents an utilisation-based, creating a relation between the average density and signal settings for undersaturated and oversaturated situations. After transferring the link-based urban fundamental diagrams to an area-based one, the following equations are presented:

$$p(u) = \frac{u\hat{Q}}{V^{av}(u)}, \tag{2.9}$$

in which

$$V^{av}(u) = \frac{V^0 ln\{1+[1-f(u)]T_{cyc}/T^0\}}{(1-u)T_{cyc}/T^0} + V^0\frac{f(u)-u}{1-u}, \tag{2.10}$$

$$f(u) = (1+\delta)u \tag{2.11}$$

$$T_{cyc} = \frac{T_{los}}{1-sf(u)} \tag{2.12}$$

Using the equations, a curve was fitted to the data of the Yokohama experiment, described in Geroliminis and Daganzo (2008). Using the same values for $\hat{Q}$ and V^0 as in Geroliminis and Daganzo (2008), a fit was realised only using δ, T_{los}/T_0 and s (see Figure 2.10).

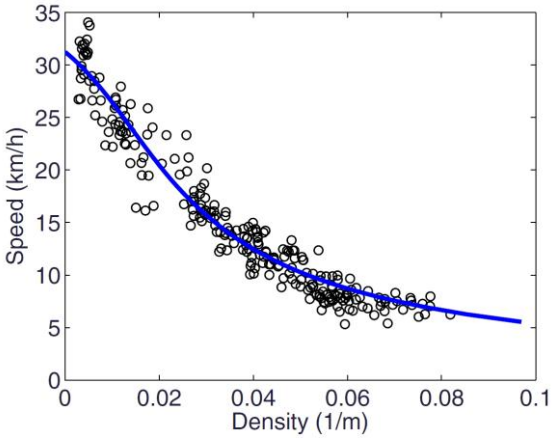


Figure 2.10: *Fundamental velocity-density relationship for a central area of Yokohama. Small circles correspond to empirical data by Kuwahara as evaluated by Daganzo and Geroliminis (2008). The blue curve is fitted using equations (2.9) – (2.12). (Helbing, 2009)*

Helbing (2009) concludes that the shape of this curve is not only dependent on the average density, but also relies on the density distribution, signal operation schemes and other factors as well.

Using a stochastic cellular automaton model, Zhang et al. (2011), made a comparison of Macroscopic Fundamental Diagrams for arterial road networks governed by different types of adaptive signal systems. By changing the boundary conditions, three different scenarios were created (1) time-independent isotropic demand[8], (2) time-independent biased demand (west-to-east bias) and (3) time-dependent isotropic demand.
The different signal systems evaluated are:

1. **SCATS-L**, which uses strings consisting of a master node and multiple slave nodes. The signal timings of each of these nodes can be adapted, in such a way that proper offsets can be created.
2. **SCATS-F**, which optimises every individual intersection, without taking the neighbouring intersections into account.
3. **SOTL**, which also optimises every individual intersection by changing the cycle time and green times, but is also capable of changing the order of the program depending on the traffic state.

For all scenarios it could be observed that the SOTL-system outperformed the SCATS systems, attaining higher network flows and sustaining them over higher network densities. This effect can be contributed to the lower variation of density created by the SOTL system. Also the variation of density seems to be strongly correlated to the presence of hysteresis.
Furthermore it was found that the shape of the MFD depends on the chosen bias and that the MFD has a steep drop in the flow just beyond the maximum flow.

[8] Isotropic demand means that the probability that a vehicle is inserted at and removed from a specific boundary point is equal for all boundary points

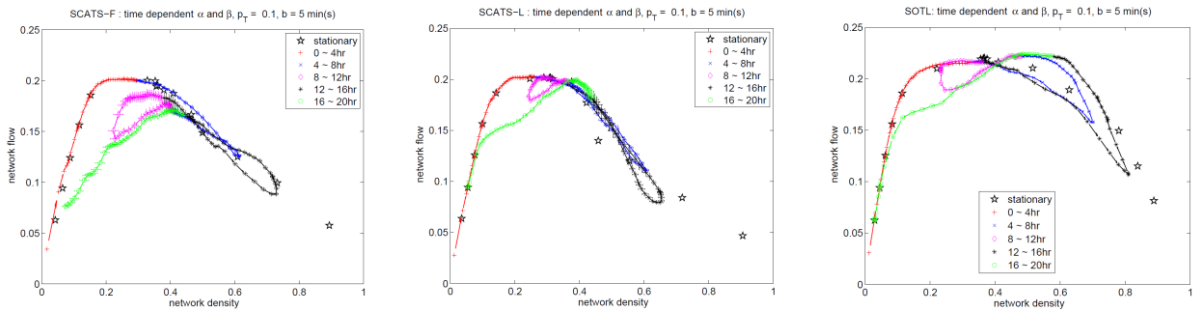


Figure 2.11: *MFDs for the time-dependent isotropic demand scenario (Zhang et al., 2011)*
From left to right: SCATS-F, SCATS-L, SOTL

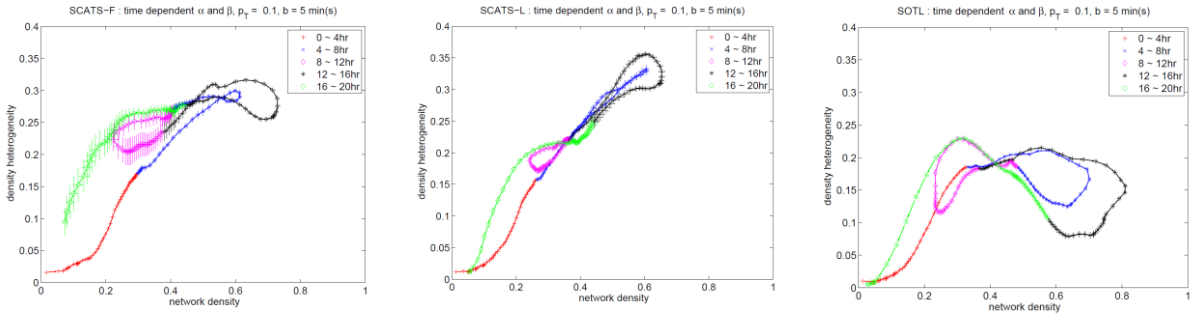


Figure 2.12: *Variation of density, belonging to the MFDs in Figure 2.11 (Zhang et al., 2011)*
From left to right: SCATS-F, SCATS-L, SOTL

2.3.8 Other work related to the MFD

In Wu et al. (2010) a so-called arterial fundamental diagram (AFD) is investigated. Although roughly the same as a MFD, it is used on a smaller scale and falls between the FD and MFD. In this paper a 900 meter arterial equipped with intelligent detectors is investigated. An MFD-like shape is obtained, but scatter is very high. An interesting observation is that two different capacity values are found, corresponding with the morning- and evening peak. This in turn can be credited to bad signal offsets during the morning peak period.

The scatter could partly be explained by the fact that when queuing for an intersection, a vehicle may stop on a detector for a period of time, prolonging the occupancy time, which in turn generates inaccurate values for the flow rate. In this paper a technique is presented to filter out this queue-over-detector (QOD) phenomenon, which relies on reconstructing the correct vehicle trajectories between different detectors and intersections. After correcting for the QOD, almost all scatter is removed from the AFD.

Although the method proposed in Wu et al. (2010) could be used to remove scatter in a MFD, this would probably be difficult to accomplish, as reconstructing vehicle trajectories on a macroscopic scale is quite hard and most scatter is already filtered out by aggregating the data. Nevertheless, the resulting errors in measurement could be substantial, especially in the congested branch of the MFD.

2.4 Fields of application

2.4.1 Perimeter control

The development of the macroscopic fundamental diagram was primarily a by-product of Daganzo (2007) as it was a part of a newly developed adaptive control approach to improve urban mobility and relieve congestion.

In paragraph 2.3.2 it has been shown that the number of traffic exiting a road is a fixed fraction of circulating traffic, which has been expressed in equation (2.2).
If one would fix the number of vehicles in the loop, used in this example, and inputs are regulated to maintain the system in a steady state, than

$$n = \int_0^{L_T} k(x)dx \tag{2.13}$$

and the distribution of cars that maximises total outflow can be expressed as

$$p(x)w(k^*(x), x) = constant \tag{2.14}$$

in which the function $k^*(x)$ maximises (2.13) subject to (2.2) and $w(k, x)$ is the kinematic wave speed. This implies that when the constant is zero, $w(k, x)$ is zero and the road should be at capacity everywhere. This lead to Daganzo's first insight:

"If conditions do not change rapidly with time, a road should not simultaneously have congested and uncongested portions."

From equation (2.14) it can also be derived that the locations with the highest exit rates $p(x)$ should have the smallest value for the kinematic wave speed $w(k, x)$; i.e. the flows closest to capacity. This lead to Daganzo's second insight:

"If conditions do not change rapidly with time, system output is maximized when flow is at capacity only along road stretches with the greatest exit rates $p(x)$; i.e. the greatest density of destinations, $\lambda(x)$."

Daganzo argued that if the conditions on which the above two insights are based should hold for networks with a much more complex geometry, the following (third) insight should be true:

"Assume $p(x)$ is fixed and conditions do not change rapidly with time. Then, the rate at which trips are served in a metropolitan area is maximized only if capacity flows occur in the neighbourhoods with the greatest $p(x)$ (i.e. the greatest density of destinations $\lambda(x)$), and elsewhere the network is either congested all over or uncongested all over."

This third insight lead to the development of the *AB strategy for cities* in which AB stands for Aggregation Based. This strategy suggest the following:

"To maximize the rate at which trips are served in a metropolitan area, a policy should maintain near-optimal accumulations in the neighbourhoods with the greatest density of destinations."

One of the ways that these near-optimal conditions can be maintained is by implementing a perimeter control strategy. In such a strategy the access to a particular subnetwork (or neighbourhood) is temporarily restricted or reduced, which can be achieved in a number of different ways, such as congestion charging, reversing lanes or adapt the signal timings. Especially the latter could be quite effective, as the number of vehicles entering the particular subnetwork can be controlled very precise. The aforementioned near-optimal accumulation, is the number of vehicles resulting in the highest production of that network. This value, or range thereof can be determined using the MFD, as this is the highest point in the diagram. In Figure 2.1 the range of accumulations resulting in the highest average network flow is referred to as the 'threshold region'. When looking at some of the figures in the previous section, this point or region can clearly be identified on all of the different MFDs. For some networks this point could even by determined analytically, for instance by using VT.

To test the AB strategy for cities defined in Daganzo (2007), Geroliminis and Daganzo (2007) applied the theory to the downtown city centre of San Francisco, using a model. This whole network was considered as being a single reservoir, in which the demand, generated at the boundary, was restricted in such a way that the accumulation within the network was maintained as close to the optimal value as possible.

They found that when applying this perimeter control strategy, the total outflow of the system increased by 34% in the same period of time, showing that perimeter control can indeed have a substantial impact on improving traffic flow in a network.

Apart from this Geroliminis and Daganzo (2007) also found that (i) the shape of the MFD seemed to be invariant of OD-tables, (ii) system recovery at high accumulations is very difficult and (iii) that a linear relation between system output O and production P exists, which is the reciprocal of average trip length $\bar{L}$ and can be described as

$$O/P = 1/\bar{L}. \tag{2.15}$$

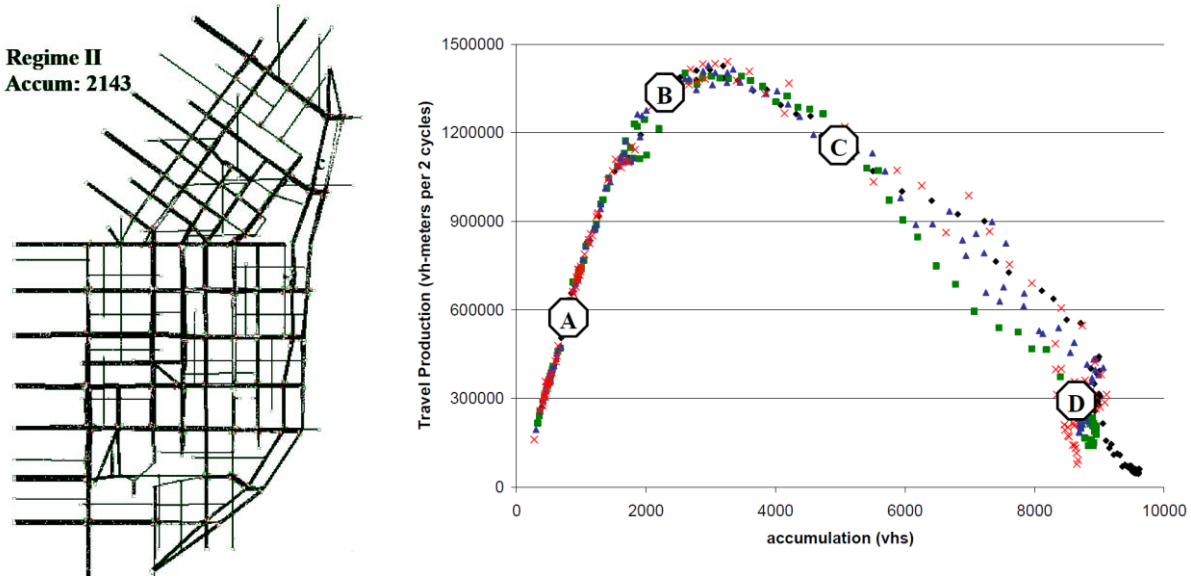


Figure 2.13: *San Francisco network and resulting MFD (Geroliminis and Daganzo, 2007)*[9]

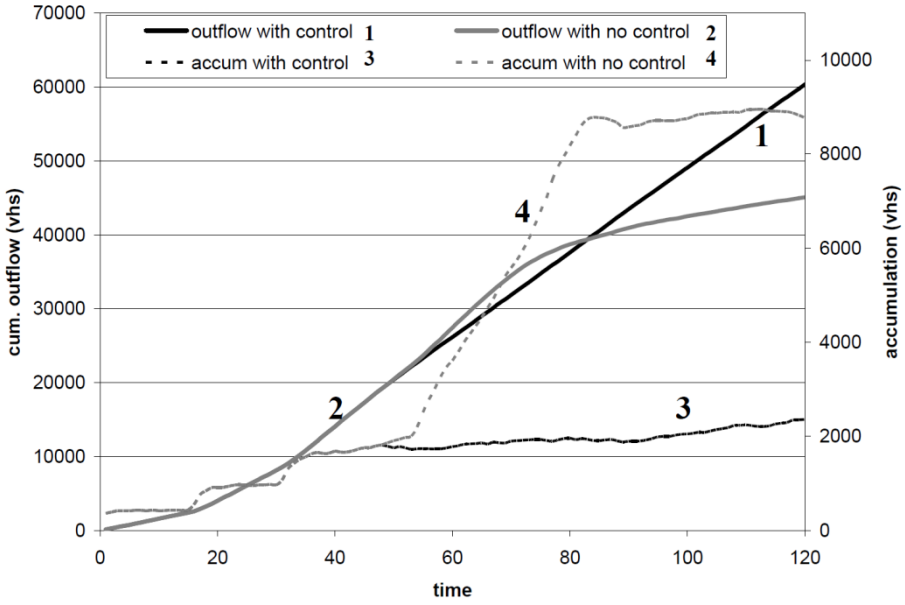


Figure 2.14: *Outflow and accumulation with time for the San Francisco network with and without control in the boundary (Geroliminis and Daganzo, 2007)*

In Yoshii et al. (2010) an area metering control method that can be applied to urban expressway road networks, has been proposed and tested. Using a mesoscopic dynamic traffic simulation model and a part of the Hanshin expressway in Osaka, Japan (see Figure 2.15), they controlled the traffic inflow at the onramps using the MFD. From the MFD the critical value for density was derived k_c, i.e. the number of vehicles in the network, which generates the highest performance. When the average network density K is above this critical value, the control is switched on. In this case, the onramps that will be controlled are chosen by checking whether the next downstream link

[9] A movie of this simulation can be found at http://www.ce.berkeley.edu/~daganzo/Simulations/MFD/MFD.html

was congested in the previous time interval. The allowed inflow rate per controlled onramp is then calculated as

$$N_l(t) = q_l(t-1)r(t), \tag{2.16}$$

$$r(t) = \left[1 + \frac{\sum_{i\epsilon B} q_i(t-1) - \{K(t-1) - k_c\} - \sum_{i\epsilon A} q_i(t-1)\Delta t}{\sum_{i\epsilon C} q_i(t-1)\Delta t}\right] \tag{2.17}$$

Where,

$N_i(t)$: controlled traffic flow rate at link i in time interval t (veh/s)

$q_i(t)$: average traffic flow rate at link i in time interval t (veh/s)

$K(t)$: aggregated traffic density at the end of time interval t (veh)

$r(t)$: multiplier at time interval t (-)

A: inflow link set towards the loop

B: outflow link set from the loop

C: controlled link set

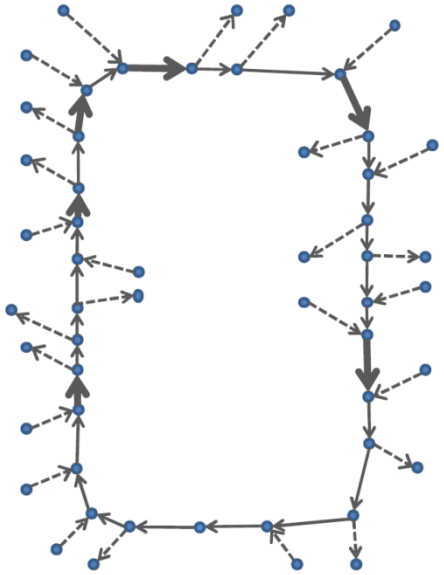


Figure 2.15: *Schematization of Hanshin Expressway study network (Yoshii et al., 2010)*

With equation (2.17) the remaining capacity is evenly divided over the controlled onramps. The area metering control method of Yoshii et al. (2010) decreased the total delay in the network by 38%, again showing that restricting access to the network can significantly improve traffic flow.

In a recent study by Geroliminis et al. (2012), the MFD is used in a two-region system, governed by a perimeter control strategy, in order to maximise the number of trips reaching their destination. Within this system, the shape of the MFD is estimated using a 3rd-degree polynome. Using a prediction model with a rolling horizon, the accumulation at $t + \Delta t$ for both regions is estimated. Based on the estimated accumulation and the estimated MFD, the amount of vehicles that may transfer from one region to the other is calculated.

In a case-study, the results of this method are compared against a greedy control strategy. This greedy control maximises the interregional flow when the accumulation is suboptimal and minimises the flow when one of the regions is congested. In case both regions are congested, the flow towards the 'more congested' region is minimised, while the flow towards the 'less congested' region is maximised.

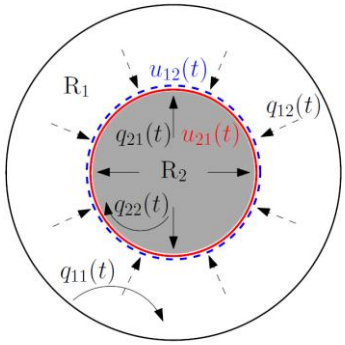


Figure 2.16: *Two region MFDs system (Geroliminis et al., 2012)*

The results of the case-study show that the MPC-strategy outperforms the greedy control strategy in every scenario. Especially when the system reaches it congested state, the MPC performs significantly better than the greedy control strategy (up to 20% better).

It should be remarked that the MFDs used in this study are assumed to be constant and do not change over time. However, as changes are made to the number of vehicles that can travel from one region to the other, one could expect that the shape of the underlying MFD would change and therefore have an impact on the effectiveness of the control strategy.
Another remark that can be made is that the effectiveness of this strategy is only tested against a greedy control strategy. However, no methods using static, or local demand adaptive perimeter controllers have been assessed. Such models would probably better represent actual traffic conditions and network operations, as they would operate like fixed-time controllers, respectively self-optimizing traffic lights.

2.4.2 Assessing DTM-strategies

Another field in which the MFD could be used, is the assessment of strategies for dynamic traffic management (DTM). As shown earlier, the shape of the MFD can be affected by things like type of infrastructure, homogeneity of densities and applied traffic control measures.
From Geroliminis and Daganzo (2008) it followed that the maximum production of an urban network depends on the bottleneck capacity caused by traffic lights (maximum g/c-ratio). The amount of scatter can be explained by homogeneity of traffic densities and the slope of the free flow branch can likely be attributed to the average network speed.
This means that when new control strategies are implemented, or changes to the infrastructure are made, the shape of the MFD will most likely change. When for example new DTM-strategies are

implemented, an increase in the production could be witnessed, or the amount of scatter would be lower. It could also be the case that the amount of points in the congested region is lowered.

Although the effects can of course be evaluated locally, this does not always give an accurate answer. It could be that a reduction of the maximum speed seems to improve traffic flow locally, but has a negative impact on a network level, as drivers change their route. Using the MFD for the before and after situation can thus show us to what degree the implemented measure has a positive (or negative) effect on the complete network.
It should be noted that this does not always work, as the MFD aggregates traffic data over a large region. Small, local improvements could then become lost in the aggregation and an improvement cannot be witnessed.

Using the MFD Qian (2009) evaluated the implementation of an extra lane and ramp metering for the Amsterdam network, as in Ji et al. (2010). The results of this are shown in Figure 2.17.

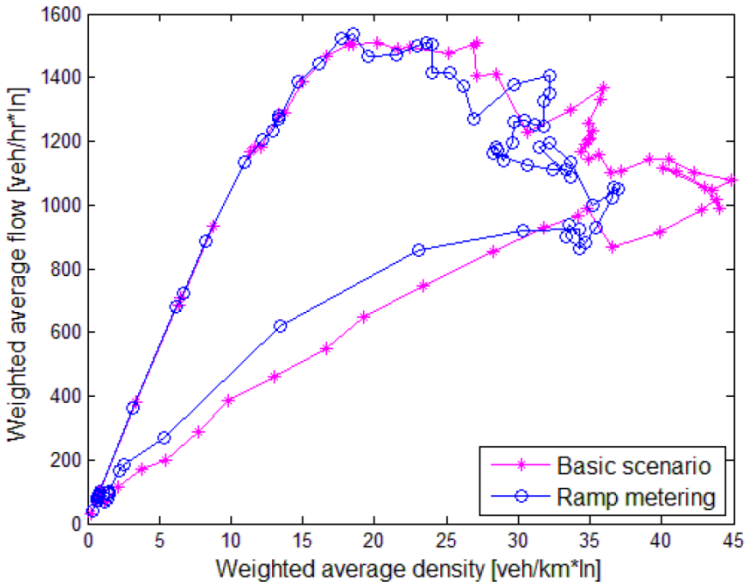


Figure 2.17: *MFD of the motorway with ramp metering implementation (Qian, 2009)*

As can be observed from Figure 2.17, implementing the ramp metering improved traffic flow, as (i) the traffic densities are kept lower and (ii) after resolving the congestion, the system restores at a higher average flow. The amount of scatter in the MFD is however quite high and using this as a tool for evaluation would be too rough. Qian (2009) therefore advices to use the MFD as a supplement to the more conventional assessment methods.

2.4.3 Routing strategies

As the subnetwork accumulation and the standard deviation of vehicle accumulation seems to be related to the network performance, Knoop et al. (2011b) investigated if traffic could be controlled using only this information. To this end Knoop et al. (2011a) tested whether this relation holds

using a grid network with periodic boundary conditions as proposed before by Mazloumian et al. (2010).
It was found that the standard deviation of subnetwork vehicle accumulation could be used as an indicator for the performance of the network. This conclusion seems to be justified because of the absence of scatter in the diagram (see Figure 2.18). At no point in the diagram two different values of performance are found at the same level of standard deviation of (subnetwork) vehicle accumulation. On the other hand this could also mean that the standard deviation increases continuously and thus every standard deviation only generates a single value for performance. Multiple runs at the same density should show if scatter arises, or that these curves are invariant to further network dynamics. In Knoop et al. (2011b), different simulations for routing strategies were made. At a low level of standard deviation, these curves all follow exactly the same line. As destinations are randomly chosen, the traffic patterns are different for each run. From this it could be concluded that a direct relationship between performance and standard deviation of density exists. This seems to be in line with the results of Mazloumian et al. (2010) which showed a direct relationship between the standard deviation of densities and network performance (equation 2.7).

As Knoop et al. (2011a) showed that the vehicle accumulation in subnetworks seemed to be a good precursor for network performance, Knoop et al. (2011b) tested different routing strategies using this information. In these strategies vehicle routes were updated every 15 minutes, based on (1) speed on all individual links, (2) average subnetwork speed and (3) estimated subnetwork speed based on accumulation in that particular subnetwork. For strategy 3 the shape of the MFD was assumed, and based on the accumulation, the subnetwork speed was determined. Two different MFD-shapes were used, the first shaped similar to a triangular fundamental diagram (Daganzo, 1997) and the second as the fundamental diagram used by Drake et al. (1967).

The results of these simulations are given in Figure 2.19. It shows that routing based on the speed of all links performs best. The 'jumps' in the line are related to the routing updates given every 15 minutes, which temporarily increases production. Without these updates, production would eventually drop to the same level as no routing. Although providing route advice based on subnetwork accumulation performs less, this method still improves traffic flow and is far less data intensive than speed routing. More importantly, these results show that it is possible to use the MFD (albeit in an assumed form) to improve traffic flow in a network.

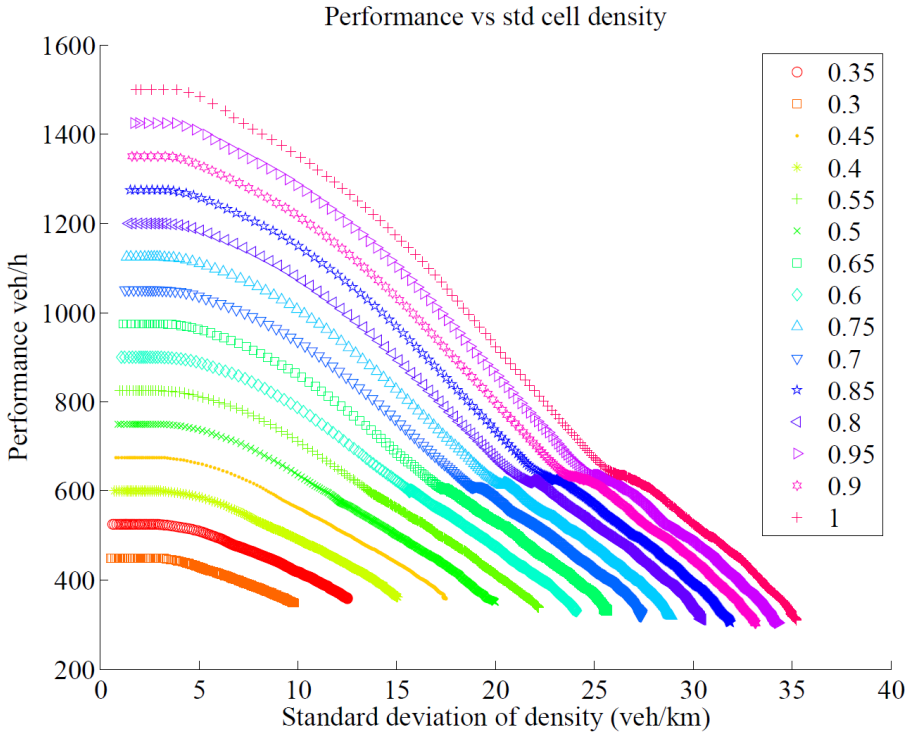


Figure 2.18: *The network performance as function of the variation in density for different network loads (Knoop et al., 2011a)*

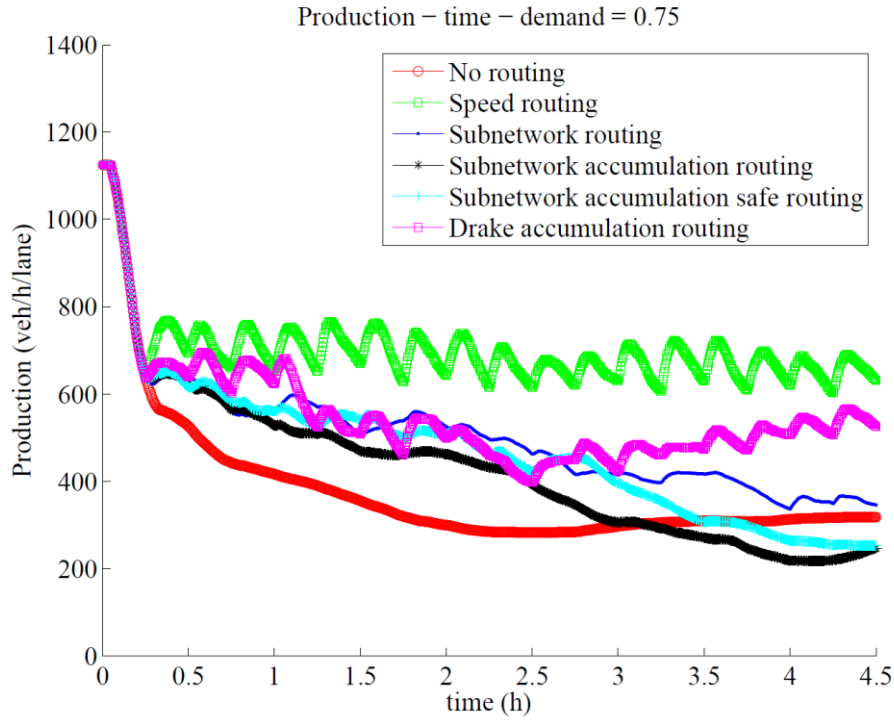


Figure 2.19: *Network production over time under different routing strategies (Knoop et al., 2011b)*

2.5 Analysis, criticism and applicability

2.5.1 Analysis of the literature

Although the research into the macroscopic fundamental diagram has only recently taken of, the amount of literature that has already been written on the subject is quite substantial.
The most notable thing is that most of these studies have aimed at finding out the finer intricacies of the MFD and only a small amount have looked at methods how to apply this knowledge.

Regularity conditions and new insights

In Daganzo and Geroliminis (2008) it was conjectured that a well-defined relationship with low scatter between production and accumulation should arise if the following 'regularity conditions' are met:

1. A slow-varying and distributed demand;
2. A redundant network ensuring that drivers have many route choices and that most links are on many desirable routes;
3. A homogeneous network with similar links;
4. Links with an approximate FD that is not significantly affected by turning movements when flow is steady.

In addition to the original regularity conditions (which still seem to hold, but can be somewhat more relaxed, as the DoD makes the MFD more applicable for heterogeneous networks), new insights have been gained, concerning factors influencing the shape of the MFD. Although most studies so far have looked at only one type of network within their research and/or have tried to explain particular anomalies within their 'own' MFD, when comparing the different MFDs throughout all these studies, the following observations are made:

- Freeway MFDs have quite a sharp transition between the free flow and the congestion branch, urban MFDs are more 'rounded off';
- Higher levels of heterogeneity of traffic and infrastructure creates more scatter in the diagram and lowers the production;
- Most MFDs show a strong relationship with low scatter between the network performance and variability of link densities.
- The effect of signal timing and offset is a key component in the shape of the MFD;
- Hysteresis only seems to occur in freeway MFDs;
- Differences in OD-demand do not have a significant impact on the shape of the MFD;
- The MFD becomes less scattered as the size of the network increases;

When looking at the studies done so far, it seems that a consistent MFD is easier to obtain for urban networks, than for freeway networks. The reasoning behind this is that weighted flow during congestion resolution is lower than during congestion onset at the same accumulation, as demand is lower when congestion resolves. Another important factor is also the uneven distribution of traffic over the freeway network. Although this factor also applies to urban networks, hysteresis seems to be absent in urban networks, which can most probably be attributed to the fact that

traffic can redistribute itself much easier in urban networks than in freeway networks, as more alternative routes are available.

Applicability to heterogeneous networks

One of the most important questions today, is how far the regularity conditions can be stretched, as to make the MFD applicable in networks that consist of different link types, have fast changing demand patterns and have an inhomogeneous distribution of traffic.

Looking at the scatter in the congested branch of most MFDs, it seems that the theory is not directly applicable to heterogeneous networks. In Geroliminis and Sun (2011a), Ji et al. (2010) and Buisson and Ladier (2009), mixed networks were used, consisting of a freeway and an urban network. The resulting MFDs are all highly scattered. In Buisson and Ladier (2009), the data was then split into two separate sets, which resulted in fairly low-scattered MFDs. The urban network however, was fairly heterogeneous, with a lot of different types of roads, implying that links do not have to be similar. The differences between the urban and freeway MFDs could be explained by hysteresis and the absence of multiple routes on the freeway.

In multiple studies, the effect of different demand patterns have been investigated (data from different days). In most cases, the resulting MFD was fairly similar. Only on a lower level of scale differences in the MFD were found (Wu et al., 2011), providing evidence that different demand patterns do not significantly influence the shape of the MFD.

One of the most interesting findings so far is that there seems to be a strong relationship between the distribution of traffic over the network and the shape of the MFD. Although the theory cannot be directly applied to heterogeneous loaded networks, using the variance in traffic can still produce accurate results.

In light of the above and especially when taking the distribution of traffic into account, it seems that decent MFDs for heterogeneous loaded networks can be produced.

Variability of densities

The effect of the distribution of traffic over the network on the shape of the MFD, was first recognised by Helbing (2009). Subsequently, Mazloumian et al. (2010), Knoop et al. (2011a, 2011b), Geroliminis and Sun (2011a) and Zhang et al. (2011) have shown that a strong relationship between the average flow and the standard deviation of vehicle densities over the network exists and that the highest flow is obtained when traffic is distributed evenly over the network (confirming Daganzo's earlier condition). This too seems to be most logical, as deviation of vehicles causes the aggregated data to contain measurements from the free flow and the congested state of the individual link's FDs, creating intermediate points that are by definition below the branches of the FD. This effect is clearly demonstrated in Cassidy et al. (2011).

These findings also suggest that control strategies should focus on distributing the traffic as evenly as possible. Perimeter control can be a strong tool for this, as it can control the number of vehicles entering different parts of the network. Perimeter control should therefore not only focus on controlling the number of vehicles in the subnetwork, but also try to distribute them as homogeneous as possible. Although not used in the form of perimeter control, improving the network performance by reducing the variability has been done in Knoop et al. (2011b).

Trip completion rate

Another important finding, is that there seems to exist an almost linear relationship between the average flow and the trip completion rate, as shown in Geroliminis and Daganzo (2007, 2008) and Geroliminis and Sun (2011a). This is important, because in real life the number of completed trips cannot be measured. When using strategies to maintain the accumulation in a network at the optimal level, it is important to know how much vehicles are present in that network. When optimal accumulation is reached, the number of vehicles that can enter the network in a certain interval, should than be equal to the number of vehicles that completed their trip, minus the number of trips that were created in that network during the interval. How to obtain the number of trips created in a network or if they can be related to the average flow as well, has not been studied yet.

Signal settings

The effect of signal settings on the MFD was first studied by Helbing (2009). In this article the effect of signal settings is incorporated in the calculation of the average network flow, which is then related to the network density. The resulting formulation thus predicts the average network flow based on the density, size of the network and signal settings.

Using Variational Theory, Daganzo and Geroliminis (2008) showed that the signals are a highly important factor in the MFD, as they form the upper bound of the MFD-curve (assuming the available infrastructure is sufficient to meet the traffic demand). Boyacı and Geroliminis (2010) extended this work by incorporating different signal cycles, offsets and street lengths, showing that badly timed signals can have a highly adverse effect on the overall performance of a network.

In Strating (2010) it was found that the shape of the MFD remained constant when using fixed time control. When using self-optimizing signals differently shaped MFDs can arise even when the control strategy is constant.

The most thorough study regarding the effect of signal settings on the MFD so far is Zhang et al. (2011), in which different signalling systems were tested. This study showed that different signal systems can result in differently shaped MFDs. They also found that the drop in network performance after congestion has set in, is much higher when demand is biased, i.e. not uniform over space. More important however is that they found a strong relationship between the shape of the MFD and the variation in density, as the signal system that created the most homogeneous conditions resulted in the highest performance, which could also be sustained over a larger density region (see Figure 2.11 and Figure 2.12).

Taking the above mentioned studies into account, it can be concluded that signals play an important role in the performance of a network (as might of course be expected). Their role within a network however might still not be fully understood. Although their main purpose is to improve traffic by separating conflicting streams, their regulating function also reduces the inflow to areas downstream and holds traffic back in upstream areas, thus contributing to a reduction in the variability in density over the links in the network. Interestingly enough, the study by Zhang et al. (2011), seems to imply that self-organizing traffic lights are better in reducing this variability, than coordinated signal systems are.

In view of the above, one could argue that an effective control strategy could be created by optimizing every intersection in such a way, that the average density of the downstream area/links remains equal to the average density of the area/links upstream (up to the first upstream intersection). Although coordination on a network level is needed, every intersection can easily be optimised using only traffic data from the upstream and downstream links.

Application

So far only a few studies have tried to apply the information contained in the MFD. The most important is perimeter control, which has been studied by Geroliminis and Daganzo (2007), Yoshii et al. (2010) and Geroliminis et al. (2012). In all of these studies, it was found that substantial improvements in the networks performance (up to 30%) could be achieved. These results are very encouraging and by themselves provide enough base to justify the continuing research towards the MFD.

Another promising field of application is to use the MFD to assess the improvements in the networks performance after making changes to the infrastructure or the way it is controlled. Although this assessment is fairly straightforward, this feature has only been exploited in Qian (2009) to investigate the global effect of ramp-metering strategies. The amount of scatter in the resulting MFDs however were too rough be able to use it as an evaluation tool. The problem is that when the network under investigation is too large, more local effects are lost in the aggregation. When the size of the network is chosen too small, the global effects are not taken into account. The latter could lead to a MFD that shows an increase in performance on a local scale, but misses the decrease of overall performance due to rerouting or spill-back.

The third field of application is routing strategies and has been studied by Knoop et al. (2011a, 2011b). The main reason to use the MFD is that the data demand to construct a MFD is lower than conventional methods. It was illustrated that by using the data of subnetwork MFDs, traffic could be rerouted away from overcrowded subnetworks, increasing the network performance.

2.5.2 Criticism

One of the acclaimed properties of the MFD is that it is independent of OD-demand. Although the results from Geroliminis and Daganzo (2007, 2008) confirmed this statement, the results from Buisson and Ladier (2009), Wu et al. (2010) and Zhang et al. (2011) stated otherwise. Although the results from Buisson and Ladier (2009) could be biased by a strike of truck drivers, Wu et al. (2010) properly explained the difference in the shape of the MFD for the AM and PM peak period, as a consequence of bad signal offsets in one of the two directions.

The study by Zhang et al. (2011) showed that the network performance degradation in the congested regime is higher when traffic is biased, which shows that the shape of the MFD is not independent of the demand pattern. It could however be the case that due to the different demand patterns the signals are operated in a different way, causing the differences in the shape of the MFD.

Another reason why the MFD might not be independent of the OD-demand, is that the theory proposed in Daganzo (2007) assumes that the exit taken by a vehicle is independent of its origin. Although this could be true for more homogeneous networks like grid or radial networks, this does not immediately hold for networks with a heterogeneous infrastructure, as (i) the number of 'exit

opportunities' is quite limited and might very well be dependent of a driver's origin and (ii) the number of alternative routes is limited.

Another important point which has not been dealt with properly is the type of fundamental diagram used for urban traffic. Most models have been using a triangular shaped fundamental diagram. Although this is a much used form of the FD, it has not been proven that this FD actually holds for urban traffic. In Helbing (2009) an attempt is made in constructing a FD for urban traffic flow, which incorporates traffic signal cycles. Helbing concluded that using the triangular FD in an urban network could produce highly inaccurate results and one has to compensate for the effect of signal timings. As such it remains to be seen if the results of the models used so far produce accurate MFDs, especially as no study so far has validated model outcomes with real data.

Regarding the application of the MFD, some remarks can be made as well, especially regarding the application of the MFD in perimeter control. In the studies done so far, substantial improvements in the networks performance were found. In Geroliminis (2007) the traffic was just held at the boundary and no use of traffic signals was made. The effects of holding this traffic back at the boundaries on the other parts of the network have not been investigated. As such it could be that the positive effects in the city centre are countered by the disrupted traffic flow in other parts of the network. The same basically applies to Yoshii et al. (2010).

2.5.3 Applicability of literature to this study

The most important question that has to be answered at this point, is to what degree the available literature is sufficient to answer our research questions as set out in section 1.4 and if it is possible to use the knowledge gained to build a proper model for evaluating the effect of signal settings on the MFD.
One of the goals of this thesis is to investigate how the shape of the MFD of a subnetwork and its perimeter is affected by adapting the signal settings on the perimeter and if these are in some way related to one another. Apart from that, the effect of different network structures is investigated as well.

No study done so far has made an effort to systematically investigate the shape of the MFD in relation to the network structure and the way it is controlled. Although some theoretical formulations for the shape of the MFD have been made, none of these have been tested if they can properly predict the shape of the MFD of any arbitrary network.
In Zhang et al. (2011) the effect of different signalling systems on the same network, using different scenarios was tested. Although a qualitative comparison of the performance of each system was made, the effects were not quantified, as to make them transferrable to other networks or scenarios.

Also the relationship between MFDs of adjacent (sub)networks has never been investigated, even though it should be expected that adjacent (sub)networks highly affect each other, as one can

form the bottleneck for the other. Especially when the MFD is used as input for a perimeter control strategy this can be of major importance, as has been stressed earlier.

A point that has been stressed in multiple articles, is that if one wants to control a city using the MFD, it should be divided in smaller, more homogeneous subnetworks (also called neighbourhoods), as to obtain well-defined MFDs. However, up to this point it is not clear how large these subnetworks should be. Also the effect on the MFD by choosing differently sized (sub)networks and (sub)network structures has not been investigated systematically, as to make any proper statement on the matter.
Although an automated method to create subnetworks has been proposed in Ji and Geroliminis (2010, 2011), this method is quite complicated. Another problem with it, is that the shape and position of each subnetworks changes over time, as the algorithm takes the traffic situation into account. When these subnetworks keep shifting, it is hard to implement a control strategy. Furthermore, this algorithm only seems to apply to grid-networks and not to heterogeneous networks with many different link types.

2.6 Conclusion

The literature study has shown that the research questions set out in this study are far from trivial and cannot be answered by literature alone. Although some research into the relationship between infrastructure, signal settings and the shape of the MFD has been done, this has never been done systematically enough, in order to be able to relate specific parameters to the shape of the MFD.

Although a number of important factors have been found that influence the shape of the MFD, these factors have not been quantified in such a way that the shape of the MFD can be predicted beforehand (at least not in a simplified manner). It is therein that the lies the challenge for this thesis.

3 Assessing factors impacting the macroscopic fundamental diagram

In the previous chapter it was concluded that studies done so far, have shown a number of factors influencing the shape of the macroscopic fundamental diagram. However, none of the studies has done this in a systematic way, in order to investigate how the shape of the MFD changes if different network structures are used, or signal settings are changed. In order to make the first step towards a framework, which should be capable of making these evaluations, it is important to gain more insight in the different parameters that make up the MFD. To this end, section 3.1 will kick off this chapter, by analysing the underlying parameters of the MFD and the relations resulting from them. Using these findings, section 3.2 will discuss how the shape of the MFD is assumed to be impacted when these parameters are changed. In section 3.3 a conclusion will present the main findings of this chapter.

3.1 Exploring the macroscopic fundamental diagram in-depth

In the previous chapter a number of factors are found, that have an influence on the shape of the MFD. Some of these factors are: network length, free flow speed, signal settings, signal offsets and the distribution of traffic over the network.
In order to get some more understanding of how any of these factors impact the shape of the MFD, we will start by simply looking at the parameters that make up the MFD.

3.1.1 Basic principles

In most cases, the MFD is a data-driven representation of traffic operations within a network and is (often) not created from a continuous function. The MFD can therefore be regarded as a scatterplot, in which each point represents a single value for the production and accumulation, measured over a certain period of time (usually between 5 and 15 minutes), which basically shows the average amount of vehicles in the network over that time period (accumulation) and how much distance these vehicles have travelled (production).

It should be stressed that the resulting diagram itself does not have a temporal component, and that the graph does not show the evolution of traffic over time. This means that any point found within the MFD, can be from any point in time over the period taken into consideration, e.g. points found on opposite sides of the diagram could stem from subsequent time intervals. Although this seems rather obvious, it is actually a commonly made mistake.

3.1.2 Production and performance

Production

On the y-axis, one normally finds the production, or the performance. The production PD is calculated by multiplying the flow q with the length of the part of the network it represents l, for a given period of time Δt and summing this for all links (or measurements) i in the network Z.

$$PD_{\Delta t} = \sum_{i \in Z} q_{i,\Delta t} l_i \tag{3.1}$$

Now as the flow is measured as $q_{i.\Delta t} = n_{veh,i,\Delta t} \frac{3.600}{\Delta t}$, the resulting value for the production shows the amount of kilometres driven per hour, which is basically an extrapolation of the current traffic state, and is therefore not equal to the distance that has been travelled over the considered time interval. This can be somewhat confusing, but is done to make the MFD independent of the chosen time interval.

As has been stressed before in paragraph 1.3.1, the production does not depict the efficiency with which the traffic is processed in the network, but only the distance that can be travelled in the network. This becomes immediately apparent from equation (3.1), because if the flow q is kept constant, but the distance l is lowered, the total production decreases. Now assume the total length of the network would be doubled at the same flow (which is just the same amount of vehicles passing each link or detector, i.e. the number of vehicles driving through the network is unchanged), then the production would also double, even though with respect to traffic, nothing has actually changed, perhaps apart from an increase in travel time.

Performance

In order to get some insight in the efficiency with which the traffic is processed, the performance is used. The performance PF is obtained by dividing the production by the total network length, which in essence is the average (weighted) flow and is expressed as

$$PF_{\Delta t} = \frac{\sum_{i \in Z} q_{i,\Delta t} l_i}{\sum_{i \in Z} l_i}. \tag{3.2}$$

As the performance takes the length of the network out of the equation, only the average flow is left, making it possible to compare different MFDs to each other, i.e. compare the efficiency with which traffic is processed.
A drawback of using the performance is that it does not give an indication of the capacity of the specific network, which could lead to a wrong interpretation of the quality of that network.

In equation (3.2), the performance is calculated using a weighted average. However, the performance can also be calculated using an unweighted average, which is no more than the average of all flow measurements, i.e. the length of the corresponding link, or part of the network represented, is not taken into account. Although using the unweighted average is less accurate, it

is sometimes used, as it does not require a detailed knowledge of the network. Especially when working with real data, taken from local detector measurements, it cannot be fully specified which part of the network is represented by every measurement location. Thus in order to free oneself from any pre-assumptions, the unweighted flow is often used. However, when the data is generated using a simulation model, one often does have this knowledge, as the data is generated for each link in the network, meaning that a weighted average should be preferred.

Production vs. Performance

In order to gain some more insight in the differences between production and performance, the different MFDs for a network are presented in Figure 3.1 to Figure 3.4. The MFDs are taken from a network identical to the network presented in Figure 1.2.

In the first two figures, three different MFDs are presented. The first MFD is based on the arterials in the network, the second MFD is based on the subnetworks and the third MFD is a combination of the arterials and the subnetworks and therefore represents the total network. The difference between these two figures is that in Figure 3.1 the production is used, and in Figure 3.2 the performance. The same is done for the last two figures, but in this case the MFD is created for both subnetworks and their perimeter.

If the network, and the different parts thereof would be assessed using the performance, one should reach the conclusion that the arterials are the best performing part of the network and that keeping the arterials at their optimal accumulation would be the best control strategy. The subnetworks on the other hand, would then be designated as 'storage areas', as their optimal accumulation is higher.

However, if the network would be assessed based on the production, one would reach the conclusion that traffic should be controlled in the subnetworks instead, as the production (total amount of distance that can be travelled) is far higher than in the arterials.
Now, to make the example even more complex, let us take a look at the performance and production of the different subnetworks and their perimeters. When looking at Figure 3.3, one finds that the performance of both subnetworks is almost equal. From this, one could easily conclude that both subnetworks are equal and that none of these subnetworks should be treated any different from the other. However, when looking at the production, it is directly found that the first subnetwork and perimeter (which are 70% larger) have a much higher production and that it would be more beneficial to the over-all network performance to redirect vehicles to the first subnetwork.

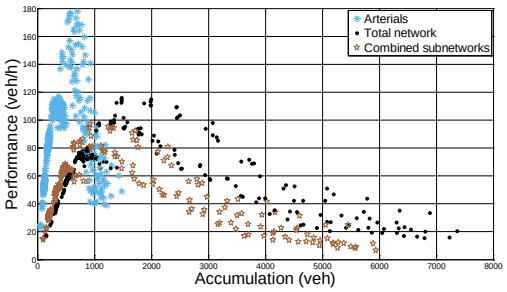


Figure 3.1: *Performance vs. Accumulation total network*

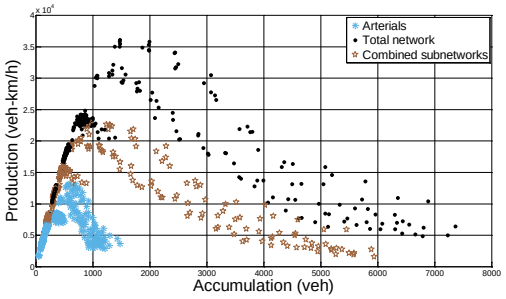


Figure 3.2: *Production vs. Accumulation total network*

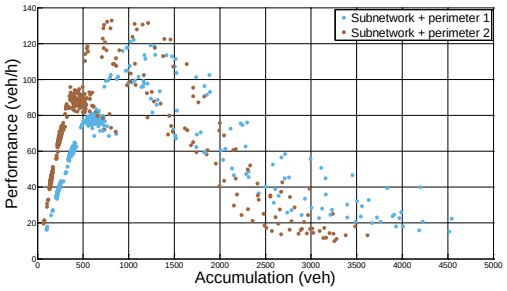


Figure 3.3: *Performance vs. Accumulation subnetworks and perimeters*

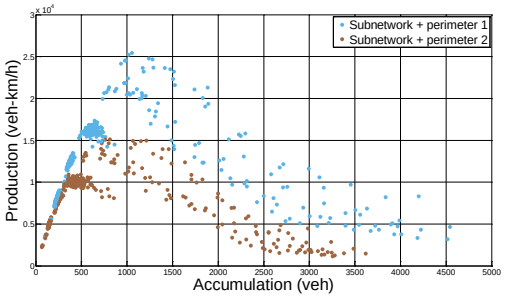


Figure 3.4: *Production vs. Accumulation subnetworks and perimeters*

What this example illustrates, is that evaluating a network solely on the production or the performance is not as straight-forward as might be expected, as they both give an entirely different view on the network. The MFD for the production is strongly related to the infrastructure, while the MFD for the performance is related to the efficiency at which the traffic is processed. It therefore seems, that if one wants to make a full assessment of the effect of changes to the network, both the effect on the production and performance should be evaluated.

Nevertheless, it can be concluded that when the accumulation in the network is sub-optimal, the traffic should be directed to that part, which has the highest performance, as traffic is processed in a more efficient way. If the accumulation in the network transgresses the optimal accumulation and spill-back occurs, traffic should be redirected to the parts of the network which have the highest production.

3.1.3 Accumulation and network density

Accumulation

On the x-axis of the MFD, one normally finds the accumulation, or the network density. The production PD is calculated by multiplying the density k over the length of the part of the network it represents l, for a given period of time Δt and summing this for all links (or measurements) i in the network Z, resulting in

$$A_{\Delta t} = \sum_{i \in Z} k_{i,\Delta t} l_i. \tag{3.3}$$

As $k_{i,\Delta t} = \frac{q_{i,\Delta t}}{u_{i,\Delta t}}$, the accumulation can also be written as

$$A_{\Delta t} = \sum_{i \in Z} \frac{q_{i,\Delta t} l_i}{u_{i,\Delta t}}, \tag{3.4}$$

making the accumulation dependent on the average speed. The average speed can in turn be expressed as a function of the average travel time tt and length of the considered part of the network, so that $u_{i,\Delta t} = \frac{l_i}{tt_{i,\Delta t}}$. This simplifies (3.4) to

$$A_{\Delta t} = \sum_{i \in Z} q_{i,\Delta t} tt_{i,\Delta t}. \tag{3.5}$$

The average travel time over a part of the network is dependent on a number of factors. One of the most well-known models for travel times is the BPR-function. Although it is highly simplified and can only be applied to a single link, it does show a number of factors influencing travel time. The BPR-function is defined as

$$tt_{i,\Delta t}(q_{i,\Delta t}) = \frac{l_i}{u_f}\left(1 + \alpha_{BPR}\left[\frac{q_{i,\Delta t}}{C_i}\right]^{\beta_{BPR}}\right). \tag{3.6}$$

From equation (3.6) it is found that the travel time depends on the free flow speed u_f, the link capacity C_i and the current traffic flow $q_{i,\Delta t}$. Apart from that, if a link ends at a traffic light, the travel time increases with the average waiting time at that signal $t_{wait,sig}$.

The above shows that the accumulation seems to depend on a number of static factors, such as the free flow speed, network length, capacity and average waiting time (in case the signal timings are fixed), and a number of dynamic factors, such as the traffic flow and the signal timings (in case the signals are demand responsive). It is worth mentioning that in equation (3.6), the free flow speed, link length and capacity are assumed to be static. However, this is not the case for networks which incorporate rush hour lanes, or adaptive speed strategies. Meaning that in order to obtain a uniform MFD, data must be derived over a period of time, in which the network is operated in the same manner.

Average network density

Instead of the accumulation, the average network density, often simply referred to as 'network density', is also commonly used. As with the performance, this takes the length of the network out of the equation, making it possible to compare the MFDs of different networks to each other. Apart from that, when using the performance on the y-axis, the units on both axes are equal to the units used in the well-known fundamental diagram. Showing the same relationship between density and intensity, but on a macroscopic level instead. The average network density K is expressed as

$$K_{\Delta t} = \frac{\sum_{i \in Z} k_{i,\Delta t} l_i}{\sum_{i \in Z} l_i}. \tag{3.7}$$

Like the performance, the average network density can also be calculated by taking the mean of all density measurements, which is called the unweighted average network density. This too is often done when working with data from real networks.

Accumulation vs. average network density

Analogous to the production and the performance, using either the accumulation or the average network density in comparing networks is not straight-forward. Again the aggregated value (in this case the accumulation) gives us insight in the (storage) capacity of the network, whereas the weighted value (the average density) shows the efficiency with which the traffic is processed.
Due to this analogy, it should go without saying that a full assessment of a network cannot be made by using either the accumulation or average network density alone and fully depends on that which the researcher is investigated in. On the one hand, if one wants to employ a specific control strategy, such as perimeter control, it is important to know the number of vehicles at which the production, or performance is maximised. On the other hand, if one wishes to quantify the effect of a certain traffic management system (such as a toll pricing scheme), employed in two different networks to each other, it could be better to use the average network density, as this scales the values the MFD of both networks to roughly the same size.
Generally speaking, it could be stated that production and accumulation can be used to assess the quantitative improvement of a network, whereas the performance and average network density can be used to assess the qualitative improvement of a network.

3.1.4 The effect of using different units

Figure 3.5 to Figure 3.8 show the effect on the MFD using the exact same dataset, but have different units on either axes. The first two figures are equal to Figure 3.1 and Figure 3.2 and show the difference between production and performance in relation to the accumulation. The last two figures show the effect when the average network density is used instead of the accumulation.

The first thing that becomes immediately clear when looking at the different types of MFDs, is that the shape of the MFDs, especially in relation to each other, differs heavily when changing the units in which it is expressed. It should however be noted that the actual shape of the MFD does not really change, as the relation between production/performance and accumulation/network density remains the same in all situations. The differences found in these MFDs are caused by scaling the MFDs as a consequence of taking the length of the different parts of the network out of the equation.
This effect is best shown by the MFD of the arterial. When taking the performance and network density into account, as is done in Figure 3.8, it can be concluded that the arterial performs better than the subnetworks, because a higher performance (at a higher network density) is attained, before spill-back occurs. Although the other figures seem to imply this is not the case, this is actually a valid conclusion. However, to derive from this conclusion that vehicles should be redirected to the arterials as much as possible is not valid, as it shows that the optimal accumulation is only half that of the subnetworks.

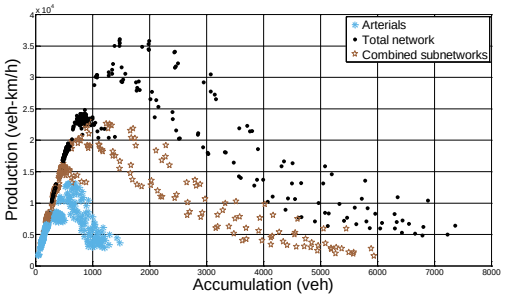


Figure 3.5: *Production vs. Accumulation*

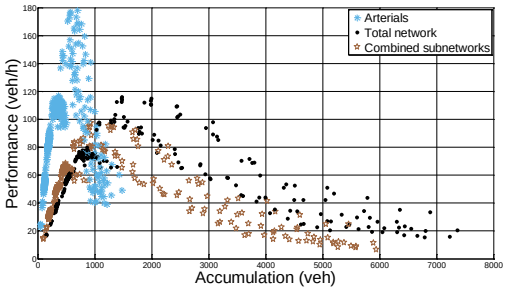


Figure 3.6: *Performance vs. Accumulation*

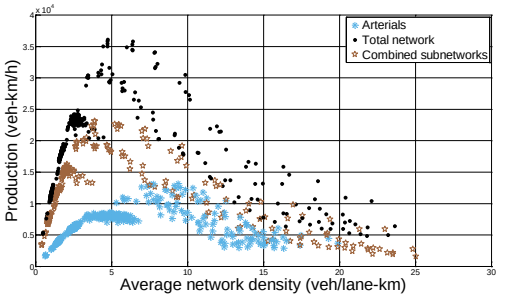


Figure 3.7: *Production vs. Average network density*

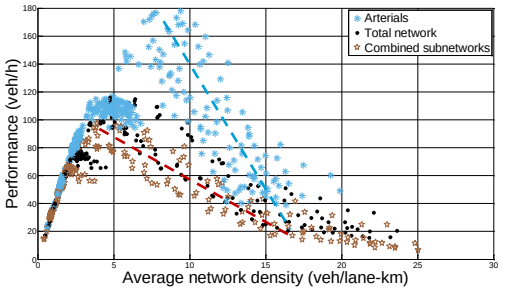


Figure 3.8: *Performance vs. Average network density (lines show backward wave speed)*

As in the previous example, the most valid control strategy in this case, is that the arterial should be kept close to its optimal accumulation, as it can process traffic more efficient than the subnetworks. In case the number of vehicles in the arterial should transgress this optimum, traffic should be stored in the subnetworks, as spill-back sets in at a much higher accumulation, due to the fact that the production (the amount of kilometres that vehicles can drive) is much higher in the subnetwork, than in the arterial. Looking at the different figures above, this conclusion cannot be reached directly from a single MFD, but needs the interpretation of a number of them. The only MFD coming close is the performance and accumulation, shown in Figure 3.6. However, this does not imply that this is the best type of MFD and should therefore always be used, because as has been explained before, this choice simply depends on the information one is interested in, or the actual data that is available.

For example, when working with data derived from real networks, the unweighted performance and unweighted network density are often used, because it is mostly not clear which part of the network each measurement represents. Also the production and accumulation are very hard to obtain for real networks, as large parts of these networks, such as the residential areas, are not covered by detectors. Nevertheless, it should be stressed that one should be cautious by drawing any conclusions from this diagram alone.

3.1.5 Average network speed, frequency and kinematic viscosity

Average network speed

In traffic flow theory, the relation between flow q, density k and speed u, $q = ku$ is one of the most well-known and most used relations. In the q-k fundamental diagram, this relation can also be found, as the angle between the flow and density shows the speed at that point. A schematic representation is given in Figure 3.9.

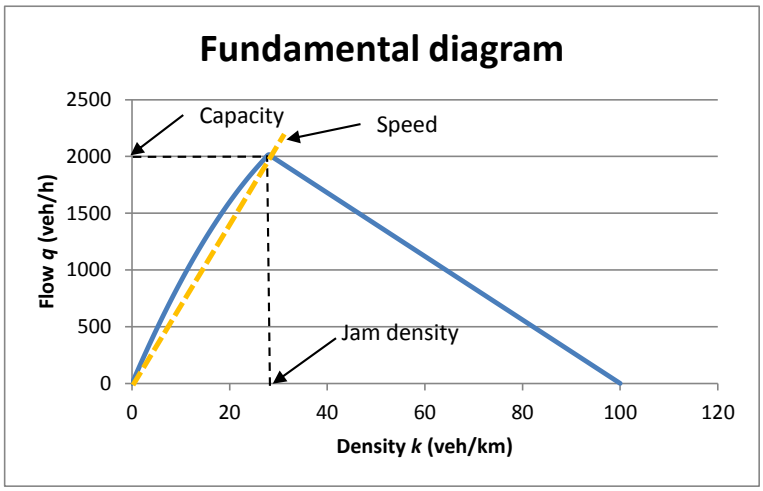


Figure 3.9: *Schematisation of a fundamental diagram relating flow and density*

The same relation between production and accumulation, or performance and network density should apply, i.e. the angle of the line should show the average network speed, as

$$U = \frac{PD}{A} = \frac{\frac{veh-km}{h}}{veh} = \frac{km}{h} \text{ and} \tag{3.8}$$

$$U = \frac{PF}{K} = \frac{\frac{veh}{h}}{\frac{veh}{km}} = \frac{km}{h}. \tag{3.9}$$

When turning to Figure 3.5 and Figure 3.8, it can be found that the average free flow speed for the subnetworks and the arterial is almost equal, at 60 / 2 = 30 km/h (derived from Figure 3.5). However, it can also be found that the arterial can keep up this speed at higher densities than the subnetwork, showing that it can maintain free flow over a larger density range than the subnetworks.
Nevertheless, this does illustrate that the free flow speed does have an impact on the shape of the MFD, as was conjectured in paragraph 3.1.3.

Apart from calculating the average free flow speed in the network, it should be possible to calculate the average network backward wave speed Ω, analogous to the fundamental diagram. Using the

tangent between two points in the MFD on the congested branch, the speed at which shock waves Ω move throughout the network can be calculated[10] as

$$\Omega = \frac{PD_2 - PD_1}{A_2 - A_1} = \frac{PF_2 - PF_1}{K_2 - K_1}. \tag{3.10}$$

Using a rough interpolation of the values found in Figure 3.8, we find that the speed at which shockwaves move through the arterials is approximately $\Omega = \frac{50-140}{15-10} = -18\ km/h$ and for the subnetworks $\Omega = \frac{30-90}{15-5} = -6\ km/h$. This also seems to be confirmed by Figure 3.5, which shows that the production of the arterial decays much more rapidly than the subnetworks, when the optimal accumulation is transgressed.

Frequency

If the same relation would be applied to the MFD relating the performance to the accumulation, we end up with

$$\frac{PF}{A} = \frac{veh/h}{veh} = \frac{1}{h}. \tag{3.11}$$

Expressed in the SI, the resulting unit is s^{-1}, which is also known as frequency. Judging from Figure 3.6, the frequency of the arterials is higher than that of the subnetworks. For instance, at its maximum performance, the frequency of the arterial is $\approx \frac{180}{750}/3.600 = 66{,}7\mu Hz$ and for the subnetworks $\approx \frac{100}{1.000}/3.600 = 27{,}8\mu Hz$. The corresponding periods are $T = \frac{1}{f} = \frac{1}{66{,}7 \cdot 10^{-6}} = 15.000\ s = 4{,}17h$ for the arterial and $10{,}0h$ for the subnetworks.

In order to illustrate how the frequency and period should be interpreted, consider a looping road, with a total distance of 1 km. If a single vehicle would drive at a speed of 100 km/h along this loop, the performance would be 100 veh/h, as the vehicle would pass an arbitrary point along the road 100 times per hour. The accumulation is only 1 vehicle. The frequency of this system is then $f = \frac{PF}{A} = \frac{100}{1}/3.600 = 27{,}8mHz$, with a period of $T = \frac{1}{f} = \frac{1}{27{,}8 \cdot 10^{-3}} = 36s$, which is the time it takes for the vehicle to complete a single lap around the road.

So basically, the frequency of the network is the amount of time it would take a single vehicle at a given accumulation to drive over all the roads in the given network. Using equations (3.1), (3.2) and (3.10), it can be found that this period (or lap time) T is given by

$$f = \frac{PF}{A} = \frac{PD}{A}\frac{1}{\sum_{i \in Z} l_i} = \frac{U}{\sum_{i \in Z} l_i} \rightarrow T = \frac{1}{f} = \frac{\sum_{i \in Z} l_i}{U}, \tag{3.12}$$

[10] This should also explains why the MFD does not have a sharp transition between the congested and uncongested regime (as in the FD), but is more rounded. The reasoning behind this is that when the optimal accumulation is transgressed and congestion sets in, queues start to grow. As queues grow, the number of vehicles hitting congestion increases exponentially, as more roads become blocked, resulting in ever faster growing queues and additional shockwaves, which in turn results in an exponential decay of the network performance (given the assumption that every vehicle can be loaded onto the network).

which is the total network length divided by the average network speed. Whether the frequency, or the period is a usable entity to assess the performance of networks is not completely clear. Nevertheless, this value could prove useful to assess the travel times on ring roads around a city.

Kinematic viscosity or diffusivity

For the MFD relating the production and the average network density, one finds

$$\frac{PD}{K} = \frac{\frac{veh-km}{h}}{\frac{veh}{km}} = \frac{km^2}{h}, \tag{3.13}$$

Which in the SI is expressed as m^2s^{-1} and is called the 'kinematic viscosity', which is the ratio of the inertial force to the density of the fluid (Keith, 1971) and is denoted with ν. The viscosity is used to measure the resistance of a fluid which is being deformed by either shear stress or tensile stress. More commonly the viscosity is regarded as the thickness of a fluid and its resistance to flow, meaning that a fluid with a low viscosity moves more easily than with a high viscosity.
In terms of traffic flow theory, the comparison between the flow of traffic and fluids is not uncommon and even some traffic models incorporate the viscous behaviour of traffic flow, such as the viscous model of Kühne (1987). If the analogy between fluids and traffic should hold, this relation would basically show the ease with which the traffic can move through a certain network. Nevertheless, the physical meaning of viscosity in traffic flow is not fully understood, but is suspected to be related to the resistance of drivers to sharp changes in speed (Zhang, 2003).

When looking at Figure 3.7, it is found that the viscosity of the arterial is lower that the viscosity of the subnetworks. Based on the previous conjecture, this should imply that traffic flows more easily through the arterials than through the subnetworks. This conclusion does not seem to farfetched, as the arterials generally consist of longer links than the subnetworks and that intersections are controlled by signals, avoiding conflicts between different traffic streams. Especially the latter argument seems reasonable, as the intersections in subnetworks are not controlled. This results in more shocks in the traffic flow, as each vehicle needs to find an appropriate gap to cross an intersection, often causing it to wait a while before an opportunity is offered. As such, this could imply that the viscosity could also show the intensity of conflicts encountered in a network.

Apart from the viscosity, the unit m^2s^{-1} is also used for thermal diffusivity, which is a unit to measure the speed with which heat moves through a substance in relation to its density. The higher the diffusivity, the more rapid heat can travel through that substance. In our case, this would mean that the higher the ratio between the production and the average network density, the more easily traffic can move through the network. Interestingly, this seems to be the opposite of the viscosity, which implies that a lower value would mean that traffic can move more easily.
In the case of the arterial and the subnetworks, this would then mean that traffic would be able to move more easily through the subnetworks than through the arterials.

However, the angle of the MFD for the total network is steeper than that of either the arterial and subnetworks. If this relation should in some way show the ease of movement, than the total network should actually be between the arterial and subnetwork, as it seems not viable that it would be easier to move through the total network, than to move through any of its parts.

As the above line of reasoning is only based on assumptions, it is not appropriate to derive any conclusions from this relation. Nevertheless, this is an interesting topic that could be investigated further. Because if this relation would hold in some way, this unit of measurement could show if making any adaptations to the layout of a network, aimed to reduce the intensity of conflicts, (such as one-way streets, roundabouts, signalising intersections, et cetera) could actually improve the production of a network. As for this thesis, this unit will not be used any further, due to the fact that its actual meaning is unclear.

3.1.6 Scatter and deviation of density

It has been found in literature that the deviation in density of the different links throughout the network, is also a strong factor influencing the shape of the MFD and is regarded by some as *the* factor explaining the scatter found in most MFDs. As such the deviation of density is introduced on the z-axis as

$$S_{\Delta t} = \sigma_{i \in Z, \Delta t}. \tag{3.14}$$

To illustrate the influence of the deviation of density, consider a network consisting of a simple looping road, that can be entered and exited at both ends, in which the road between every entry and exit point is a link, resulting in a network with 2 links. Every link has the same length and is governed by the same fundamental diagram, as is shown in Figure 3.10.

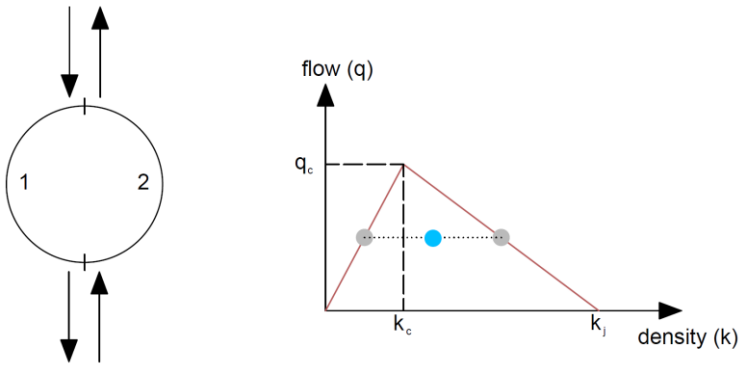


Figure 3.10: *Simple network and fundamental diagram governing network links*

By regulating the inflow and outflow at every entry and exit point, the network can be theoretically controlled in such a way, that a steady state, with any arbitrary density for each link can be achieved.

Assume that the inflow and outflow for every point is equal, making the density for both links equal. As the flow is drawn from the same fundamental diagram, the resulting average flow, for

any average density will be exactly equal to the shape of the original fundamental diagram. As the density is equal for each link, the deviation of the density is zero.
As a second case, assume that the density on the first link is zero and on the second link is equal to the critical density k_c. The resulting performance is now half the maximum capacity q_c, with an average density of half the critical density . This means that the resulting point in the MFD is also found along the same line. So, regardless of the number of links in the network, as long as all links are either congested or uncongested, the shape of the MFD is linear.
As a third case, assume that the inflow and outflow are regulated in such a way that the density of link 1 is half the critical density and that the density of link 2 is halfway between the critical density k_c and the jam density k_j, as shown in Figure 3.10. As in the previous case, the resulting flow is equal for both links. However, the resulting point on the MFD will end up below the original fundamental diagram (the blue point) and also the deviation is now larger than zero. From this, it should go without saying, that the larger the deviation of density is, i.e. both points on the FD are further apart, that the total flow would become lower.

Apart from explaining the scatter in the MFD, which is caused due to the fact that individual link data is derived from the different branches of the FD, the example also illustrates that the production of performance of a network is not directly linked to the deviation of density. This is because the same flow is found, at different deviations of the density. And as the deviation density is related to the accumulation or network density, one should expect a strong relationship between these parameters. Figure 3.12 clearly illustrates this, as the resulting relationship between the network density and the deviation of density is almost linear. The relation between the performance and the deviation of density on the other hand is like the MFD a concave function and is roughly similar in shape. In Figure 3.13 these plots have been combined, showing how the deviation of density changes as the average density increases. In this diagram, it is clearly visible that the change in colour evolves almost linear with the average network density over the complete performance range.

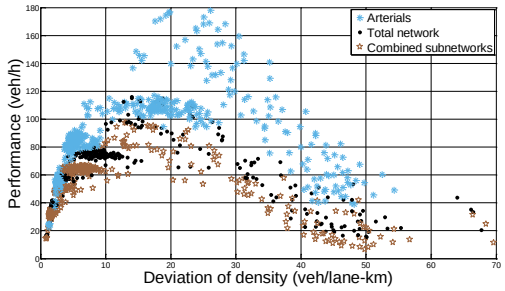


Figure 3.11: *Performance vs. Deviation of density*

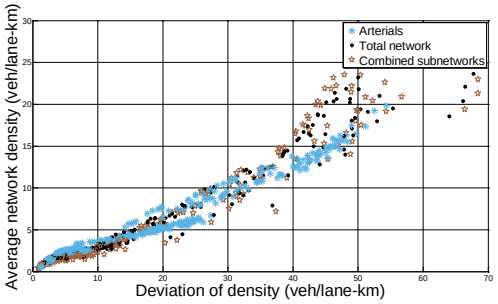


Figure 3.12: *Average network density vs. Deviation of density*

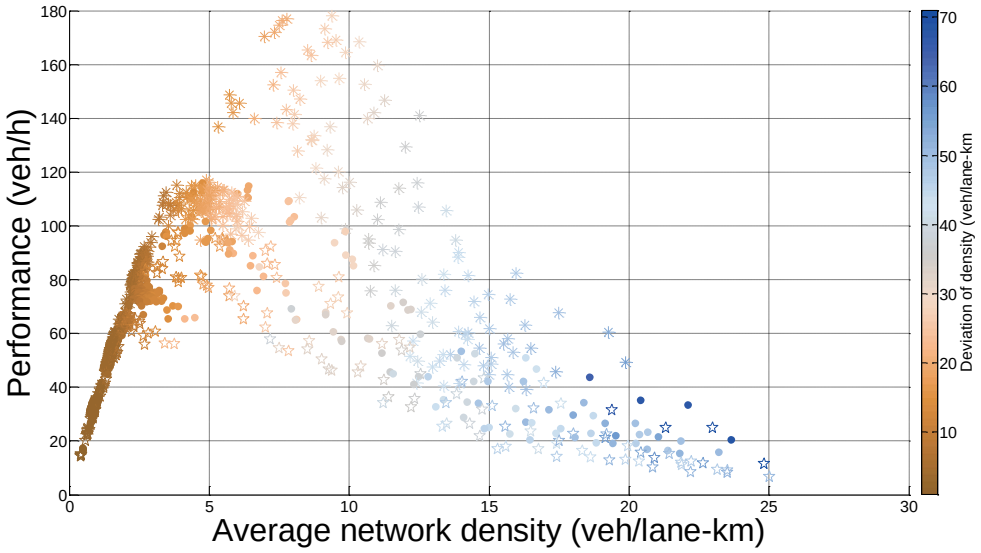


Figure 3.13: *Performance vs. Average network density vs Deviation of density*

In this paragraph and in the example it has been shown that the deviation of the density is an important factor in explaining the scatter found in the MFD and that this deviation is almost linearly related to the accumulation or network density. The maximum production of performance can therefore only be reached if all links are at critical density.

Although the deviation of density is an important factor in explaining the MFD, the evaluative value is somewhat less, because the range of scatter already gives an indication of this deviation. Apart from that, when relating the deviation of density to the performance, no significant differences were found from the relation between the performance and the average network density. This implies that incorporating the deviation of density would not lead to a MFD which shows less scatter.

3.2 Factors influencing the shape of the MFD

3.2.1 Assessment of factors

In the previous section, the MFD has been subjected to a thorough investigation, in which the data found on either axes, was broken down to its bare parameters, in order to find the factors that make up the MFD and therefore influence its shape.

The most important factors that have been found to influence the shape of the MFD are:

- Length of the network and underlying links;
- Capacity of the links;
- Free flow speed;
- Average waiting time at signals/signal settings;
- Spatial distribution of traffic over the network.

To assess how the MFD should theoretically react to changes in each of these factors, let us repeat the formulations found earlier.

Production: $$PD_{\Delta t} = \sum_{i \in Z} q_{i,\Delta t} l_i, \quad (3.15)$$

Performance: $$PF_{\Delta t} = \frac{\sum_{i \in Z} q_{i,\Delta t} l_i}{\sum_{i \in Z} l_i}, \quad (3.16)$$

Accumulation: $$A_{\Delta t} = \sum_{i \in Z} k_{i,\Delta t} l_i, \quad (3.17)$$

Average network density: $$K_{\Delta t} = \frac{\sum_{i \in Z} k_{i,\Delta t} l_i}{\sum_{i \in Z} l_i}. \quad (3.18)$$

For which the accumulation and average network density were broken down further to

Accumulation: $$A_{\Delta t} = \sum_{i \in Z} q_{i,\Delta t} tt_{i,\Delta t}, \quad (3.19)$$

Average network density: $$K_{\Delta t} = \frac{\sum_{i \in Z} q_{i,\Delta t} tt_{i,\Delta t}}{\sum_{i \in Z} l_i}, \quad (3.20)$$

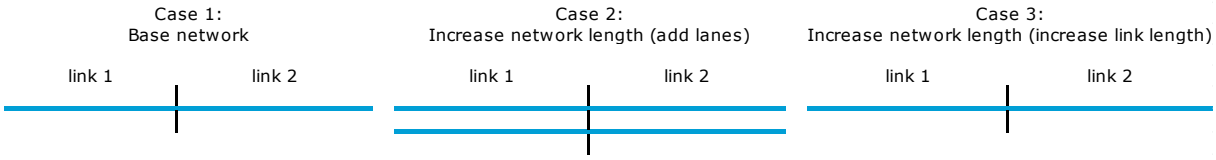
with $$tt_{i,\Delta t}(q_{i,\Delta t}) = \frac{l_i}{u_f}\left(1 + \alpha_{BPR}\left[\frac{q_{i,\Delta t}}{C_i}\right]^{\beta_{BPR}}\right) + tw_i. \quad (3.21)$$

3.2.2 The effect of changing network length

Changing the length of the network should impact the MFD in multiple ways, depending on the way these changes are made. Assuming the flow is unaffected (this could for instance happen if additional lanes are added to a link), it should be expected that the production of the total network should remain the same, because the same flow is divided equally over the lanes.
Consider the simple network shown in Figure 3.14, consisting of 2 links, both having only a single lane. In the second case, an additional lane is added and the flow is divided equally over both lanes.

The example shows that the production and the accumulation remain unchanged after an additional is added, which is logical because $PD_{\Delta t} = \sum_{i \in Z} \frac{q_{i,\Delta t}}{n_{lanes,i}} l_i n_{lanes,i}$ and $A_{\Delta t} = \sum_{i \in Z} \frac{k_{i,\Delta t}}{n_{lanes,i}} l_i n_{lanes,i} = \sum_{i \in Z} \frac{q_{i,\Delta t}/u_{i,\Delta t}}{n_{lanes,i}} l_i n_{lanes,i}$, showing that the number of lanes is factored out of the equation and both the production and accumulation depend on the original flow. The performance and network density on the other hand are halved, as they both are divided by the network length, which has doubled.

Now assume that instead of adding additional lanes, the network length is increased, by increasing the length of the links itself. As can be seen in case 3 in Figure 3.14, the production and accumulation have doubled, whereas the performance and network density have remained the same. Apart from the effect this has on the MFD, this also implies that if the average travel distance between an origin and destination is increased or decreased, the shape of the PD-A MFD changes, making the MFD dependent on the OD-patterns as well.
However, if we assume that a flow of 500 veh/h is equal to the capacity of each lane and that this capacity is fully utilised after additional lanes have been added, than all the values will become the same as when the link length is increased.



q = 500 veh/h/lane	q = 500 veh/h/lane	q = 250 veh/h/lane	q = 250 veh/h/lane	q = 500 veh/h/lane	q = 500 veh/h/lane
l = 2 km	l = 1 km	l = 2 km	l = 1 km	l = 4 km	l = 2 km
u = 50 km/h	u = 50 km/h	u = 50 km/h	u = 50 km/h	u = 50 km/h	u = 50 km/h
n_lane = 1	n_lane = 1	n_lane = 2	n_lane = 2	n_lane = 1	n_lane = 1
Production	1.500	Production	1.500	Production	3.000
Performance	500	Performance	250	Performance	500
Accumulation	30	Accumulation	30	Accumulation	60
Network density	10	Network density	5	Network density	10
Average speed	50	Average speed	50	Average speed	50
Frequency	17	Frequency	8	Frequency	8

Figure 3.14: *Effect of increasing network length by adding additional lanes or lengthening links*

What this example shows, is that adding additional infrastructure, influences the shape of the MFD in different ways, depending on how these additional roads are added and which type of MFD is used. If for instance the effect would be measured directly after adding additional lanes, then there would be no effect on the PD-A MFD, while the PF-K MFD would imply that the network would perform even worse, as the performance and network density would have been halved, which seems to be completely opposed to what one would actually expect. Apart from that, even when the new infrastructure would be completely utilised, the PF-K MFD would show that the network has not improved in any way. Only when the production or the accumulation would be used, one can witness an improvement in the network.

Although an improvement in the network can be found using the right relation, the most troubling thing is that none of these relations actually show that the average travel time has actually increased in the second case, causing the total network to actually perform worse than it did before. Even more bothersome would be the fact that if the total link length would be halved (which also halves the average travel time), this would not show in any MFD, and actually only a decrease in the production and accumulation would be witnessed.

The only thing that could hint to an actual deterioration of the network performance is the frequency. The frequency has decreased, which shows that a vehicle that would travel over the complete network, will only be seen half the times than it would have been before, i.e. its period $T = 1/f$ has doubled. But although this is true for the third case, this also implies that the travel time in the second case has increased, even though it actually remained the same.

All in all, it is assumed that changing the total length of the network will not result in any change to the shape of the PF-K MFD, as the total network length is factored out in the equations making up the diagram. The only case in which a difference might be witnessed, is if the traffic volumes are kept equal in both cases. However, when increasing the traffic volumes to the capacity of the network, no substantial changes should be witnessed.

As for the PD-A MFD, a linear increase to both the production and the accumulation is expected. This implies that when the PF-K MFD for two networks are roughly equal, that the PD-A MFDs should have the exact same shape, albeit at a different scale. From this it is conjectured that the length of the network only scales the MFD over both axes and does not affect its shape. Therefore, the PF-K MFD should be able to compare the impact of factors, other than the network length, on different network types to each other.

3.2.3 The effect of changing the link capacity

If the capacity of the links within the network can be raised (for example by using traffic management, reducing conflicts, wider lanes, et cetera), it is expected that the MFD is exponentially shortened over its x-axis. This assumption is made based on equation (3.21), which shows that an increase in the capacity of the links leads to a reduction of the q/C-ratio and therefore lowers the travel time. Due to the fact that the scaling parameter β_{BPR} is larger than 1 (otherwise travel times would decrease when the q/C-ratio increases), the travel time decreases when the capacity is increased. But because the total travel time is also dependent on the free flow travel time (which remains equal), the relative reduction in travel time increases exponentially, as the influence of the free flow travel time becomes less when the flow increases. This means that the reduction of the accumulation/network density should become stronger at a higher production/ performance. For the free flow branch, this effect is assumed to be minimal.

Apart from the accumulation, the production/performance should be affected as well. Because of the increase of capacity, the point at which congestion sets in is at a higher production or performance. Also due to the increase of the capacity, the average speed can be maintained at higher accumulations. It is therefore expected that the free flow branch of the MFD remains linear over a larger production or performance range.

It is expected that the optimal accumulation should shift to the right due to the fact that the production is assumed to increase and the angle of the free flow branch will roughly remain the same. Nevertheless, as the jam density remains the same for all links, the maximum accumulation should not change, meaning that the MFD is fixed on both of its endpoints.

Now, if the capacity increases, which is assumed to also increases the maximum production and the free flow speed remains unchanged, then it should go without saying that the slope of the congested branch should also be much steeper. This then results in a sharper transition from the free flow state to the congested state, as is shown in Figure 3.15. It is expected that this effect should be the same regardless of the type of MFD that is used.

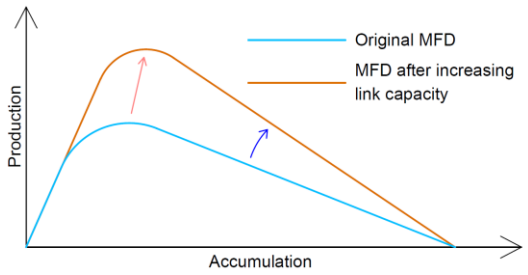


Figure 3.15: *Assumed effect on the MFD of changing the link capacity*

3.2.4 The effect of free flow speed on the MFD

Changing the free flow speed, i.e. adapting the maximum allowed speed, should affect both the production and the accumulation of the MFD. Regarding the accumulation, equation (3.21) shows that increasing the speed shortens travel times and therefore causes a reduction in the accumulation, in case flows are kept equal. On the production side, it is expected that the production and the performance will increase, as an increase in speed often results in an increase of the maximum capacity (up to ≈ 90 km/h).

Apart from the above, it has also been shown that the angle of the linear branch of the PD-A, or PF-K MFD, shows the average network speed. So as the speed of each link is increased, the average network speed should increase as well, resulting in a steeper angle of the linear branch of the MFD.

However, as with the capacity before, the maximum accumulation should not change, as this value is assumed to be bounded by the jam density, i.e. the number of vehicles that maximally fits into the network. In this case it is not completely clear to which side the optimal accumulation will shift. If it is assumed that the link capacity is linearly related to the speed, than the critical density will not change, meaning that the optimal accumulation will remain unchanged as well. However at higher speeds, it is often found that the capacity decreases, causing a reduction in the critical density, lowering the optimal accumulation.

As this thesis will mainly investigate urban networks, operating at lower speeds, it is assumed that the link capacity is linear to the speed. This then implies that the change in the maximum production should also be linear to the change in the maximum speed.

A rough interpretation of the effect changing the free flow speed has on the MFD is shown in Figure 3.16. It is expected that this effect should be the same regardless of the type of MFD that is used.

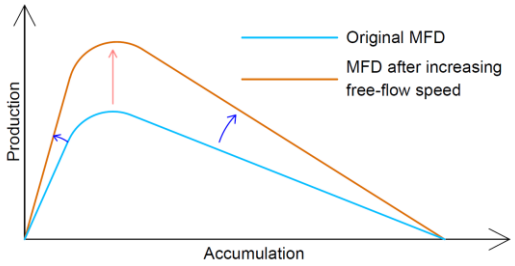


Figure 3.16: *Assumed effect on the MFD of changing the free flow speed*

3.2.5 The effect of signal settings on the MFD

General

One of the factors that has been found in literature to have a strong effect on the shape of the MFD, is the timing of signals at the intersections in the network and the offset between different signals. A thorough study of the effect of different signal timings has been done by Boyacı and Geroliminis (2010), which showed that the performance of a network could be seriously affected in case signal settings and offsets are chosen incorrectly. The reasoning behind this is that if vehicles are given green at the first intersection, but arrive at the next intersection while the signal is red, the flow is severely restricted by this, even though the second intersection might offer lots of green time. As such, the bottleneck capacity is a combination of signal timings, signal offsets, traffic flow, link length, and driving speed (see Figure 2.9).

This means that the effect of different signal settings on the shape of the MFD cannot be directly quantified by taking the signal timings of the different intersections, as the bottleneck capacity is an intricate combination of many factors. Apart from that, the fact that turn rates and signal timings differ from intersection to intersection, it seems to be almost impossible to exactly calculate this effect.
However, if the capacity of each intersection is made redundant enough, the effect of spill-back can be mitigated, as the intersection should be able to serve all traffic. Even if vehicles should always arrive during a red phase, the incoming links can be emptied completely during the next green phase, which causes each intersection to operate independent of each other intersection.
If every intersection would operate independent of every other intersection, this would imply that the capacity of every link depends on the maximum outflow of the first downstream bottleneck, which holds for either controlled and uncontrolled intersections. The maximum production therefore becomes

$$PD_{\Delta t,c} = \sum_{i \in Z} q_{i,\Delta t,c}\, l_i. \tag{3.22}$$

Now, it should go without saying that linking every single link in the network to the corresponding bottleneck seems to be impossible. Nevertheless, this formulation does imply that if the maximum outflow rate at intersections is increased, e.g. by changing signal timings or adding lanes, that the production and the performance should increase.

The rush and its super state

In paragraph 1.3.2 it has already been shown that the total outflow of a subnetwork can never be larger than the summed capacity of all the links connecting that subnetwork to the rest of the network.
Interestingly enough, the maximum production or performance in such a network, can become larger than the bottleneck production or performance.

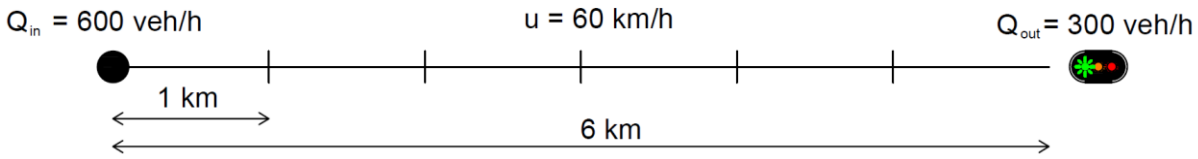


Figure 3.17: *Sample network showing the effect of early rush*

In Figure 3.17 a simple linear network is shown, consisting of six link, with a length of 1 km each. The network is fed with $Q_{in} = 600\, veh/h$ at one end and is controlled by a traffic signal with an outflow of $Q_{out} = 300\, veh/h$. The speed in the network is chosen at 60 km/h, making the travel time per link 1 minute. The maximum density of each link is assumed to be 10 veh/km.
At t_0 the network is considered to be empty, and production and performance are 0. After one minute, 10 vehicles have been loaded onto the network, which are all in the first link. The measured flow in link 1 is 600 veh/h and 0 veh/h for the five other links, making the production $PD_{\Delta t} = \sum_{i \in Z} q_{i,\Delta t} l_i = 600\, vehkm/h$ and the performance $PF_{\Delta t} = \frac{\sum_{i \in Z} q_{i,\Delta t} l_i}{\sum_i l_i} = \frac{PD_{\Delta t}}{\sum_i l_i} = \frac{600}{6} = 100\, veh/h.$

The accumulation in the network is $A_{\Delta t} = \sum_{i \in Z} k_{i,\Delta t} l_i = \sum_{i \in Z} \frac{q_{i,\Delta t} l_i}{u_{i,\Delta t}} = 1 \cdot \frac{600 \cdot 1}{60} = 10veh$ and the average density is $K_{\Delta t} = \frac{\sum_{i \in Z} k_{i,\Delta t} l_i}{\sum_i l_i} = \frac{A_{\Delta t}}{\sum_i l_i} = \frac{10}{6} = 1{,}66$.

In Table 3.1, the values for the different MFD parameters for the first six minutes are given, as the network remains in free flow over this period.

Table 3.1: *MFD parameters for example network showing early rush*

Time (min)	Production (veh-km/h)	Performance (veh/h)	Accumulation (veh)	Network density (veh/km)
Free flow				
1	600	100	10	1,66
2	1.200	200	20	3,33
3	1.800	300	30	5,00
4	2.400	400	40	6,66
5	3.000	500	50	8,33
6	3.600	600	60	10,00
Congestion				
7	1.800	300	60	10,00

As can be seen from the table, the MFD for the first six minutes is simply linear and the maximum production is 3.600 veh-km/h, at an accumulation of 60 vehicles.

Now, consider what happens in the 7th minute. The network is fully loaded and only 300 veh/h, i.e. 5 veh/min can exit the network. The consequence of this is that the maximum flow on each link in the network becomes 300 veh/h for the rest of the simulation period, resulting in a maximum production of 1.800 veh/h, as can be seen in Table 3.1.

This means that the MFD shows a significant drop in production and performance at the same accumulation/density. Although the statement that the maximum production is 3.600 veh-km/h and that the optimal accumulation therefore is 60 vehicles is not untrue, this is actually not the most optimal state for this network. From the above, one can derive that the most optimal accumulation would be 30 vehicles, as this maintains the network in its free flow state. Production and performance however, are not maximised in this way.

Although not specifically shown, the above 'problem' has already been identified, as one of the regularity conditions to obtain a well-defined, low scatter MFD is "a slow-varying and distributed demand". And indeed it is true that if the network would have been loaded more homogeneously, that this 'super state' would not have occurred. However, this effect cannot be disregarded, as real networks often experience this loading pattern during the morning peak period, in which the initial network density is often low. As this effect can be related to rush hour periods, we will refer to this phenomenon as 'the rush'.

The super states created during the rush actually reside outside the boundary of the 'steady-state' MFD, which is caused by the fact that the production is derived from the maximum capacity of the individual links and not from the bottleneck capacity.

A further complication to this problem is that the severity of this effect is related to the number of vehicles loaded onto the network. For example, if $Q_{in} = 400veh/h$, than the maximum production would become 2.400 veh-km/h, at an accumulation of 40 vehicles, changing the actual shape of the MFD. Assuming that (i) all links in a network are capable of meeting the traffic demand, (ii) traffic distributes itself over all links in the network and (iii) traffic from all origins r are loaded uniformly onto the network, than the production PD_c and accumulation A_c at capacity, in that network should become equal to

$$PD_c = \sum_{r \in Z} q_r \sum_{i \in Z} l_i, \tag{3.23}$$

$$A_c = \sum_{i \in Z} k_i l_i = \sum_{r \in Z} q_r \sum_{i \in Z} \frac{l_i}{u_i}, \tag{3.24}$$

which shows that the shape of the MFD is not invariant to the number of trips made in a heterogeneously loaded networks, making it highly unlikely that a well-defined MFD can be obtained for such a network. This example also shows that differently shaped MFDs can be obtained, even if the network is controlled in the same way, making it highly complex to directly relate signal settings to the shape of the MFD.

However, the biggest implication of 'the rush' is that it results in an inaccurate MFD, incorrectly representing the network, because these states cannot be sustained under any circumstances. The 'steady-state' maximum production and accumulation of a network, i.e. the production and accumulation which can be sustained over a longer period of time, cannot be accurately derived, as they are obscured by inaccurate points in the MFD.
Another problem with this is that basing a control strategy on the optimal accumulation, will almost always fail, as traffic should end up in congestion after a short while.

Conclusion

Regarding the accumulation and the network density in general, these are expected to increase as the average waiting time at signals increases, because of equation (3.21). The increase in the accumulation is assumed to be non-linear with the increase of the average waiting time, as the travel time is composed of the travel time over the link and the waiting time. As the (free flow) travel time over the link should be unaffected by the presence of the traffic signal at its end, this part of the equation will remain the same. Apart from that, not every link is affected by a change in the signal settings and as such, these changes do not have an impact on the total network. Generally speaking, it is assumed that the accumulation slightly increases if the average waiting time at signals is elongated. Due to the restriction this places on the outflow, the capacity of the link should lower, resulting in a decreasing production.

In both the PF-K and the PD-A MFD, this should show as a shift of the maximal production and optimal accumulation to the lower left. The consequence of this is also that the angle of the free flow branch is lowered, meaning that the average network speed has decreased. A rough interpretation of the effect that changing the free flow speed has on the MFD is shown in Figure 3.18. As with the changes to the capacity and the free flow speed, the maximum accumulation in the network is not expected to change, meaning that the speed at which shockwaves move through-out the network is also lowered.

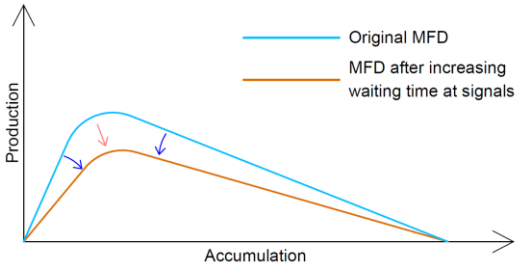


Figure 3.18: *Assumed effect on the MFD of changing the waiting time at signals*

3.2.6 The effect of changing the deviation of density

In paragraph 3.1.6, it has been shown that when the density of the links within the network differs significantly, especially when a part of the network is in free flow and another part is congested, or the shape of the fundamental diagram of the individual links in the network differs, scatter arises in the MFD. These scatter points, are almost always found below the optimal shape of the MFD. Traffic management strategies aimed at reducing the deviation of density (DoD), such as redirecting traffic or changing the signal timings, can therefore improve the performance of a network.

Changing the distribution of traffic over the network should theoretically not affect the accumulation, as the same amount of traffic would still be present in the network. However, redistributing traffic over the network more evenly can avoid pockets of congestion forming, which decreases travel times and result in lower accumulations at a higher performance. Nevertheless, as this redistribution does not physically change anything to the network (apart from changes to the signal timings in case of demand-adaptive signals), the actual shape of the MFD will not change, but will only become more pronounced. The reason for this is, that the outer envelope of the MFD is assumed to be determined by;

- The angle of the free flow branch, determined by the average free flow speed;
- The optimal accumulation, obtained if all the links are at critical density;
- The maximum production, attained when all links are at maximum capacity
- The maximum accumulation, found when the complete network is in gridlock i.e. all links are at jam density.

Changing the distribution of traffic over the network, does not affect any of the above mentioned parameters. The only thing that a reduction in the deviation of density does, is create less scatter and bring the MFD points closer to their outer envelope.

However, reducing the deviation of density does not always improve traffic flow. If for instance the total network is uncongested and assuming that all links are governed by the same triangular fundamental diagram, than the flow on every link is derived from the same linear relationship between flow and density (as has been shown in paragraph 3.1.6 and Figure 3.10). The resulting

shape of the MFD is than unaffected, even though traffic is not distributed evenly over the network. The same applies if all the links in the network are congested.
Basically, this means that even though MFDs are commonly highly scattered because traffic is distributed unevenly, their basic shape, i.e. the outer envelope of the MFD, can still be witnessed.

An example of this is given in Figure 3.19. This figure is created by randomly drawing the length $l_i = rand(1;5)km$ and free flow speed $u_f = rand(40;80)$ of 5 links. From this, the FD for each link is created, with jam density $k_j = 110\,veh/km$, critical density $k_c = rand(20;30)$ and maximum capacity $q_c = u_f k_c rand(0{,}8;1{,}2)$. Then 25.000 random densities are drawn for each link and the corresponding flow and deviation of density is calculated, resulting in the figure below.

Figure 3.19 shows that the basic shape of the MFD is still roughly triangular, even though the length and the FD of the underlying links are different. The maximum production found in the MFD is 21.117 veh-km/h, at an accumulation of 340 vehicles. Making the average network speed 62,1 km/h.

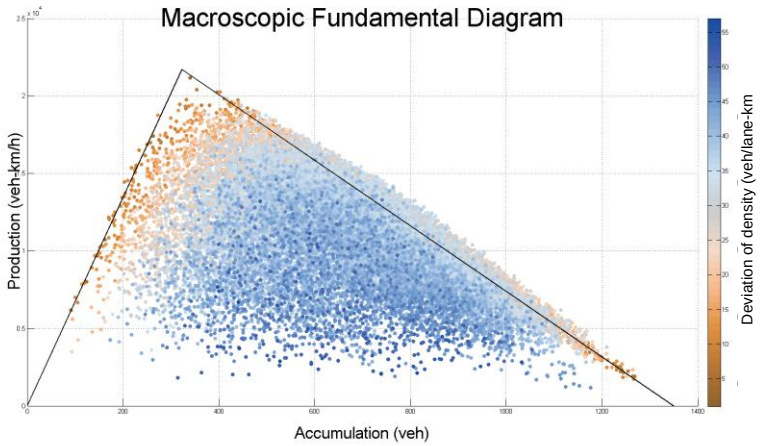


Figure 3.19: *Deviation of density vs. Production and Accumulation with envelope*

The actual sum of the link capacities and link lengths came out at 21.694 and the optimal accumulation, derived from the critical density is 323 vehicles, making the average network speed 67,2 km/h. Based on the total length of the links and the jam density, the maximum accumulation should be 1350 vehicles, which again is found in the graph.

Enveloping the MFD

The figure shows that most of the points of the MFD are found in a certain region, bounded by the average network speed, link capacity, critical density and jam density. Interestingly enough, a decent amount of points is found outside this region. On the free flow side this can be explained by the fact that the randomly generated density for the link with speeds lower than the average speed was far lower than for the other links, which had a higher speed, resulting in a point just outside the boundary. On the congested branch, the number of points outside the bounded area is far higher, and also an interesting bend is found at an accumulation of 750 vehicles. The same line of

reasoning can be applied to this part, with the difference that in this case, the culprit is the higher capacity of certain links, resulting in a higher lying branch in the FD.

From these findings it can be derived that the outer envelope of the MFD is not simply bounded by the sum of the aforementioned factors, but consists of a concatenation of the FD slopes of each of the links. On the free flow branch, these slopes should be concatenated in order of the highest speed, which sorts the steepness of the slopes of the FDs. For the congested branch, the links are sorted in order of the weakest descending congested branch of the FD, i.e. $q_c/(k_j - k_c)$. The MFD with the new envelope is shown in Figure 3.20, in which every point in the MFD is inside the envelope. Also the envelope follows the bend in the congested branch neatly. This then implies that if the critical density and capacity of every link in the network are known, that it should be possible to create the outer envelope of the MFD of any network. It should however be noted, that in reality the link capacity is not the maximum capacity of the link itself, but the maximum capacity of the downstream bottleneck by which it is actually governed, as has been shown in paragraph 3.2.5. This is not the case in this example, as the dynamics of traffic are not taken into account.

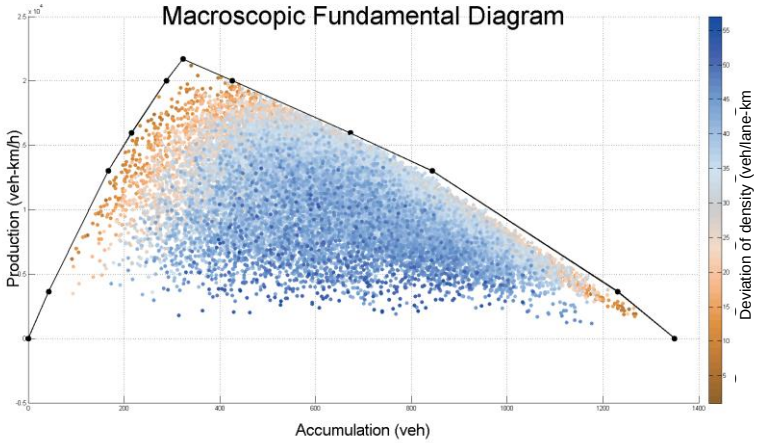


Figure 3.20: *Deviation of density vs. Production and Accumulation with improved envelope*

Hitting the wall

Another interesting effect, is that along the border of the MFD in the congested branch, a lot of points are found that have a very high deviation of density. The reason that these points are still along the border of the MFD, is that although the densities are far apart from each other, they are all in the congested region and draw from a roughly same shaped FD, resulting in a MFD-point, which is also found along this slope. The result of this, is that the exact same points on the MFD are found over a large range of the DoD. To illustrate this, Figure 3.21 shows a 3D version of the MFD, in which the height of the points is the deviation of density. The point of view is chosen in such a way, that one looks along the slope of the congested branch, with the point of maximum production closest to the viewer.

The figure shows that on the congested branch, there is a large range of the DoD resulting in the same production and accumulation, which is portrayed as, what can best be described as a 'wall'. Only when a certain DoD threshold is passed, the production and accumulation are lowered, i.e.

the MFD starts curving inward. What is also found from the figure (not shown) is that underneath the surface, there is a large empty area. This means that any combination of production and accumulation that does not reside along the border of the MFD, only exists within a specific (small) range of the DoD. This range should become smaller when more measurements are added.

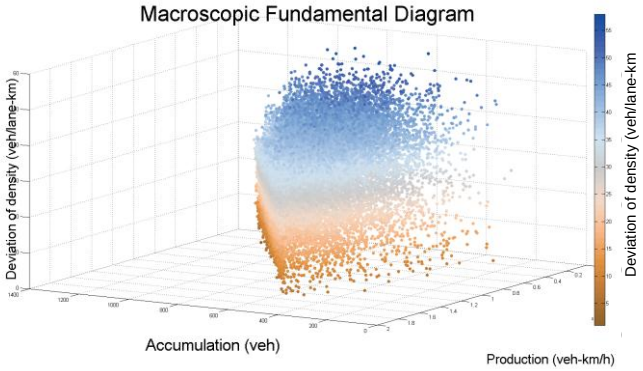


Figure 3.21: *Deviation of density vs. Production and Accumulation 3D*

Although fairly straightforward, it should be noted that although the production and accumulation are bound to a specific range of DoD, this does not work the other way around, i.e. the production and accumulation cannot be derived from the DoD.

Implications for traffic management

The above findings shed some light on the effect that the spatial distribution of traffic has on the MFD. One of the more interesting implications of these findings is that a traffic management strategy, aimed at improving the performance by reducing the deviation of density, should not always aim at reducing the deviation to zero. Instead, the production is optimised when the deviation of the critical density is zero. This then implies that in order to optimise the performance, vehicles should be redirected to the links having the highest critical density and/or capacity, i.e. the part of the network with the highest average free flow speed. This strategy however, becomes less effective when the optimal accumulation in the network is transgressed, as a wider range of the DoD would still result in the same production and accumulation. In this case, the strategy should aim to keep the DoD below the threshold value.

Conclusion

In regard to the shape of the MFD, it can be stated that changing the distribution of traffic over the network, in order to reduce the deviation of density, does not severely influence the shape of the MFD. It does however reduce the scatter and produce a sharper and better defined MFD. A general interpretation of how the MFD is affected is shown in Figure 3.22.

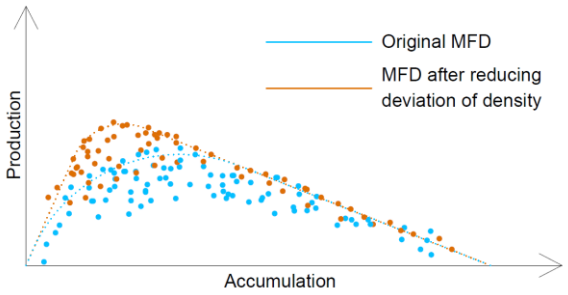


Figure 3.22: *Assumed effect on the MFD of reducing the deviation of density*

3.3 Conclusion

3.3.1 Factors impacting the MFD

In this chapter the MFD has been subjected to a thorough analysis, in order to find a number of factors that affect the MFD and how each of these factors influences the shape of the MFD. The most important factors that have been identified, including their effect on the MFD are:

- **Network length**: An increase of the length should result in a linear increase of both the production and the accumulation. The MFD relating performance and average network density is assumed to remain unchanged.
- **Link capacity**: Improving the capacity of the links is assumed to increase the production and the performance of the network. The optimal accumulation is assumed to increase as well. The maximum accumulation and network density however, should remain unchanged.
- **Free flow speed**: Changing the free flow speed of the links (increasing the speed limit, or improving the average flow), should result in a higher production and performance, as this should also increase the capacity. The accumulation and network density are assumed to remain unchanged. Because the angle of the linear branch in the PD-A and PF-K MFD is also the average network speed (as with the FD), this angle should become steeper as well. As the capacity of a link is often linearly related to the speed on that link (up to a certain speed), it is assumed that the production should grow linearly with the increase of the speed.
- **Signal settings** (average waiting time): Changing the signal timings should severely impact both sides of the MFD. Increasing the length of the red phase, leads to a prolonged waiting time at the signal, lowering the maximum capacity of that link, causing a decrease in the maximum production and performance. The optimal accumulation and network density are assumed to become higher, as longer waiting times result in longer travel times. The relative effect is assumed to be non-linear, e.g. doubling the waiting time should result in only a slight reduction of the production.
- **Distribution of traffic over the network** (deviation of density): Distributing the traffic more evenly over the network should not change the basic shape of the MFD, i.e. the outer envelope of the MFD, and only reduce the amount of scatter, resulting in a more well-defined shape.

3.3.2 The effect of using different types of MFD

In this chapter, the effect of using different types of MFDs (PD-A, PF-K, PD-K and PF-A) has been investigated as well. From this it is found that using a different relation results in a completely different MFD, in which every MFD tells something different and should be used in a different way. The PF-K MFD is one of the most commonly used and enables to investigate how well traffic is processed in the network and can be used to compare different traffic management strategies, or networks to each other. On the other hand, effects can only be qualified and not be quantified.

In order to quantify the effect of specific measures or changes to the network, the PD-A MFD is better suited, as it incorporates the actual network length. However, the drawback of this diagram is that it only shows how far and how much vehicles can travel through the network and does not show improvements in traffic flow. Even more so, if the average travel distance and travel time

should become larger, the production and accumulation increase, making the PD-A MFD an unreliable tool to make any statements about the quality of traffic flow in the network.

The PF-A and PD-K are not commonly used, as the relation between the axes does not show the same relationship as in the FD. However these MFDs are actually much better in showing the relative effect of different (sub)networks. Especially the PF-A MFD is interesting, as the steepness of the angle shows the frequency of the network, which is the inverse of the period and is the time needed to travel over every link in that network. Apart from that, it shows most of the information needed to compare different networks. On the one hand it shows which network performs better (the network that should be kept at maximum production) and on the other hand it shows the accumulation, providing information on the network most suitable to 'store' traffic, in case the accumulation in one of the networks becomes too high.

Whether or not the PD-K MFD is usable, is not completely clear. Although different networks can be compared to each other in both a quantitative and qualitative way, the meaning of the relation between the production and the network density is not fully understood (viscosity or diffusivity) making it currently unsuitable to use.

3.3.3 The rush and super states

Another important phenomenon that should be pointed out is the so-called 'rush' and the 'super states' this creates. The 'rush' is a period of time in which links remain in free flow and can reach their maximum capacity, before they become congested due to a downstream bottleneck, reducing the effective maximum capacity of these links. The result of this is that very high productions and performances are found over a short period, which cannot be sustained over a longer period of time. These points called 'super states', are therefore outside the boundary of the 'steady-state' MFD, leading to a MFD that is a not an accurate representation of the network.
The result of this is that the maximum production and optimal accumulation found in a MFD, cannot be directly utilised as control parameters, or to assess the quality of the network.

4 Research approach

Using the insights gained in the previous chapter, this chapter will present the general approach with which the research questions will be answered. To this end, section 4.1 will shortly discuss the research questions and the overall objective for this thesis. Based on these objectives, an appropriate test method to answer these questions is set forth in section 4.2. After choosing the test method, section 4.3 will present the general research approach, in order to answer the research questions, and section 4.4 will present the different types of networks that will be used as a base for comparing the different effects. The actual effects that will be investigated are presented in section 1.1, and a hypothesis for each of these effects is made in section 4.6, based on the insights gained in the previous chapter.

4.1 Research objective

The aim of this thesis is to investigate if (1) a relationship between the shape of the macroscopic fundamental diagram and the structure of the underlying network exists and on which factors this depends and (2) if the shape of the MFD of a subnetwork and its perimeter are affected by different signal settings and how this does affect the applicability of the MFD for control strategies. Although both these questions share a common interest in the MFD, it is important to note that these questions are in fact very different from one another, which will result in a different approach to each of the questions.

The first question focusses on the MFD of a complete network and in what way MFDs of different networks differ from each other, meaning that different types of networks are compared to each other. If and where differences arise, an attempt is made to explain these differences, using the insights gained from chapter 3. Apart from that, it will also be investigated if it is possible to relate and/or quantify certain network properties to the shape of the MFD.

The second question deals with the assumed interrelationship between a subnetwork and its perimeter and should form a base for further research into the applicability of the MFD in perimeter control. In 1.3.2 it was hypothesised that the way in which either of them should be controlled, has an impact on the other and that an optimum control should exist. In order to answer this question, every subnetwork and perimeter is studied separately, only comparing to other subnetworks and perimeters, to investigate whether some sort of general principle applies. This means that the resulting MFDs are not related to all the factors making up the MFD, but only to the changes in the signal settings.

In order to answer each research question, an answer to the following sub questions/research aspects should be found first:

- Does a general shape for the macroscopic fundamental diagram exist or is this network specific?
- Which factors influence the shape of the macroscopic fundamental diagram, and is it possible to quantify the effect of each of these factors?
- How is the MFD affected by changes in the network structure?
- Does a relationship between the macroscopic fundamental diagram of the subnetwork and its perimeter exist?
- How are the macroscopic fundamental diagram of the subnetwork and its perimeter affected by changes in the signal settings?
- Can the macroscopic fundamental diagram be used as an input for control strategies?

4.2 Choosing a test method

The first choice that has to be made, is what test methods are available and which method is best suited to answer the research questions at hand. Generally speaking, the research questions can be answered in three different ways: using real-life data, simulation and a mathematical approach. Below an overview of these methods, with their benefits and drawbacks is presented. Based on this assessment, a test method is selected that is best suited to answer the questions presented in the previous paragraph.

Real-life data

Using real-life data is the most representative of all methods and is basically free from any assumptions. Also no further simulation is needed.

However, real-life data has a substantial number of drawbacks. The first one is that if one wants to systematically investigate effects of network structure and signal settings, the data should be derived from a network, in which this also has been done in a systematic way. However, significant changes to the network structure are rarely made and signals are almost never operated in a systematic manner over a complete network.

Apart from that, it turns out that in none of the available datasets, the actual signal settings and log-files of intersection detectors are available. A second problem with these datasets is, that detectors are almost exclusively placed on the network arterials and are almost absent for the lower-level roads in the different subnetworks. As such, no statement can be made about the traffic state in such a subnetwork.

Mathematical approach

Although the mathematical approach can be said to be systematic by nature, the resulting formulations still have to be verified, either by using real-life data or simulation. The fact remains that a theoretical formulation contains a large amount of assumptions and that network dynamics are often not taken into account explicitly. Another drawback is that derivation of these formulations is highly complex and requires a high degree of mathematical knowledge.

Simulation

Although simulation also hinges on a lot of assumptions, the network dynamics are often taken more into account than in a mathematical approach. Apart from that, simulations offer the benefit that all parameters, apart from those that have to be evaluated, can be kept constant, making them most useful for a systematic approach. A major drawback of simulation is that substantial validation and calibration of the model is needed to obtain proper results.

4.2.2 Trade-off

Taking the above into account, the use of real-life data can be dismissed almost directly. The available datasets are often incomplete and do not have the data needed to answer our research questions, as data for subnetworks and signal settings is absent.

A mathematical derivation could be used to answer the questions, but the results would still have to be verified by simulation or existing data. Looking at the complexity of such formulations (see Helbing, 2009), the required mathematical knowledge to derive such formulations is quite substantial and the amount of time needed to obtain this knowledge would exceed the amount of time available.

The use of simulations provides the most flexibility and seems to be the most suited to solve our research questions, as it is possible to accurately control most aspects of the simulation environment. This has the benefit that differences in the simulation results can be properly attributed to the parameters that have been changed within these simulations. Also different network structures are easy to implement and changes to signal settings can be easily made in a systematic way.

4.3 General research approach

4.3.1 The macroscopic fundamental diagram vs. network structure

The first research question set out in this thesis is, is if a relationship exists between the shape of the MFD and the structure of the underlying network. In order to get a clearer understanding of this research question, it should first be explained what is understood by network structure. In this thesis, the network structure is defined as the way in which different types of roads are laid out in the network and the size of subnetworks created by the main infrastructure. For instance, if the central arterial in the network shown in Figure 1.2 would be removed, or additional arterials would be added, the structure of the network would change. This because the number of access points from the subnetwork to the perimeter increases, vehicle routes are different and arterials are operated in a different manner than subnetworks, due to the presence of controlled intersections.

In order to investigate how the shape of the MFD is affected by making changes to the network structure, a number of different networks has to be created, in which the length and layout of the arterials and the subnetworks can be changed. To make a decent comparison between the networks, all other factors should be kept equal.

Nevertheless, a couple of elements can be changed in order to investigate the impact of certain factors. Changing the layout of the network will change the length of the network which should

undeniably change the shape of the MFD. The effect of changing the free flow speed and capacity can be investigated, by creating different types of links, each with different speeds and capacities. As for our networks, this can be realised by assigning different speed limits to the subnetwork and the arterials, which should also result in different capacities of these links. Also the effect of signal settings on the MFD can be investigated, as the different networks should have a different number of intersections and the signal settings are adapted to match the traffic demand. Regarding the distribution of traffic over the network, this cannot be influenced, as this is a property of traffic itself.[11]

Apart from the above, it is also interesting to investigate to what degree the MFDs of networks with the same general layout differ from each other. Apart from comparing the MFDs of similar networks to each other, this can also be done for different, similarly shaped subnetworks and/or perimeters. Although free flow speed has also been found to affect the shape of the MFD, the maximum speed on the subnetwork and perimeter is kept constant throughout the different networks. When different speeds would have to be evaluated the number of simulations would increase substantially, increasing the amount of simulation time needed. Due to time restrictions this has been omitted.

4.3.2 The macroscopic fundamental diagram vs. signal settings

To assess the effect of signal settings on the shape of the MFD of the subnetwork and perimeter and to investigate how both are related to each other, as stated in the second research question, the following aspects are varied:

- Total lane length[12] of the perimeter;
- Total lane length of the subnetwork;
- Total inflow and outflow capacity of the signals between perimeter and subnetwork.

Other factors that can influence the shape of the MFD are kept equal, so that the differences in the MFD can be fully attributed to changes to the signals.
In order to investigate the relation between the subnetwork and perimeter, it should be tested if the maximum production and optimal accumulation for either network change when the timing of the signals between the subnetwork and perimeter are changed. Based on these results, a statement can be made regarding the applicability of the MFD as an input for control strategies.

4.4 Networks to be used

In order to investigate the effect of the network structure on the shape of the MFD, and the effect that different signal settings have on the MFD of the subnetwork and perimeter, a number of different networks are needed, that are differently structured and contain different types/differently shaped subnetworks and perimeters.
As it is chosen to use a microscopic simulation model (see section 5.2), the size of the network should not be chosen too large, as this would severely impact the simulation time and would mean

[11] This is of course not necessarily true, as rerouting can be used to distribute traffic more evenly over the network. This however is not done, due to the complexity of implementation of such measures in a simulation model.
[12] Total lane length is defined as the length of each link, multiplied by the number of lanes of that link ($\sum_{i \in Z} l_i n_{lanes,i}$).

that the number of simulation that can be run, would be too small. The networks shown in Figure 4.1 to Figure 4.5 will be used to answer our research questions. The 5 networks are created in such a way, that the effect of a number of network elements on the MFD can be tested, and the size and shape of each of these networks has been chosen in such a way that it should be able to compare the resulting MFDs to one another and differences between them can be properly attributed to a specific parameter.

The first network is a simple network that consists of a single subnetwork, with a single perimeter and serves as the base network. In the second two networks, an additional arterial is added to the network, creating two subnetworks. In the last two networks a second central arterial is added, creating four subnetworks. In both sets of networks, the arterials are either placed in the centre of the network in order to create evenly shaped subnetworks, or are placed asymmetrical to create differently shaped subnetworks. Apart from the changes this makes to the shape of the subnetwork, the main reason is to investigate if the MFD is affected in any way by the physical layout of the network, i.e. the network structure.

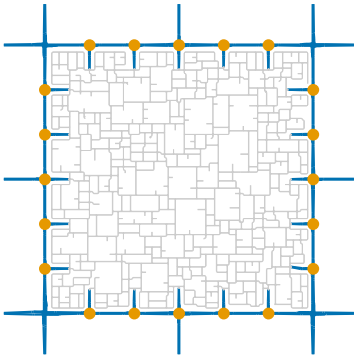


Figure 4.1: *Basic network*

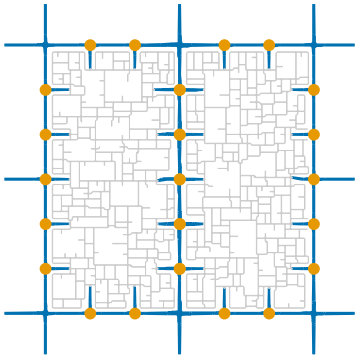


Figure 4.2: *Network with single symmetrical arterial*

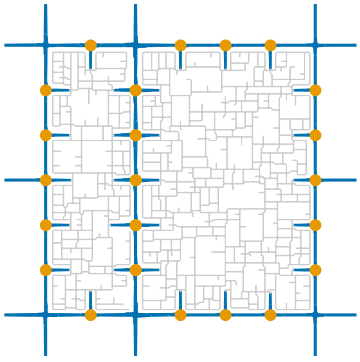


Figure 4.3: *Network with single asymmetrical arterial*

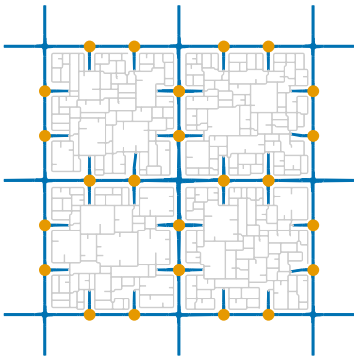


Figure 4.4: *Network with two symmetrical arterials*

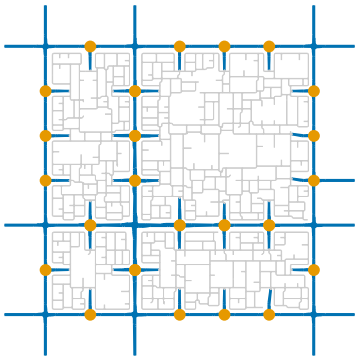
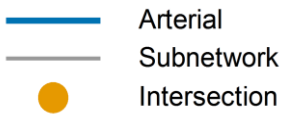


Figure 4.5: *Network with two asymmetrical arterials*



4.5 Effects to be investigated

4.5.1 Adding additional arterials

The first effect that will be investigated, is how the shape of the MFD, i.e. the production, performance, accumulation and/or network density react when additional arterials are added. This is not only investigated on a network scale, but the impact on the arterials and the subnetworks will be measured as well.

To this end, three basic sets of networks are distinguished. In the first network, the arterials consists of a single ring road, around one subnetwork. In the next two networks, an additional arterial is placed in the centre of the network and in the last two networks, even two additional arterials are added to the base network.

4.5.2 Asymmetric arterial placement

The second effect that will be investigated, is if the physical layout of the network (the network structure) has an influence on the shape of the MFD. As shown in Figure 4.2 to Figure 4.5, two different network groups have been created two answer this question. In the first group one additional arterial is added and in the second group two additional arterials are added. In order to investigate how the MFD is affected, the location of these central arterials is shifted a single intersection off-centre, directly creating asymmetrical networks and differently shaped subnetworks. The effect will not only be analysed for the total network, but also for the arterial and the subnetworks.

4.5.3 Different traffic loading patterns

It has been found in paragraph 3.2.5 that the MFD can be seriously affected by different traffic loading patterns, creating so-called 'super states' residing outside the boundary of the 'steady-state' MFD. The underlying phenomenon has been dubbed 'the rush', as this effect should be mainly found during rush hours, as the loading patterns show strong variations during this period. In order to investigate how the rush affects the shape of the MFD, the different traffic loading patterns and amount of vehicles will be analysed and will be related to the MFD, in which special attention will be paid on how production changes over time.

4.5.4 Stochasticity in network design

In order to investigate to what degree small differences in the shape of the arterials and the subnetwork and the total length of all the links might impact the MFD, two network similar to the first network will be created. The general layout of these networks is the same, but small differences in the shape of the subnetwork are present, due to stochastic design of the road network within each subnetwork. Also the number of lanes of roads along the arterial and signal settings differ for the various networks, as the underlying OD pattern for each of these networks is different, because the OD-matrices have been randomly generated.

4.5.5 Scalability of differently shaped subnetworks and perimeters

Similar to paragraph 4.5.4, the effect of stochasticity in network design can be investigated, by comparing the MFDs of both subnetworks and perimeter of the network shown in Figure 4.2, as both are equally sized.
The same goes for the subnetworks and perimeters of the network shown in Figure 4.3 and the north-western and south-eastern subnetworks and perimeters of the network in Figure 4.5. And although differently shaped, the length of the perimeter and the number of intersections of the north-western subnetworks of networks the last two networks are actually equal. Also the size of these subnetworks is roughly equal, providing a decent base for comparison.

Apart from comparing equally sized subnetworks and perimeters, it is also investigated if the MFDs from differently sized subnetworks and perimeters can be scaled to each other and to what degree differences in scale can be attributed to differences in the length of the network, number of intersections and the amount of trips made in the network.

4.5.6 The effect of signal settings on subnetwork and perimeter

It has been hypothesised that the way in which either the subnetwork or the perimeter is controlled, has an impact on the performance of the other, implying that the shape of their MFDs is highly sensitive in changes to the signal settings,
To assess how the shape of the MFD of the subnetwork and perimeter are affected by changes to the signal timing. Changing the timing of the signals is done by placing a fictive traffic demand on each signal controlling the flow from the perimeter into the subnetwork, or vice versa. Throughout different simulations, this fictive demand can be varied, to investigate to what degree the MFD is affected by the different signal settings and to see if the performance or production can actually be improved while doing this.
The changes to the signal timings are made first to the signals regulating the inflow to the subnetwork from the perimeter, and is referred to as 'controlled subnetwork inflow'. In the second set of simulations, the signal timings of the signals regulating the outflow from the subnetwork to the perimeter. In this way it can be investigated what happens if the flow on the perimeter is restricted and what happens if the flow in the subnetwork is restricted.
This effect is investigated for every subnetwork and its perimeter individually and is not done for all the subnetworks in the network at the same time. As such, the results cannot be compared to the MFD of the total network.

4.6 A hypothesis for the different effects

4.6.1 Adding additional arterials

Based on the insights gained in chapter 3, the most important factors that are changed are the network length, free flow speed and the amount of signals in the network. Below the effect on the different parameters is given for the arterials, subnetworks and total network.

Arterials:

- **Maximum production (higher):** As the arterials have increased in length, the maximum production should increase as well, as more distance can be travelled.
- **Maximum performance (lower):** The maximum performance is expected to be lower, as the average number of intersections per unit of distance will increase (higher conflict ratio), hampering the overall flow of the arterial.
- **Maximum accumulation (higher):** As with the maximum production, the maximum accumulation should increase as well, as more vehicles can be stored in the arterials.
- **Optimal accumulation and average network density (lower):** Both the optimal accumulation and optimal average network density are expected to drop. The reasoning behind this, is that the increased amount of conflicts reduces the maximum capacity of the links in the arterials. Apart from that, as more arterials are present, more traffic will use these arterials, leading to higher cycle times of the intersections, resulting in longer waiting times for each turn, which again reduces the capacity of the links in the network. Due to the reduction in the capacity, the critical density is assumed to be lowered as well, in case the average network speed remains unchanged.
- **Deviation of density (higher):** Assuming that all traffic in the subnetwork searches the shortest route to the arterials and that the origins and destinations in the subnetwork are equally distributed, every part of the arterial should attract roughly the same amount of traffic in the base network. In case a central arterial is added, half of the traffic of each subnetwork is drawn to this arterial, making it roughly twice as busy as the arterials on the edge of the network, meaning that traffic is distributed more unevenly over the network, increasing the deviation of density.

Subnetworks

- **Maximum production (lower):** Due to the longer arterials, the subnetworks should become a little shorter, decreasing production. Also the increased amount of traffic signals hinders the flow of traffic in the subnetwork, decreasing the production as well.
- **Maximum performance (lower):** The maximum performance is also expected to be lower, as more of the links in the subnetworks are now restricted by the bottleneck capacity of the traffic lights.
- **Maximum accumulation (lower):** As with the maximum production, the maximum accumulation should decrease as well, as less vehicles can be stored in the subnetworks.
- **Optimal accumulation and average network density (lower):** Both the optimal accumulation and optimal average density for the subnetworks are expected to become lower, as the increased amount of intersections reduces the capacity of the links in the subnetworks.
- **Deviation of density (unknown):** What exactly happens to the deviation of density of the subnetworks is unclear. On the one hand the reduced size of the subnetworks should lead to a more evenly distributed demand, but on the other hand, due to the increased number of intersections more vehicles are held back at the outside of the subnetwork, leading to high density areas near the perimeter access points, creating a higher imbalance in the density.

Total network

- **Maximum production (lower):** Production of arterials is assumed to increase, while production of the subnetworks is assumed to decrease. As the subnetworks form the greatest part of the network and because the additional traffic signals increase waiting times, thus lowering the capacity, the maximum production of the total network is assumed to be lowered.
- **Maximum performance (lower):** As for both the arterial and the subnetwork the performance is assumed to become lower, the maximum performance of the total network should be lower as well.
- **Maximum accumulation (lower):** Due to a decrease in the size of the total network, the maximum accumulation is assumed to be lowered.
- **Optimal accumulation and average network density (lower):** As for both the arterial and the subnetwork the accumulation is assumed to become lower, the optimal accumulation of the total network should be lower as well. Apart from that, adding additional arterials should lead to a reduction of the average travel times, as routes are shortened. Due to the reduction of travel time, the optimal accumulation should lower as well.
- **Deviation of density (higher):** Due to the added arterial, the number of routes taken throughout the network should become less and traffic will most probably cluster around the central arterials, leading to a less homogeneous distribution of traffic over the network. This should then result in a higher deviation of density.

4.6.2 Asymmetric arterial placement

Taking into account all the factors making up the MFD, none of them seems to be dependent on the actual lay-out of the network itself, but seem to depend on factors related to the individual links. However, it has been found that the actual capacity of a link is not related to that link, but to the capacity of the first downstream bottleneck. Although one of the subnetworks might become smaller and therefore more of the links are affected by the bottlenecks created by the intersections on the perimeter, this effect should be cancelled out by the other larger subnetwork(s).
However, some changes in the shape of the MFD might be witnessed in case traffic patterns are asymmetrical as well. If for instance the average number of trips originating from the smaller subnetwork might be higher than in the larger networks, than more trips benefit from the shorter routes and improved travel time, leading to a slightly lower accumulation.
Assuming that the OD-demand is spread equal over the network, ***the shape of the MFD is not expected to change.***

Only the deviation of density is assumed to increase, as the smaller subnetwork(s) are assumed to be hindered more by the intersections around it, causing high densities on these links. The other subnetwork(s) however, should have lower densities, leading to a higher deviation of density. This however does not have an effect on the shape of the MFD, apart from an increased amount of scatter.

4.6.3 Different traffic loading patterns

As has been shown in paragraph 3.2.5, the way and the speed with which traffic is loaded onto the network, can seriously affect the shape of the MFD, creating 'super states' that reside outside the

boundary of the 'steady-state' MFD. How far outside the boundaries these states are found is assumed to depend on the speed with which vehicles are loaded onto the network. Assuming that vehicles are loaded onto the network linear in time, ***changes in the maximum production and optimal accumulation should be roughly linearly related to the number of vehicles loaded onto the network per unit of time***. Meaning that if double the number of trips should be served in the same network within the same amount of time, the maximum production found is roughly twice as high.

4.6.4 Stochasticity in network design

As no substantial changes are made to any of the network parameters, besides the network length, the production and accumulation are assumed to be linearly related to the length of each network, meaning that ***the MFD relating performance and average network density should almost be the same for the different networks***. Any differences between the MFDs should therefore be contributed to differences in signal settings, or are assumed to be caused by small differences in the network layout, creating blocks along which gridlock can arise more easily.

4.6.5 Scalability of differently shaped subnetworks and perimeters

It is not expected that differences in the MFDs of subnetworks could be attributed to the size of the subnetwork (area), or the circumference of its perimeter. As in the previous paragraph, it is assumed that the shape of the MFD, i.e. the maximum production, optimal accumulation and maximum accumulation can all be largely contributed to the lane length of the subnetwork or perimeter. However, the shape of the MFD is assumed to be highly dependent on signal settings, which lowers its production and increases the optimal accumulation. As such, it is expected that the relation between the network length and the number of intersections, i.e. the total waiting at signals controlling the links under investigation, should explain remaining differences in the MFD.

Based on the above it is conjectured that ***the MFD relating performance and average network density should almost be the same for the different subnetworks and perimeters, as long as the ratio between network length and the total waiting time at signals is the same.***

It should be worth mentioning that this conjecture assumes that traffic conditions between different networks are the same and homogeneous enough, to avoid any 'super states' from arising.

4.6.6 The effect of signal settings on subnetwork and perimeter

The effect of signal settings on the shape of the MFD of the subnetwork and the perimeter is tested in two ways. The first method is controlling the flow from the perimeter to the subnetwork and the second is controlling the flow from the subnetwork to the perimeter. The described effect is therefore not the expected effect in relation to the original results (as described in the previous paragraphs), but is the predicted effect of increasing the total amount of green time of the different traffic signals controlling the subnetwork and perimeter.

The effects described below are based on controlling the subnetwork inflow. For the effects of controlled subnetwork outflow, the effects are assumed to be the other way around, i.e. the effects of the subnetwork outflow are equal to the perimeter inflow and vice versa.

Perimeter

- **Maximum production (higher):** As the total amount of green time given to vehicles in the perimeter becomes larger, production should increase, due to an increase of the system capacity.
- **Maximum performance (higher):** As the production is assumed to increase and no changes in the infrastructure are made, the maximum performance is assumed to increase as well.
- **Maximum accumulation (equal):** As no physical changes to the network are made, the total number of vehicles that can reside in the network remains unchanged, resulting in the maximum accumulation to remain equal as well.
- **Optimal accumulation and average network density (lower):** As green times increase, more traffic can be processed by the perimeter, as the improved signal settings lead to a reduction in travel times, resulting in a lower optimal accumulation.
- **Deviation of density (lower):** As the amount of green time is increased, less pockets of congestion can occur on the arterial, leading to a more even distribution of traffic over the network.

Subnetwork

- **Maximum production (equal):** No changes are made to the structure of the subnetwork, or the timing of its controlling signals, thus keeping production equal.
- **Maximum performance (equal):** As the production is assumed to remain equal and no changes in the infrastructure are made, the maximum performance is assumed to remain equal as well.
- **Maximum accumulation (equal):** As no physical changes to the network are made, the total number of vehicles that can reside in the network remains unchanged, resulting in the maximum accumulation to remain equal as well.
- **Optimal accumulation and average network density (equal):** No physical changes to any of the elements governing the subnetwork is made, meaning that its MFD should stay the same.
- **Deviation of density (unknown):** As the amount of green time is increased, more traffic ends up in the subnetwork. Whether this results in a more even or uneven distribution of traffic is not completely clear.

Overload and optimal signal timings

It should be noted that it is expected that at some point increasing the green time of the traffic signals does not result in more traffic flowing into the subnetwork, as this can already be served at lower green times, thus meaning that certain effects will level of, e.g. the production will not increase any further after a certain point.

It is also expected that the effects described above, reverse at some point and the total network performance decays heavily. The reason for this is, that when the amount of green time given to

signals regulating traffic into the subnetwork, this is done at the expense of the amount of green time of the other signals. If the amount of green time becomes too high, not enough time is left for the other signals to serve all traffic, resulting in spill-back and heavy congestion (or even gridlock) in both the perimeter and the subnetwork.

All in all, increasing the subnetwork inflow is not assumed to benefit the subnetwork in any way other than that this creates more space on the perimeter, resulting in less spill-back and better outflow from the subnetwork, or vice-versa in regard to controlling the outflow.

It is assumed there is a certain optimal timing, in which both systems benefit the most from each other. However, it is not expected that in this case the production, performance, accumulation or average network density should become higher than in the first set of simulations in which the inflow or outflow is not controlled. This because in the previous case, all signal timings are adapted to the specific flow of every turn of every intersection. Imposing the same green time on all the signals controlling the inflow to or outflow from the subnetwork, do not cater to the specific needs of each single turn, and should therefore always disrupt the balance present in the network.

5 Modelling framework

The literature study in chapter 2 has shown that the relation between the shape of the macroscopic fundamental diagram and the underlying network structure and the way that network is controlled, has only been investigated mathematically. So far, no study has systematically investigated this relation using real data or a simulation model. Within this chapter a framework will be presented, able to model the impact of different network structures and signal settings on the shape of the MFD. First the general framework is presented in section 5.1. In section 5.2 different model types and programs are discussed and a trade-off between these models and programs is made. Subsequently, the model and algorithms needed to feed the simulator are presented in sections 5.3 and 5.4. Section 5.5 discusses how the output of the simulation is transformed into the data needed to construct the desired MFDs.

5.1 General framework

The model used within this thesis is composed of four main components.

- A microscopic dynamic traffic assignment (DTA) simulator (VISSIM);
- A network creation model;
- A network structure simulation algorithm and a controlled in-/outflow simulation algorithm;
- A data processor.

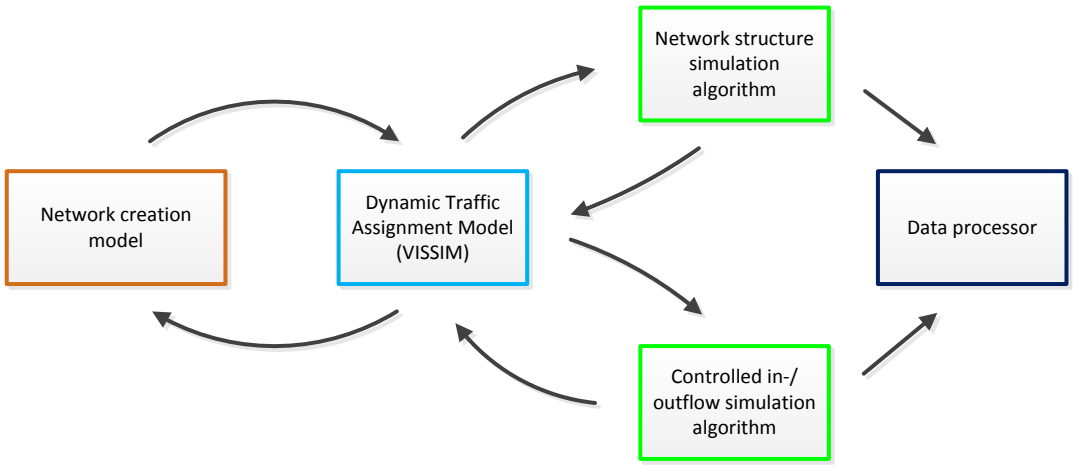


Figure 5.1: *General framework*

In the subsequent paragraphs a detailed description on how the framework shown in Figure 5.1 was created and the underlying choices is set forth.

5.2 Simulation environment

5.2.1 Type of simulation model

As paragraph 4.2 concluded that simulation is the most appropriate method to answer the research questions, the next choice that has to be made, is what type of simulator has to be used. The different simulators are often categorised by the level of detail incorporated in the model. The main types are: microscopic, macroscopic and mesoscopic. Below a description of each model has been given, based on California Department of Transportation (2011).

Simulation models

Microscopic simulation model

In a microscopic simulation model, each vehicle is modelled on an individual basis, taking into account specific preferences of the driver and/or vehicle characteristics. The assignment of vehicles is done dynamically, in which factors as acceleration, following distance and gap acceptance are taken into account. Drawback of these models is that the large set of variables results in large calculation times. Also the calibration of microscopic models is highly complex, as vehicle and driver characteristics should be tuned to match actual behaviour, which often consists of parameters which are hard to measure in real life.

Macroscopic simulation model

A macroscopic model runs simulations by using aggregated values for flow, density and speed on a link-level. Individual vehicles are not tracked and vehicle streams are routed based on split rates at the nodes of the model. The traffic process at intersections is simplified and most often calculates the movements over the intersection by taking the traffic densities and intensities of the incoming and outgoing links into account.

The absolute advantage of macroscopic models is that the computational time is low, as detailed vehicle interactions do not have to be simulated. The obvious drawback is that vehicle interactions are not modelled explicitly, which makes it more difficult to evaluate control strategies.

Mesoscopic simulation model

A mesoscopic simulation model combines elements of the microscopic and macroscopic simulation model. Individual vehicles are modelled and routes are chosen on an individual basis, based on link travel times. Assignment of the vehicles to the network is done on a macroscopic level, i.e. by using aggregate data for flow, density and speed.

Trade-off

The research questions could be answered by using any of the models described above. If all of the models would meet the criteria, one should prefer a macroscopic model, as it takes the least amount of time for calculation.

A drawback of using macroscopic and mesoscopic models is that the actual modelling of the traffic process at intersections is done in far less detail. As we are especially interested in this part of the

network, using a microscopic simulation model seems to be more suited to answer our research questions.
Another benefit is that a microscopic model gives more insight in the traffic process at intersections, as one can accurately observe the movement of individual vehicles.
The drawback of using a microscopic model is that the network has to be coded in a high level of detail and its high calculation times. Also calibrating the model takes a lot of effort, but can be circumvented when only fictional networks are simulated.

As it is important to get substantial insight in the traffic process at intersections and to be able to accurately control signal settings (both offline and online), it has been chosen to use a microscopic simulation model, as none of the available macroscopic models contained a node model which met those demands. Especially the requirement that the signal timings should be adaptable during the simulation was met by none of the macroscopic node models. [13] This because macroscopic node models calculate travel delays over the node as a function of the downstream traffic demand and upstream capacity, instead of using the exact timing of signals.
In Muller et al. (2011), for instance, it was found that using the Webster formula to calculate the cycle time (which is often used in macroscopic models), often leads to inaccurate cycle times, as the cycle time for an intersection is chosen based on the highest cycle time of any of the conflict groups. In this case, it is not taken into account that cycle times can be shortened by using extended green, or is elongated as different conflict groups have to 'wait' on each other due to different clearance times.

5.2.2 Choice of simulation program

Specific model requirements

In order to find the right simulation model, a number of requirements has been formulated, which should be met by the model.

- As the effect of different signal settings on the MFD is an important part of this thesis, it is highly important that each of these models should have a detailed node model, in which all signals of an intersection could be adapted individually, to accurately measure the effect of these changes;
- Vehicles have to be assigned to the network dynamically, as only in this way the effect of signal settings can be properly assessed.
- The simulation model should be able to communicate with an external programming environment, which in our case is Matlab. Within Matlab the timing of different signals can be adapted and the simulation model should be controlled externally to load different networks and control scenarios;
- The simulation model must have the possibility to adapt signal timings during the simulation itself;

[13] This requirement is not absolutely necessary for this thesis, but should be met nonetheless, as this makes it easily possible to incorporate demand adaptive control in further research, which can be used for testing perimeter control strategies.

- Different types of networks should be available for the simulation model, or should be easy to create;
- The model should be able to generate different time-based paths, as to simulate the effect of changes in driver behaviour when making changes to the timing of the signals in the network;
- The model must be able to generate the desired data for the MFD (traffic density and intensity for every individual link).

Trade-off

The number of microscopic simulation models available is quite limited and the following models have been considered: DYNASMART, Paramics and VISSIM.

- Paramics was quickly discarded, as the program was not directly available, as it suffered from implementation problems.
- DYNASMART was considered a serious candidate, but was discarded as the node model was not detailed enough and functioned more like a mesoscopic model. Although a node model could be developed, the actual architecture of DYNASMART is closed and implementing a new model would not be possible.
- VISSIM checks almost all of the boxes. It has a specific module to create signal schemes and gives decent insight in the traffic process at intersections. It has a built-in, easy-to-use COM-interface to communicate with Matlab and signal schemes can be implemented externally. It generates the required data and incorporates route generate, path choice and departure time modelling. Other factors are its availability and familiarity to the author.

In light of all of the above, VISSIM (version 5.20-02) has been chosen as the simulation program. Apart from its advantages, it also has some drawbacks. The main drawbacks are the extensive calculation time needed and the fact that creating different networks is a substantial task, as each model has to be designed in a high level of detail.

5.3 Network creation model

5.3.1 General

Within this thesis multiple network types will be evaluated. However, only a few networks are available for VISSIM and mostly represent existing networks. However, as we want to systematically evaluate different networks, the networks should be largely similar, in order to relate specific changes in the MFD to the network structure, or the way it is controlled.

Because sufficient networks are not available, new networks are created. As these networks are fully fictional, neither the

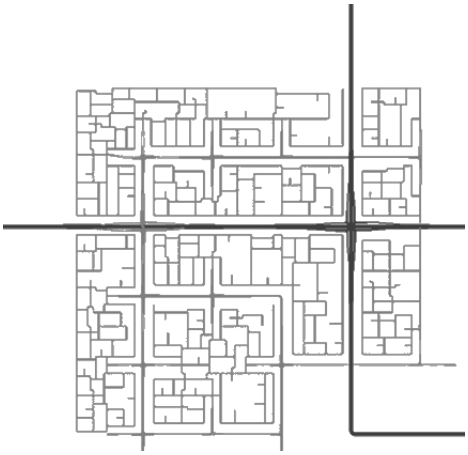


Figure 5.2: *Example of VISSIM network created with the network creation model*

network layout, origin-destination matrix, signal settings and route choice are available and these have to be created from scratch. Because multiple large-size networks have to be created, it would be too time-consuming to create each complete network by hand.

As such a network creation model has been developed, which fully automates the creation of the various network components, signal timings and origin-destination matrix. This also has the added benefit that general adaptations to the network model are fairly easy to make and do not involve manually adapting every single component in the model itself. Another benefit is that the network is constructed in a uniform and consistent manner, making it a lot easier to read and allocate the output data and transform this data into MFDs.

The downside of using a fictional network with a made-up origin-destination matrix, is that the physical layout of the network and various signal settings are not completely in tune with the actual traffic patterns. In order to somewhat overcome this problem, the network is updated, using simulation results from VISSIM, to better match the traffic patterns. Within these updates, the number of lanes of every link are adapted to match the demand and for the intersections, the signal timings and the number of auxiliary lanes (*Dutch: 'opstelvakken')* are tuned in such a way that the intersection can meet the traffic demand.

5.3.2 Description of the model

The basis of the network creation model is formed by an Excel-sheet in which the general layout of the network is created, containing the different intersections and all the roads between them. Within the sheet, only the main infrastructure is designed. The areas enclosed by the main infrastructure is converted into subnetworks, thus resulting in a network consisting of subnetworks and perimeters.
In Figure 5.3 an outline of the different components/steps of the network creation model is shown. In the following steps a summary of the model workings is presented. An extensive explanation of the model is given in Appendix A.

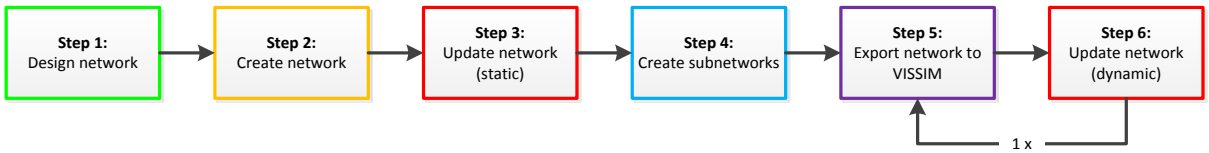


Figure 5.3: *Main components in network creation algorithm*

Step 1: Design network

In the first step a general design of the network is made in Excel. Using predefined values, an abstraction of a network is made. This network can then be imported into Matlab, which transforms it into a network that can be used in VISSIM. Within this design, a function can be assigned to every intersection in the grid. An example of such an Excel grid is given in Figure 5.4. The coloured fields (with the exception of the field along the border) within the grid are the intersections, which are connected by different links. In each of the intersections and links a number can be found. For each intersection, this number corresponds with the type of intersection and what other types of

intersections it can connect to. For links (not coloured) the number before the comma represents whether it is an arterial link, or a freeway link. The number behind the comma (only used for arterial links) represents the speed on that link.
This layout forms the basis for each network and is the only thing that has to be done by hand. After this has been done, and all parameters for the model are set, one can sit back and wait until the network is generated, simulated and the resulting MFDs and datasets are produced.

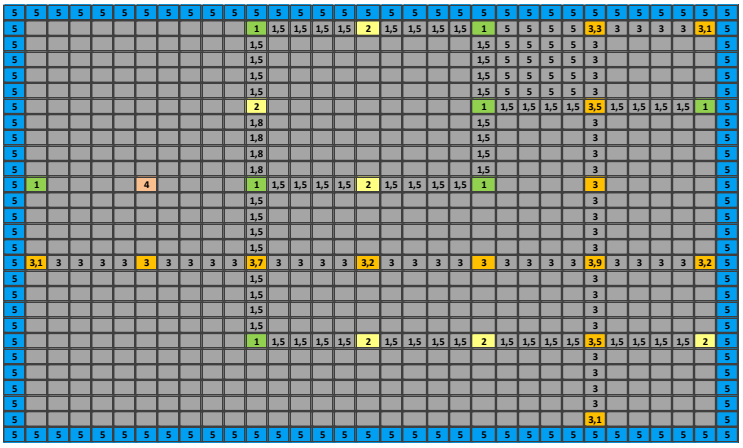


Figure 5.4: *Example of network layout in Excel*

Step 2: Create network

Within the second step, the previously designed network is imported in Matlab. Based on the function assigned to every intersection, that intersection looks in all four directions, if an adjacent intersection (that is in the same vertical or horizontal line) is present which it can connect to. With this, the basic nodes and links of the network are created. A Dijkstra-algorithm is then used to remove any node that is not connected to the network. The result of this is shown in Figure 5.5.

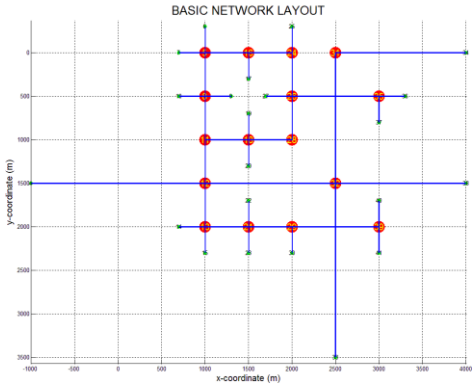


Figure 5.5: *Network after converting to links and nodes and removing 'dangling nodes'*

Next all nodes are converted into intersections and their basic physical layout is created, including full signal schemes, by placing a fictive demand of $q_{init,int,tn} = 100\ veh/h$ on each turn of the intersection. Although only one lane is assigned to each active turn in the basic intersection, the algorithm can generate the layout for any intersection with 12 turns, independent on the number of lanes of each turn. Using the equation for an ellipse, the physical curve between every turn is calculated. Next these curves are transferred to a grid, and using an overlay, the conflict areas are determined (see Figure 5.6) and clearance times are calculated.

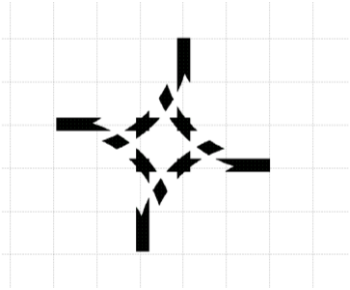


Figure 5.6: Conflict areas

With these conflict areas, the different conflict groups and control structures can be determined. Using the fictive traffic demand, saturation flow and saturation adjustments at every turn, the cycle times for every conflict group in every control structure is determined. Using an optimization method, the most optimal cycle time for each structure is calculated, resulting in a full signal scheme with start green, end green, extended green and start red.

After the signal scheme is calculated, the length of each lane is calculated using a Poisson-distribution for the arrival rate, in order to avoid blocking back of auxiliary lanes. Next the link-node configuration of the intersection is created and the intersection is put into the total network. Using the incoming and outgoing lanes of each intersection, the remaining links in the arterial are created, by connecting the adjacent intersections to each other. Nodes located on the outside of the network are converted into feeders, i.e. locations with an origin and destination zone.

Next a fictive origin-destination matrix is constructed, in which the trips between all origins and destinations are generated using a uniform distribution. The total number of trips in the network that should be made in the network is also an input in the Excel sheet. First the number of OD-pairs are calculated, after which the total demand is divided by the number of OD-pairs. As a uniform distribution is used, the upper bound is set to twice this number, with a lower bound of zero. After all trips are randomly drawn, the OD-matrix is scaled to roughly match the number of trips given.

Using the network frame as created, the free flow travel times between each OD-pair are calculated using the link speeds and the average delay for each turn at the intersections. With these travel times a k-shortest path algorithm determines routes for each OD-pair and a Deterministic User Equilibrium (DUE) assignment, using the Frank-Wolfe algorithm and the Method of Successive Averages (MSA) is performed, resulting in traffic intensities for each link on the arterial.

Step 3: Update network (static)

Using the link intensities obtained in the previous step, the links and intersections are updated in order to match the network demand. The number of lanes for each link in the arterial and each freeway links is determined by dividing the traffic intensity over the lane capacity.
For intersections, the newly generated traffic intensities are used to recalculate the signal schemes and cycle times. If the cycle time is larger than the maximum allowed cycle time, an additional lanes is added to the turn with the highest intensity per lane. When this is done, the new signal scheme

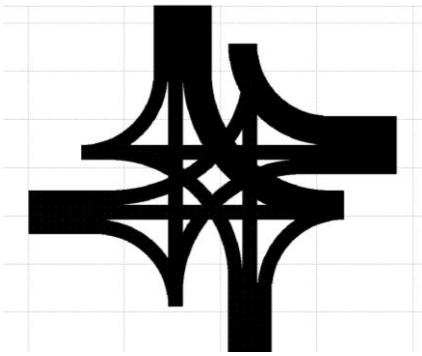


Figure 5.7: Updated intersection

is calculated. This process is repeated until the cycle time is below the maximum allowed cycle time. In Figure 5.7 an example of such an updated intersection is shown.

Step 4: Create subnetworks

A next important step is to create the subnetworks. The subnetworks are created as the areas between or next to the arterial, as is shown in Figure 5.8. After the area of each subnetwork has been obtained, a random street pattern is inserted into this subnetwork. The street pattern is obtained by adding virtual building blocks in the grid of each subnetwork. The shape and width of each block are based on characteristic components in a residential area, such as flats, houses, parks and sport parks. The distribution of each of these blocks is also loosely based on existing residential areas.

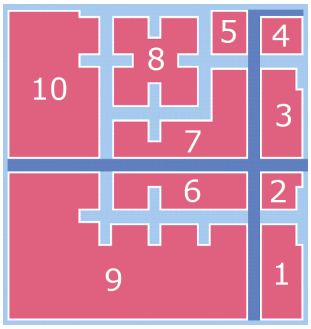


Figure 5.8: *Basic subnetwork grid*

To generate a subnetwork, first a 'safety zone' is added, surrounding the subnetwork and creating enough space for the arterials and freeways to fit into. Next the roads from the intersections on the arterials leading into the subnetwork are added and a base road along the edge of the subnetwork is added. Than a randomly generated block is added, and a road is created around it. Blocks are added from the lower left to the upper right, until no block can be fitted into the subnetwork anymore. The result of this for subnetwork 9 of ***Figure 5.8***, is shown in Figure 5.9.

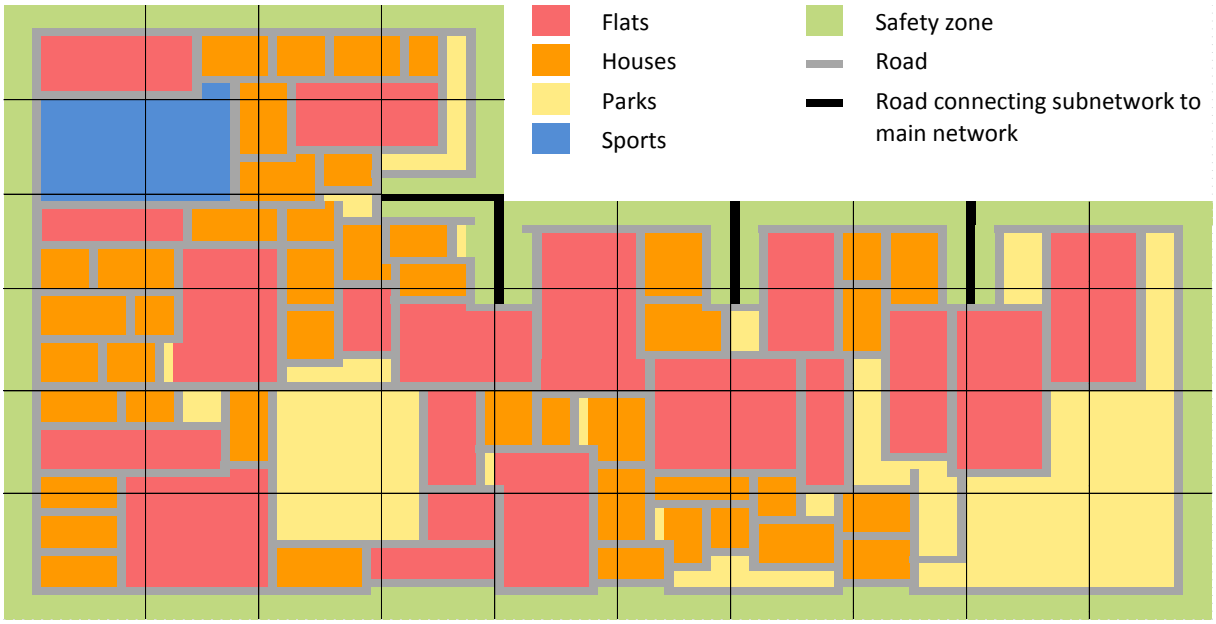


Figure 5.9: *Subnetwork after blocks are inserted*

After the complete subnetwork is filled up, the building blocks discarded and feeders are added in the subnetwork. To this end, the subnetwork is divided into smaller sections (shown by the cross-pattern in Figure 5.9) and a feeder is generated randomly within each section, and is then connected to the road network.

Next, all of the trips of each OD-pair located at the end of the roads connecting the subnetwork to the arterial, are distributed over the feeders in the subnetwork, based on the closeness of each feeder to the original location. Trips between OD-pairs located in the same subnetwork are discarded, based on the assumption that travel distances within the subnetwork are sufficiently small, to make them by other means of transport, such as walking or cycling.

Finally the street pattern is converted to a node-link pattern, resulting in the network shown in Figure 5.10. Each link in this graph is made bi-directional and curves for each of the turns within the intersections are created. All of the intersections in the subnetwork are uncontrolled.

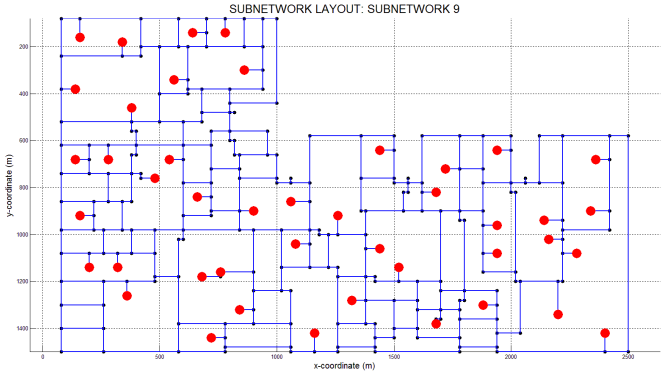


Figure 5.10: *Subnetwork after adding feeders and conversion to links and nodes*

Step 5: Export network to VISSIM

In this step, all of the various components designed and created in the previous steps are exported to VISSIM. In order to do so, the layout of the intersections has to be redesigned, as VISSIM uses a link-connector structure instead of the link-node structure designed in Matlab. In redesigning the intersections, the main task is redefining the way lanes from the network connect to the different lanes of each turn. When this is done, all links and connectors in the intersection can be created. After all of the intersections have been created, they are (where applicable) connected to the different subnetworks and the network is finally completed.
With the total network done, the network, OD-matrix, signal settings and simulation settings are written to a number of text files, which enable visualization and simulation of the network in VISSIM. This conversion is shown in Figure 5.11 and Figure 5.12.

Step 6: Update network (dynamic)

After the network is created and has been converted to VISSIM, the network is updated using simulation results from VISSIM. This process can be run completely automated, as VISSIM can be controlled from Matlab. After VISSIM is started, the network is loaded, simulation parameters are set and an initial number of simulations ($n_{runs,VISSIM} = 10$) is run to obtain a substantial path set. After the initial simulations have been run, a last simulation is run in which traffic volumes are read from the various links and connectors. Based on these volumes, the number of lanes of the arterial and freeway links are updated and the intersections are fully recreated, following the same process as in step 3. Then step 5 is repeated, resulting in a fully updated network. As with the previous

network, a number of initial simulation runs is made to obtain a path set for the complete network. At this stage the network is finalised and the network is exported to a new folder, containing all data of the network, ready to produce MFDs.

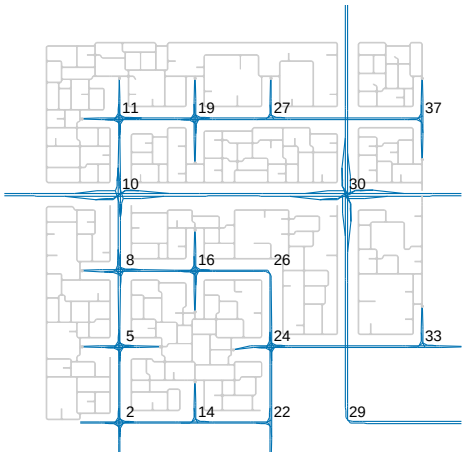


Figure 5.11: *Network in Matlab*

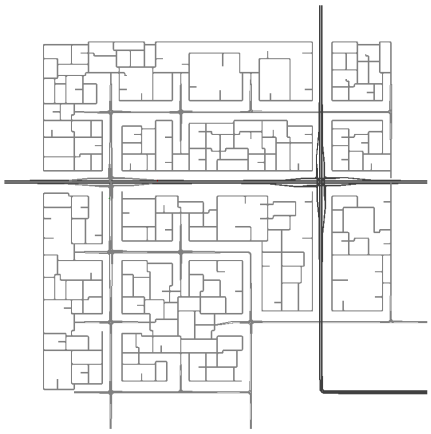


Figure 5.12: *Network in VISSIM*

5.3.3 Further applications of the network creation model

The model that has been developed, is not only usable for this study, but can be used for many other applications. The reason that this model has been created, is because no other model was available that could generate networks and accurately design large amounts of intersections, including their signal schemes.

As such, this tool should fill a void for anyone who wants to test a simple principle in a basic urban network. A network shown in Figure 5.11 takes about 7 minutes to generate in Matlab and 39 minutes total to convert it to VISSIM and run the dynamic update.[14] Considering the amount of time normally needed to create a complete network for a microscopic simulator, this can save a decent amount of time.

Also the system produces the network in a link-node format, making it possible to use the generated networks in a macroscopic model. Additional programming will have to be done, but the main principle is already available. Due to this integration, the model has a high flexibility and could very well be used for a wide range of applications, while circumventing the problem of converting a macroscopic to a microscopic model.

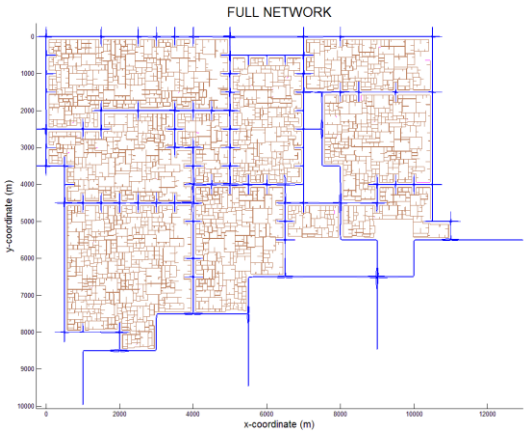


Figure 5.13: *Abstraction of the Hague*

[14] Windows 7 Performance-index: Processor 7.4, Memory 7.4, Graphical 5.7, Primary hard disk 5.9

Furthermore, an abstraction of a large network can be made, in order to do some simple computations. In Figure 5.13 an abstraction of the city of the Hague is shown, which took 1.5 hours of manual labour and 5 hours of computational time. By incorporating different block sizes for different types of area, more realistic subnetworks can easily be created.
Finally (but not exhaustive), the model can be used to create environments for driving simulators. In the newer versions of VISSIM, functionality has been added to connect networks that have been created in VISSIM to driving simulators. Although the current layout of the networks is still very rigid, it does provide an easy way to create large basic networks, which can then be adapted to meet the demands. This should be much faster, then if they were built completely from scratch.

5.4 Simulation algorithms[15]

The third component in the framework are the algorithm that actually run our simulations and generate data needed to create our MFDs.

5.4.1 Obtain data to evaluate the effects of network structure on the MFD

At first, a set of simulations is run for each network, in order to obtain data to create MFDs for the complete network, the arterials and the subnetworks. In Figure 5.14 an outline of the algorithm is presented.

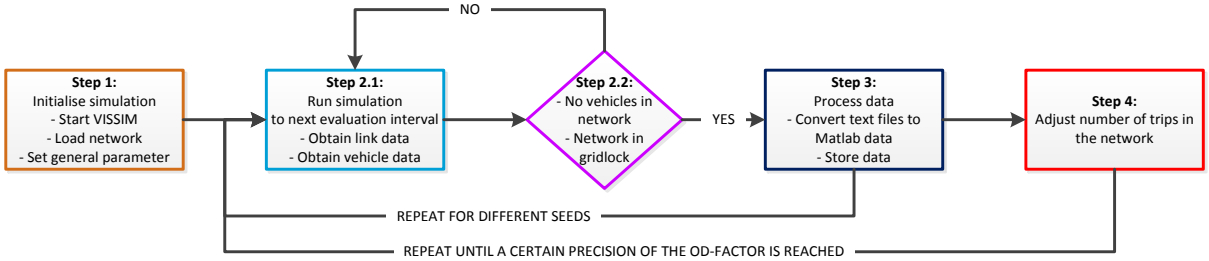


Figure 5.14: *Simulation algorithm to evaluate the effect of network structure*

Step 1: Start simulation

VISSIM is called from Matlab using a virtual COM-interface, after which the network is loaded and a number of parameters are set and the simulation is started. The complete simulation process is controlled from Matlab and is fully automated.

Step 2.1: Run simulation and obtain data

Within the simulation, all vehicles are loaded on the network over a period t_{OD}. The simulation is run for a period $t_{tot,sim}$, as to enable all vehicles to reach their destinations.

At a specified interval $t_{data,i}$, the simulation is paused and data from the simulation is either obtained directly from VISSIM by Matlab through the COM-interface, or written to a text file and processed by Matlab after the simulation has finished. For every link in the network, the average density, speed, flow and cumulative number of vehicles past over the last iteration are obtained. For each vehicle in the network, the vehicle number and simulation time are written to a file (which is processed after the simulation has ended), in order to assess the average travel time for each

[15] In this section, numerous variables are presented. The actual values assigned to each of these variables are presented in chapter 6.

vehicle in the network. In order to obtain accurate travel times, this evaluation interval $t_{data,veh}$ is set far lower.

Step 2.2: Check if simulation has ended

The simulation is ended in case the network is empty, i.e. all vehicles have been loaded onto the network and have reached their destination, or the network is in gridlock.
To assess if the network is in a state of gridlock, the link number of every vehicle is obtained online at each $t_{data,i}$. It is then checked how much per cent of the vehicles in the network has not changed to a different link over the last interval. If this percentage transgresses a certain threshold value ρ_{gl}, the network is assumed to be in a state of gridlock and the simulation is ended. It is also assumed that gridlock has set in, when the total simulation time $t_{tot,sim}$ has ended and still a number of vehicles reside in the network.

Step 3: Process data

During this step, the data obtained from the simulation is stored in a structure in Matlab. If the data was written to text files, these text files are read and data in them is converted to matrices, containing link data and vehicle data. For a detailed description, see section 5.5.

The simulation is repeated for n_{seed} times, using the exact same settings for the total network. However a 'random seed' is generated, which sets a random number generator in VISSIM in motion, changing the moment at which is vehicle is loaded onto the network and therefore changing the traffic patterns and simulation results.

Step 4: Adjust number of trips in the network

When the simulation has been run for n_{seed} times, the total number of trips between each origin and destination are multiplied by a factor $OD_{fact} = OD_{fact} + OD_{incr}$. The factor OD_{incr} is changed, depending on the outcome of the simulations. The number of trips are always derived from the original (unchanged) OD-matrix.
When n_{seed} simulations have been run, it is checked if any of the simulations ended in gridlock, or if the network was empty for all seeds. If the state of the network has changed (empty to gridlock, or vice versa), the factor OD_{incr} is multiplied by -0.5. If this is not the case OD_{incr} remains unchanged, i.e. the total number of trips is increased again. If however OD_{fact} would end up a factor it has already been, OD_{incr} is multiplied by -0.5 as well. [16]
In this way, the point at which gridlock occurs, is found iteratively. This is done to obtain a substantial amount of data points around the point where the network is between free flow and gridlock, i.e. the point in the MFD of maximum production and optimal accumulation, which is our main point of interest. This loop and the simulations are ended after a precision of $|OD_{incr}| < OD_{prec}$ has been reached.

[16] When for example gridlock occurs at 2,19 times the original number of trips, and the simulations are initiated with $OD_{fact} = 0.5$, $OD_{incr} = 0.5$ and $OD_{prec} = 0.1$, the sequence of simulations would be: 1,00 -> 1,50 -> 2,00 -> 2,50 -> 2,25 -> 2,13 -> 2,19.

5.4.2 Obtain data to evaluate the effects signal settings on the MFD

The goal of the controlled in-/outflow simulation algorithm is to obtain MFDs for the different subnetworks and perimeters in the network, under different signal settings and control scenarios. As the network has been designed by the creation model to cope with the total traffic demand, one should not expect to always obtain a full MFD, as gridlock should not occur. Although adaptations to the signal settings could result in gridlock, this should not always be the case, as the network could be sufficiently redundant. To overcome this, the number of vehicles is increased continuously, until gridlock is achieved, resulting in a sufficient number of data points for the MFDs. An outline of this algorithm is presented in Figure 5.15.

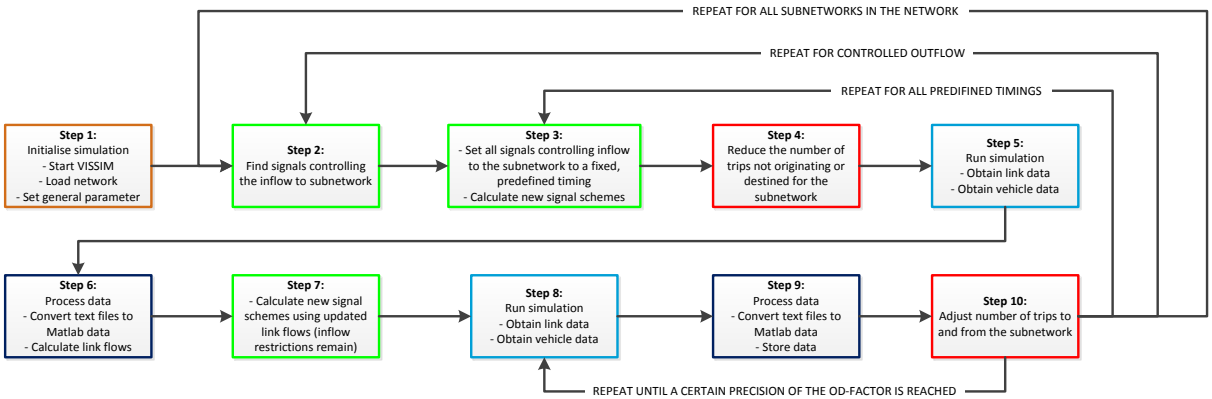


Figure 5.15: *Simulation algorithm to evaluate the effect of signal settings.*

Step 1: Initialise simulation

See step 1 of paragraph 5.4.1.

Step 2: Find signals controlling the inflow or outflow to each subnetwork

After the simulation has been initialised, the signals controlling the inflow or the outflow to each of the subnetworks are searched and stored in a separate matrix.

Step 3: Set the signals controlling subnetwork inflow or outflow to a fixed timing

After the signals have been found that control the inflow or outflow to the subnetwork, the first subnetwork and perimeter are selected and each signal controlling inflow to the subnetwork is set to a fixed value q_{ctrl} (depending on the control scenario), in order to accommodate a specific number of vehicles per hour.

To illustrate this, consider the intersection shown in Figure 5.16 and suppose a subnetwork is connected on the western side of this intersection. When controlling the outflow from the subnetwork to the perimeter, the restriction is placed on signals 7, 8 and 9. When each traffic light is restricted to an intensity of $q_{ctrl} = 100\, veh/h/signal$, the resulting number of vehicles that can leave the subnetwork is maximally 300 per hour. In case the inflow to the subnetwork is controlled, these restrictions are placed on signals 2, 6 and 10 instead.

After the outflow intensities/timings of the controlling signals are fixed, a new signal scheme for every intersection in the network is calculated. Although the original signal scheme algorithm in the network creation model also calculates extended green times, this is blocked for the signals controlling the inflow and outflow. This is done because the green time of the right hand turn is often extended, as conflicting turns are often blocked by the adjacent through-going turn. When the use of extended green should be permitted, the total permitted outflow at these turns could become larger than specified by the control value.

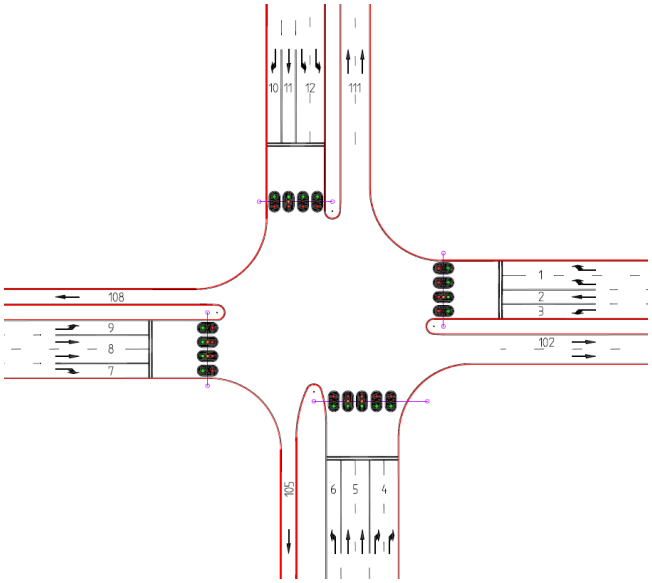


Figure 5.16: *Example of intersection and lane numbering*

Step 4: Reduce number of trips not originating from or destined for the subnetwork

The next step is to reduce the number of trips that do not start or end in the subnetwork under evaluation, i.e. do not have an origin and/or destination within the subnetwork, with a factor δ_{red}. The reason that this is done, is because of these vehicles, congestion on the perimeter would occur sooner, causing queues to grow into the subnetwork. As a consequence of this, the number of vehicles in the subnetwork will increase faster than normal, which could result in gridlock. The resulting gridlock in the subnetwork (or on the perimeter for that matter) cannot be fully attributed to the way the specific subnetwork is controlled, as the spill-back is a consequence of insufficient capacity elsewhere in the network. As a result, it is not possible to fully relate the shape of the MFD to the way the subnetwork is controlled, or the number of vehicles that are present in the perimeter, which is basically the main goal of this algorithm.

It should be noted that because of this reduction, the resulting MFDs of the subnetwork and perimeter cannot be fully compared to the MFDs of the simulations described in paragraph 5.4.1. This however should not form a problem, as we are mainly interested in how the MFD of the subnetwork and its perimeter are affected by the different signal settings, not how this affects the rest of the network.

Step 5: Run simulation

In this algorithm the simulations are carried out in the same way as in step 2.1 and step 2.2 of paragraph 5.4.1.

Step 6: Process data

From the data obtained from the simulation, the intensities on all the links are calculated. This is done by summing up the number of vehicles that have passed each of the links during the simulation and dividing this by the total time the simulation has run. If the simulation has ended in

gridlock, the end time is set to the time in which $\rho_{gl} < \rho_f$. The number of vehicles passed up to that moment is then used and scaled to the time t_{OD}, to obtain accurate traffic intensities for all links.

Step 7: Set remaining signal timings to accommodate flow on perimeter

After the VISSIM simulation has been run once and the data has been processed, the resulting traffic intensities of all links are used to update the remaining signals of each intersection, to match the traffic demand at the uncontrolled turns.

This is done in order to prevent any further spill-back from the perimeter into the subnetwork, as a consequence of insufficient outflow capacity at the other traffic signals on the perimeter, the intensities of the uncontrolled turns are increased until a cycle time t_{cyc} close to the maximum allowed cycle time $t_{cyc,max}$ is found.

In doing this, the capacity of each turn should become higher than the actual traffic demand, meaning that vehicles using the uncontrolled turns, can easily traverse the intersection without causing spill-back onto the arterials or into the subnetworks, as a consequence of insufficient green time at the intersection, and thus not adversely influencing the MFDs for the subnetwork and perimeter.

Step 8 – 10: Run simulation, process data and adjust number of trips in the network

When the definitive signal schemes of all the intersections have been found, the simulations are run in the same way as within the previous algorithm (paragraph 5.4.1). The difference with the previous algorithm is, that only the trips that have an origin and/or destination within the subnetwork under investigation are changed. Apart from that, the method with the changing OD-factor is kept exactly the same. The simulations are ended when $|OD_{incr}| < OD_{prec}$.

Step 3 to 10 are repeated for the different control scenarios (different q_{ctrl}). After all simulation for the different control scenarios for the inflow have been run, the complete process is repeated for the controlled outflow. All of these simulations are carried out for each subnetwork in each network.

5.5 Data processor

5.5.1 General

During the simulations, a substantial amount of data is gathered. The data gathered for links consists of flow, density, speed and cumulative number of vehicles passed. Data gathered for vehicles consists of the simulation time and vehicle numbers. This data must then be converted into MFDs and performance indicators to support the analysis in chapter 1 and 8.

5.5.2 Creating MFD data

When the simulation is run, it is paused at fixed intervals and data is either directly read by Matlab from VISSIM using the COM-interface, or data is written to a column-based text-file which is converted to data after the simulation has ended. In any case, the resulting dataset is a matrix

containing the flow, speed, density and cumulative number of vehicles for every link in the network at every time interval.

MFDs are created for the total network, the arterials, the combined subnetworks, individual subnetworks and perimeters[17].
The production and the accumulation are calculated by summing up all the measurements of all the links of the particular part of the network, multiplied by the link length, for each time interval, as given by equation (3.15) and (3.17) respectively. The data is not multiplied by the number of lanes of each link, as VISSIM already aggregates the flow and density over all the lanes for each link.
To obtain the performance and the average network density, the production and accumulation are divided by the network length, in which the number of lanes are discounted. The equations used are respectively (3.16) and (3.18).
The deviation of density is simply obtained by calculating the standard deviation of the link densities of each part of the network, conform equation (3.14).

What type of MFDs are created from this data depends on the simulation scenario and is discussed in section 6.2.

5.5.3 Creating vehicle data

During the simulation, the number of each vehicle and the corresponding simulation time are obtained at an interval of $t_{data,veh}$. With only this data, all the vehicle data needed can be calculated. The data below is used to assess the quality with which traffic is processed in each of the simulations.

- **Total number of vehicles to be loaded onto the network (veh):** Summation of all trips in the OD-matrix;
- **Total number of vehicles loaded onto the network (veh):** Summation of unique vehicle numbers in the dataset;
- **Total number of vehicles reaching destination[18] (veh):** Subtraction of the number of vehicles left in the network at the last timestep, from the total number of vehicles loaded;
- **Total time spent in the network (s):** Summation of the difference between the start time and the end time of each vehicle. Vehicles left in the network are assigned the time the simulation ended;
- **Average travel time (s):** Total time spent in the network divided by the number of vehicles loaded onto the network;
- **Average travel time of vehicles reaching destination (s):** Average travel time of each vehicle that has exited the network before the last simulation step;
- **Average free flow travel time (s):** Median of the travel time of the first 1.000 vehicles that were loaded onto the network. This is done under the assumption that the first 1.000 vehicles experience free flow, as the network is empty at the start of the simulation;

[17] An arterial link is added to a subnetworks perimeter, if the distance of both its start- and end node, to any point of that subnetwork is less than 300 metres.
[18] Also referred to as 'network output'

- **Total delay (s):** First the total free flow travel time is determined by multiplying the average free flow travel time by the number of vehicles loaded onto the network. This value is then subtracted from the total time spent in the network;
- **Average delay (s):** Total delay divided by the number of vehicles loaded onto the network;
- **Average delay of vehicles reaching destination (s):** Average travel time of vehicles reaching destination minus the average free flow travel time.

Some notes regarding delays:
The average delay is not a completely reliable value, as this only show the average delay that has been experienced up to the moment the network has ended in gridlock. Actual delays are therefore much higher.

Another problem that makes the average delay and the average delay of vehicles reaching destination unreliable, is the fact that within VISSIM, the start time given to a vehicle is the moment it actually enters the network. Normally this does not pose a problem, but in case the network is congested and congestion spills back to the origins, the vehicles cannot be loaded. This means that the actual start time of the vehicle is earlier than the time assigned to it by VISSIM and therefore the resulting delay should be higher. No option in VISSIM is available to correct for this. This problem can be circumvented by extending the length of the links at the feeders. Although this can be done for the feeders at the edge of the network, this is not possible for the feeders within the subnetworks, as the available space is insufficient.
Another method that has been tested to obtain more accurate delay times, is to assume VISSIM wants to load vehicles on the network linear over time. Then by dividing the number of trips of each origin over the period that vehicles should be loaded, an assumed start time for every vehicle can be calculated. These assumed start times can then be assigned to each vehicle, in the order as they have left the origin. However, it has been found that the actual loading pattern can significantly differ from a linear pattern and thus resulting in wrong values for delays.

In light of the above, it is concluded that the average vehicle delay, as derived from the VISSIM data, is an unreliable value. Nevertheless, this value will still be presented, as it does give some insight in the quality of the network. However, the vehicle delays should not be used to compare different networks or strategies to each other, without the support of other performance indicators.

A note regarding total time spent and total delay
Although the total time spent and the total delay in the network are calculated, these are only used as an auxiliary value and not to assess the network quality. The reason why these values should not be used, is because the amount of traffic loaded onto the network differs for various simulations. Using the total time spent and the total delay can create the impression that traffic improves, when actually the number of vehicles is reduced. Using the average travel time and average delay is therefore much more appropriate.

6 Simulation setup

In the previous chapter, the general building blocks that should answer our research questions have been presented. In this chapter we will further elaborate on how these building blocks are used within the simulation. To this end, section 6.1 will present the different networks that will be simulated and section 6.2 will present the parameters used for the simulation algorithms and the type of MFDs and data that will form the output of the simulations. In order to obtain this data, a smoothly running simulation is of the outmost importance. However, to obtain a smooth running simulation, some settings in VISSIM are changed, which are discussed in section 6.3.

6.1 Networks

To assess the impact of network structure and signal settings on the MFD, 7 different networks are created, derived from the networks presented in section 4.4. The size and shape of each of these networks has been chosen in such a way that the resulting MFDs can be compared to one another and differences between them can be properly attributed to a specific parameter. The different networks are constructed using the network creation model and are shown in Figure 6.1. All networks are $w_{nw,x} = w_{nw,y} = 3\,km$, in which the distance between every intersection is $d_{int-int} = 500m$. The speed of the arterial links is set to $u_{max,art} = 50\ km/h$ and for the subnetwork to $u_{max,sn} = 30\ km/h$.

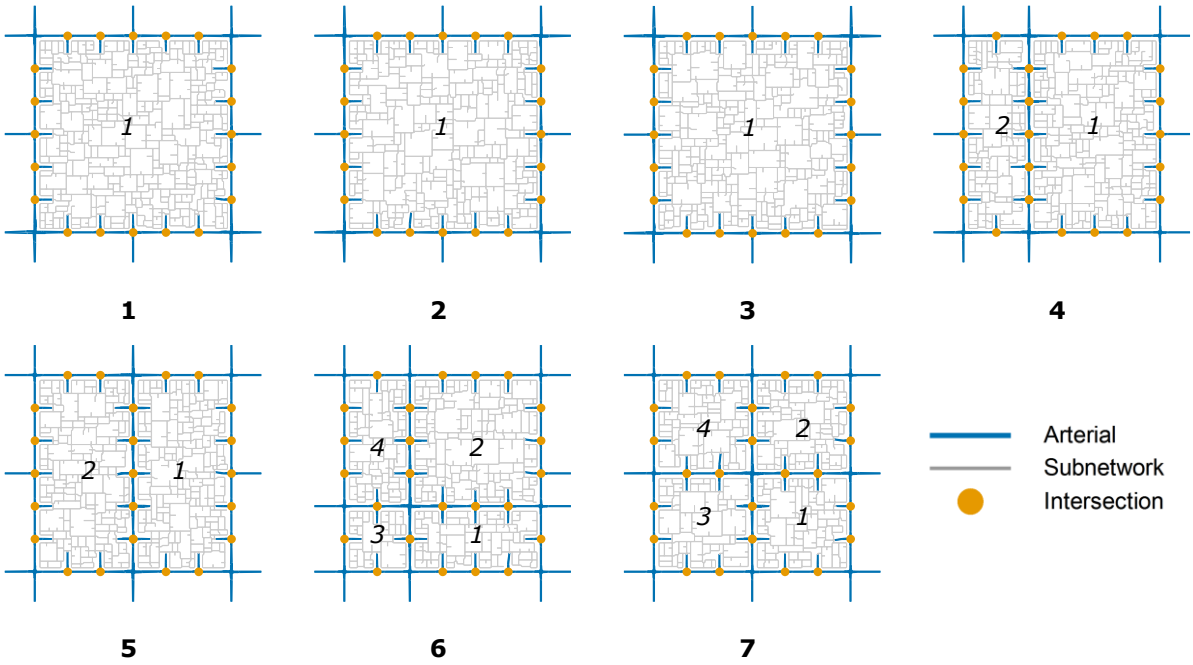


Figure 6.1: *Different networks used (including numbering of network and subnetworks)*

The layout of the subnetworks is generated at random, but is based on the same set of parameters. The number of trips between each origin and destination is drawn from the same

uniform distribution. The resulting OD-matrix is then scaled, to generate a total traffic demand of $n_{trips,tot} = 30.000\ veh$ for all networks, which are loaded onto the network over a period of $t_{OD} = 7.200\ s$.

As the effect of different signal timings has to be tested, it is preferable that all intersections on the perimeter are signalised. Because of this, the value for α_{SLOP} in the Slop intensity criteria (as described in step 3.1 of Appendix A) has been set to zero.

In Table 6.1 and Table 6.2 the different parameters for each network, respectively subnetwork are shown.

Table 6.1: *Parameters for the different networks*

Network	Total arterial length (km)	Total lane length of arterial (km)	Total lane length of subnetworks (km)	Number of intersections (-)
1	18,0	88,97	248,23	24
2	18,0	90,05	222,07	24
3	18,0	89,50	228,56	24
4	21,0	98,66	237,24	29
5	21,0	96,24	246,09	29
6	24,0	101,72	231,08	33
7	24,0	102,11	229,99	33

Table 6.2: *Parameters for the different subnetworks*

Sub-network number	Length (km)	Width (km)	Area (km2)	Total perimeter length (km)	Total lane length of perimeter (km)	Number of inter-sections (-)	Total lane length of subnetwork (km)
1.1	3,0	3,0	9,0	12,0	81,22	20	248,23
2.1	3,0	3,0	9,0	12,0	82,36	20	222,07
3.1	3,0	3,0	9,0	12,0	81,75	20	228,56
4.1	3,0	2,0	6,0	10,0	62,62	16	162,22
4.2	3,0	1,0	3,0	8,0	56,95	12	75,02
5.1	3,0	1,5	4,5	9,0	59,88	14	126,22
5.2	3,0	1,5	4,5	9,0	57,81	14	119,87
6.1	1,0	2,0	2,0	6,0	38,42	8	51,73
6.2	2,0	2,0	4,0	8,0	49,99	12	101,89
6.3	1,0	1,0	1,0	4,0	26,20	4	23,40
6.4	2,0	1,0	2,0	6,0	36,60	8	54,06
7.1	1,5	1,5	2,3	6,0	39,00	8	58,48
7.2	1,5	1,5	2,3	6,0	38,17	8	62,68
7.3	1,5	1,5	2,3	6,0	38,48	8	50,73
7.4	1,5	1,5	2,3	6,0	37,35	8	58,10

6.2 Simulation scenarios

6.2.1 Scenario 1: The effect of network structure on the shape of the MFD

The algorithm used to evaluate the effect of network structure on the shape of the MFD is described in paragraph 5.4.1. The parameters used in the simulation and the MFDs and data generated are presented below.

Simulation parameters

To obtain a sufficient amount of data, the simulation is run 5 times at every step/OD-factor, i.e. $n_{seeds} = 5$. It has not been tested whether more simulations could improve the results and if 5 runs is fully sufficient to obtain all the data. However, due to time restrictions, no more than 5 runs per OD-factor have been chosen. Nevertheless, for every network it took at least 3 and at most 5 steps to get to the gridlock-point, resulting in a total of 15, respectively 25 simulations for each network.

All traffic is loaded onto the network over a period of $t_{OD} = 7.200s$ and the simulation is ended if no vehicle is in the network, when the end of the simulation time $t_{sim,tot} = 14.400s$ is reached, or the gridlock threshold to $\rho_{gl} = 90\%$ is passed.

The evaluation intervals, i.e. the interval with which data for links and vehicles is obtained from the simulation is set to $t_{data,i} = 300s$ and $t_{data,veh} = 10s$. This results in a minimum of 120 points in the MFD for every simulation step/OD-factor.

The scaling and increment factors used to obtain the gridlock point are initially set to $OD_{fact} = 0.5$, $OD_{incr} = 0.5$ and $OD_{prec} = 0.1$ at the start of every simulation.

Output

The data obtained from each of the simulation (140 in total) is converted into MFD data, as described in paragraph 5.5.2. MFDs are created for the arterials, the combined subnetworks and the total network. For each of these network parts, the PD-A, PF-K, PD-K and PF-A MFD are created, which are added in Appendix B. An example of one of the resulting MFDs is shown in Figure 6.2.

As the deviation of density is used on the z-axis, PD-S, PF-S, A-S and K-S MFDs are generated as well. All of these MFDs are added digitally as Appendix C.
Furthermore the data has been used to reconstruct the final traffic state of each simulation, in order to assess where congestion is located. These are added in Appendix C as well.

In order to compare the results of different simulations to each other, the maximum performance and optimal accumulation are derived for each MFD constructed.

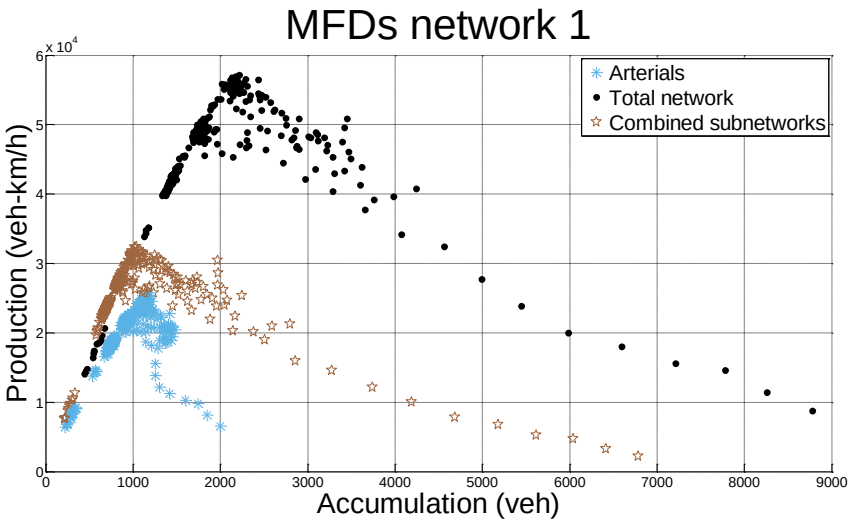


Figure 6.2: *Example of a macroscopic fundamental diagram of network 1*

Regarding the vehicle data, all data as described in paragraph 5.5.3 is generated. This data is added to the figures portraying the final traffic states, which are added in Appendix C. In the analysis in chapter 7, only the number of vehicles loaded onto the network, the average free flow travel time, the average travel time of vehicles reaching their destination, and the average delay of vehicles reaching their destination is presented.

6.2.2 Scenario 2: The effect of signal settings on the subnetwork and perimeter

A description of the algorithm used to evaluate the effect of signal settings on the MFD of a subnetwork and its perimeter is given in paragraph 1.1.1.

Simulation parameters

Apart from the number of seeds ($n_{seeds} = 1$), the same parameters as in the previous scenario are used. To assess the effect of different signal settings on the subnetwork and the perimeter MFD, the timings of the signals controlling the inflow to the subnetwork from the perimeter are set to a fixed value, in order to accommodate a specific number of vehicles per hour. The timings used are $q_{ctrl} = [100; 200; 300; 400; 500]\ veh/h$. When controlling the inflow, the maximum cycle time is set to $t_{cyc,max} = 240s$ and when controlling the outflow $t_{cyc,max} = 150s$. This is done so a viable signal scheme for every intersections can be generated, for each of the control scenarios. Especially the cycle times for the controlled inflow can be high, as all three controlled streams conflict with each other.

After the signals schemes have been calculated, all trips not originating from, or destined for the subnetwork are reduced to $\delta_{red} = 10\%$, in order to obtain 'clean' MFDs that are not affected by traffic that have no relation to the subnetwork under evaluation. The point at which the network goes from free flow to gridlock, is set to $\rho_f = 20\%$.

Output

The data obtained from each of the simulation (703 in total) is converted into MFD data, as described in paragraph 5.5.2. MFDs are created for each subnetwork and perimeter. Separate MFDs are constructed for controlled inflow and controlled outflow. In each of the MFDs, a distinction is made between the results at each different signal timing.
For these MFDs, the measurements of the first run at each signal setting are discarded. This is done, because the first simulation is used to obtain data from the network after fixing the timing of the controlling signals. This data is used to calculate new signal schemes. As these signal schemes are substantially different, this has an impact on the MFD. To single out the effect of the different signal timings at the controlling signals in the MFD, the first simulation run has to be discarded.

Again, the PD-A, PF-K, PD-K and PF-A MFD are created, which are added in Appendix C. The PF-K MFDs are added in Appendix B. An example of one of the resulting MFDs is shown in Figure 6.3. The MFDs showing the relation with any of these parameters to the deviation of density, and the aforementioned vehicle data are calculated and added to Appendix C as well.

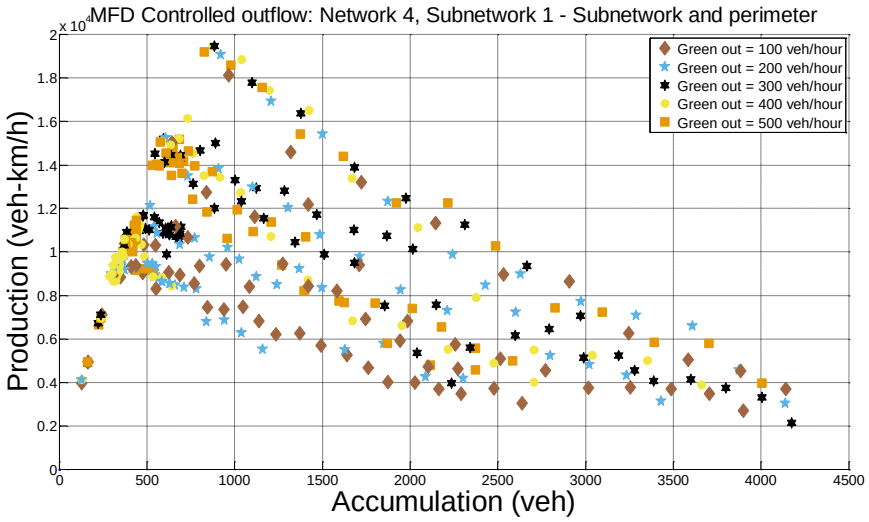


Figure 6.3: *Example of a macroscopic fundamental diagram for controlled subnetwork outflow*

6.3 Simulation settings

Apart from the different simulation settings described in the previous paragraphs and incorporated in the network creation model, additional parameters have been changed, in order to obtain a smooth running simulation. In the subsequent paragraphs the most important settings are discussed.

6.3.1 Driving behaviour

The car following model used in the VISSIM simulations is the Wiedemann-model. In this model different parameters for the driving behaviour of vehicles are incorporated. The basic principle is that the speed of each vehicle depends on the speed of the preceding vehicle(s). Within the model, four different driving modes are used (from PTV Vision, 2009):

- **Free driving**: No influence of preceding vehicles;
- **Approaching**: Applying a low deceleration rate due to the preceding vehicle driving at a lower speed;
- **Following**: Maintaining speed of predecessor, taking the safety distance into account;
- **Braking**: Applying a medium to high deceleration rate due to the safety distance falling below the desired distance.

The resulting acceleration and deceleration for each of these modes is described by a different set of formulations, based on speed, speed difference, distance and individual characteristics of driver and vehicle.
VISSIM offers two car following models, Wiedemann 74 and Wiedemann 99. Within this thesis the Wiedemann 74 model is used, as this is the most suitable for urban traffic (according to PTV Vision, 2009).
In the Wiedemann 74 model, a number of parameters can be changed in order to tune the capacity flow of the links in the network. These parameters and their values are:

- Average standstill distance ($ax = 2{,}00$);
- Additive part of safety distance ($bx_{add} = 2{,}00$);
- Multiplicative part of safety distance ($bx_{mult} = 3{,}00$).

As illustrated by PTV Vision (2009), using the speed distribution of 50 km/h, these values should lead to a capacity flow of $q_c \approx 2000 \frac{veh/h}{lane}$, which is a generally used value for urban networks.

The last parameter that has been changed is the minimum look ahead distance, which has been changed from 0,00 to 50,00 meter. Although not always realistic, the adaptation of this parameter improves the simulation, as vehicles better anticipate to a merging or diverging point (which circumvents some simulation problems).
Apart from the minimum look ahead distance, none of the other parameters affecting driver behaviour have been changed and the parameters used by VISSIM have been adopted. This seems justified, as fictional networks are used and therefore no calibration is necessary.

6.3.2 Vehicle compositions

Another important change that has been made to the basic simulation parameters, is the composition of vehicles. Although incorporating different vehicle types improves the realism of the simulation and could also impact the shape of the MFD, all non-car vehicles (trucks, busses, etc.) are removed.
This is done in order to simplify the simulation and network coding. However, the most important reason is that these larger vehicles can get 'stuck' in certain corners, leading to gridlock in the network as they cannot move. This can be solved by making local changes to the network, but as the model is controlled by Matlab, this has shown to lead to errors in the model, due to incorrect link lengths or non-uniform coding of links and connectors. Although the problem seems to have been solved, these vehicles are still not used, as the risk of this happening is still present, meaning that the simulation outcomes would be biased and incorrect MFDs are presented.

6.3.3 Path choice and assignment

Little changes have been made to the path choice and assignment in VISSIM, as this is all readily available in VISSIM and saves a substantial amount of programming time. In VISSIM the travel cost for all 'edges' (a set of links between zones, or in a zone) is calculated, and the cost of every path $M_{pt,r,s}$ is a summation of these individual costs, which is the inverse of the total utility UT_{pt} of that path. The travel costs are computed from expected travel times $M_{VIS,ett}$, travel distances $M_{VIS,td}$, financial costs $M_{VIS,fin}$ and a general supplement $M_{VIS,sup}$ for each edge. The total costs per 'edge' are then calculated as

$$M_{VIS,edge} = M_{VIS,ett} t_{VIS,edge} + M_{VIS,td} l_{VIS,edge} + M_{VIS,fin} m_{VIS,costs} + \sum M_{VIS,sup}. \quad (6.1)$$

No additional research into these costs have been done and the standard factors of $M_{VIS,ett} = 1{,}0$, $M_{VIS,td} = 0{,}0$ and $M_{VIS,fin} = 1{,}0$. $M_{VIS,sup}$ cannot be changed.

After the utility for each path has been calculated, traffic is assigned to each path pt within the the set of routes R using the Kirchhoff distribution formula

$$p(R_{pt}) = \frac{UT_{pt}^{\mu}}{\sum_j UT_{pt}^{\mu}}. \quad (6.2)$$

During the course of the simulation, travel times change, changing the utility of each of the routes found. To this end, every $R_{sim,upd,\Delta t} = 600\ s$ new routes are searched during the simulation, supplementing the existing set of paths, which has been initialised during the creation of the network, as described in step 6 of paragraph 5.3.2 and/or Appendix A. In order to speed up the simulation, the total number of paths between each origin and destination is limited to $n_{pt,r,s} = 50$, which is still very substantial. The total travel costs of each path cannot exceed $M_{pt,r,s} \leq \alpha_{det}\, min(B_{r,s})$ in which $\alpha_{det} = 1{,}5$ and $B_{r,s}$ is the set containing all $M_{pt,r,s}$.

6.3.4 Warm-up period

It should be noted that it has been chosen not to use a warm-up period for the simulation. Although this is generally preferred, to obtain more homogeneous traffic conditions and obtain better results, this has not been added in the model. Due to the current complexity of the model, adding a warm-up period was omitted because of time restrictions. If the network creation model is used for future research, it is recommended to add such a feature.

Nevertheless, not adding a warm-up period should not have a highly adverse effect on the simulation results, as the density of origins and destinations is very high in the network, meaning that within a short period of time (t≈900 s) the network is loaded to normal conditions. As such, most of the data points obtained are derived from normal states.

Apart from that, not adding a warm-up period also clearly shows the effect of the rush, one of the more interesting phenomena found in the MFD and something of specific interest.

7 Effect of network structure on the macroscopic fundamental diagram

This chapter will start by presenting the results of the simulation in section 7.1, a first assessment of the different MFDs will be made in section 7.2, discussing general differences in shapes, and differences between MFD types.
Next the analysis of the effect of different network structures on the MFD is broken up in different paragraphs, each investigating a separate effect, based on the hypotheses made in section 4.6. First, the effect of changing the network structure by adding additional arterials is investigated in section 7.3, after which the effect of the placement of the arterials is presented in section 7.4. As it is shown that the shape of the MFD could be seriously affected by changes to traffic and network loading patterns, this effect is investigated in more detail in section 7.5. Then in order to investigate whether or not similar networks produce similar MFDs, the first three networks are compared to each other in section 7.6. The same is then applied to the different subnetworks and perimeters of the remaining networks in section 7.7, where it is investigated if the MFDs of the different subnetworks and perimeters are actually scalable. Section 7.8 will end with a conclusion summing up the different insights gained from the various analyses made in this chapter.

7.1 Simulation results

7.1.1 Presentation of simulation results

As it has been found in chapter 3, that a reasonable comparison can only be made when the different MFDs are taken into account, the four different types of MFDs for each network are presented in Appendix B. In these figures, the MFD of the total network, the arterials and the combined subnetworks are shown.

The data points in each of the MFDs are derived from multiple simulations runs, in which the amount of traffic and the traffic patterns were different for each simulation, as has been explained in paragraph 6.2.1. Although the amount of traffic might severely impact the shape of the MFD, as we have seen in paragraph 3.2.5, this will not be taken into account in the analysis within this section, as this would result in too much different parameters to make a proper analysis. The effect of the different traffic patterns is nonetheless covered in section 7.5.

7.1.2 Performance indicators

In order to be able to assess the changes in the MFDs, Table 7.1 shows the production, performance, optimal and maximal accumulation and the optimal average network density for the arterials, the combined subnetworks and the total network.

Table 7.1: *Performance indicators regarding the MFD*

Network	Production (veh-km/h)	Optimal Accumulation (veh)	Performance (veh/h)	Optimal Network density (veh/lane-km)	Maximum accumulation (veh)
Arterials					
1	25.771	1.190	436	20,1	2.272
2	26.469	1.217	444	20,4	2.404
3	26.720	1.271	451	21,5	2.292
4	13.209	696	178	9,4	1.356
5	13.813	555	185	7,4	1.041
6	15.687	752	180	8,6	1.335
7	15.117	731	176	8,5	1.196
Combined subnetworks					
1	32.410	1.036	131	4,2	7.210
2	32.190	1.188	145	5,3	5.972
3	31.229	973	137	4,3	5.576
4	23.146	922	98	3,9	6.306
5	23.791	1.009	97	4,1	6.316
6	24.868	1.049	108	4,5	6.840
7	23.804	1.026	103	4,5	6.366
Total network					
1	57.167	2.214	186	7,2	9.972
2	58.038	2.377	206	8,4	9.107
3	57.661	2.242	200	7,8	8.618
4	36.100	1.477	116	4,7	7.614
5	36.931	1.423	115	4,4	7.836
6	40.233	1.805	126	5,7	8.576
7	38.813	1.839	123	5,8	7.730

In Table 7.2, the number of vehicles, average travel time and average delay are given for each network. This data is derived from the highest OD-factor, in which the network did not end up in gridlock and was able to fully serve all traffic loaded onto the network. As set forth in paragraph 6.2.1, at every OD-factor, 5 simulations were run using different seeds. The result of this was that in some of the networks, at the same OD-factor, the network ended up fully empty in one simulation, and in gridlock in another simulation. In this case it has still been chosen to use the data from this OD-factor, but use the average of the simulations (seeds) that did not end in gridlock, instead of all the simulations. In order to keep the results transparent, the number of times that the network did not end up in gridlock, is mentioned as well.

Table 7.2: Performance indicators regarding travel times

Network	OD-factor	Number of simulations not ended in gridlock	Number of vehicles loaded (veh)	Average travel time (s)	Average free flow travel time (s)	Average delay (s)
1	0,88	5 / 5	22.499	453	420	33
2	1	2 / 5	30.030	553	453	100
3	1	2 / 5	29.989	518	450	68
4	0,5	4 / 5	15.003	404	313	92
5	0,5	2 / 5	14.999	386	320	66
6	0,5	2 / 5	15.002	431	310	97
7	0,63	5 / 5	18.923	339	301	38

7.2 First assessment and general observations

7.2.1 The shape of the MFD

The first thing that is noticed when looking at the MFDs of the different networks, is that the shape of the MFDs of the base networks (the first three networks) differs substantially from the MFDs of the other networks, which have additional arterials and multiple subnetworks. In the MFDs of the first three networks, almost no scatter is present and the transition from free flow to congestion is very sharp. In the MFDs of the other networks a lot more scatter is found and the transition is less well-defined and more rounded.

It has been found that the scatter in the MFD is caused by the fact that certain parts of the network are in free flow, while others are congested, which tells that the traffic is distributed unevenly over the network. The amount of scatter found in the MFDs of the last four networks thus shows that traffic is distributed unevenly, causing the more rounded shape in the MFDs. In the first three networks, traffic conditions over the network seem to be very homogeneous resulting in the clearly defined transition, implying that the network is close to its capacity.

Another thing that can be noticed in almost every MFD, is that the congested branch of the MFD is not decreasing monotonically as expected, but first starts with a steep descent, which generally starts levelling out when the network is coming closer to the maximum accumulation. An illustration of this is given in Figure 7.1. An explanation for this effect is that at higher accumulations more links are in the congested state, which decreases the amount of scatter, resulting in the MFD coming closer to its 'real' (outer) shape. Another possibility is that gridlock seems to occur at densities of 25 veh/lane-km, while the expected maximum density should be around 100 veh/lane-km. An explanation why the network is in gridlock at these low

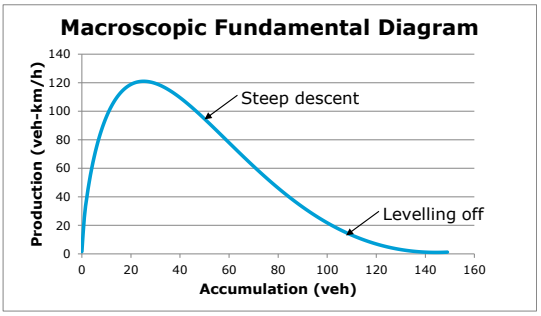


Figure 7.1: Example of a macroscopic fundamental diagram levelling off at higher accumulation

densities could be, because only a part (≈25%) of the network is completely congested, while the remaining part is almost completely empty.

Using different MFD types

As has been illustrated in chapter 3, using a different type of MFD, changes how the quality of a network is perceived. Looking at all the different MFDs, the effects described in chapter 3 are found for all networks. Especially the perceived difference between the arterials and the subnetworks is clearly present for all networks. Whereas the PF-K MFDs all show that the performance and the average network density of the arterials score higher, the PD-A MFDs show that the production and accumulation are often higher in the subnetworks than in the arterials. The exception to this are the first three networks, in which the maximum production and optimal accumulation of the subnetworks and the arterials are roughly equal.

7.3 Adding additional arterials

7.3.1 Hypothesis

In paragraph 4.6.1 it was hypothesised that adding additional arterials to the network, thus creating multiple subnetworks should result in a higher production and maximal accumulation of the arterials due to the increased lane length, while lowering the performance and the optimal accumulation due to an increase of the average number of conflicts, i.e. the number of intersections per kilometre of arterial. Also the deviation should increase, as a result of a heterogeneous distribution of traffic over the arterials, leading to a highly scattered MFD.
For the subnetwork it is expected that all of its parameters should decrease, except for the deviation of density, for which no prediction could be made.
All in all, the production, performance, accumulation and average network density are all expected to decrease for the total network. The only expected increase should be the deviation of density, but this also has a negative impact on the MFD.
It should however be noted that this does not directly mean that adding additional should not be done, as it is assumed to shorten travel times, which is positive for the road user. However, this effect cannot be obtained from the MFD.

7.3.2 Analysis of the results

From the different MFDs presented in Appendix B and the performance indicators in Table 7.1 it can be found that almost none of the parameters improves when additional arterials are added. At first glance, all of the parameters appear to be lower than in the base networks. In order to quantify the effect, Table 7.3 shows the relative difference of the different parameters, in relation to the mean of the three base networks. The results for production and optimal accumulation are not compensated for differences in network length.

From the table it can be found that every parameter scores lower than in the base networks. The only parameter that is not severely affected is the optimal accumulation and network density of the subnetworks.

Table 7.3: *Performance difference of networks 4 – 7 in respect to networks 1 - 3*

Network	Production	Optimal Accumulation	Performance	Optimal Network density	Maximum accumulation
Arterials					
4	50,18%	56,73%	40,13%	45,36%	58,36%
5	52,48%	45,30%	41,62%	35,93%	44,83%
6	59,60%	61,36%	40,60%	41,80%	57,46%
7	57,43%	59,63%	39,67%	41,18%	51,48%
Combined subnetworks					
4	72,46%	86,50%	71,01%	84,59%	100,85%
5	74,48%	94,73%	70,37%	89,31%	101,01%
6	77,85%	98,43%	78,33%	98,82%	109,40%
7	74,52%	96,34%	75,33%	97,18%	101,80%
Total network					
4	62,65%	64,85%	58,70%	60,72%	82,47%
5	64,09%	62,47%	58,28%	56,77%	84,88%
6	69,82%	79,24%	64,04%	72,63%	92,89%
7	67,36%	80,76%	62,22%	74,55%	83,73%

Arterials

Maximum production (lower): It was hypothesised that due to the increased length of the arterial, the maximum production should increase. However, the production is adversely affected, and almost half that of base networks. Interestingly enough, it cannot be stated that adding arterials always has a negative effect on the production, as network 6 and 7 seem to be less effected than network 4 and 5.

Maximum performance (lower): It was hypothesised that the maximum performance would decrease due to the increased conflict ratio. Although the performance has indeed decreased, this cannot be attributed to this factor, as the conflict ratio of network 4 and 5 is only 10% higher than the base networks and for network 6 and 7, this is only 20%. Also the fact that almost no difference between the 4 networks is present, seems to contradict this assumption.

Maximum accumulation (lower): The maximum accumulation was expected to increase, due to an increase in the length of the arterials. However, the maximum accumulation has actually been found to decrease by almost half, while the length of the arterial has increased by 7 – 14%. A reasonable explanation for this is that due to the increased deviation of density, more and larger 'pockets' (empty space) arise in the arterial, lowering the maximum accumulation. Nevertheless,

the maximum accumulation should unequivocally be bounded to the total length of the network, meaning that the MFD obtained is far beneath the boundaries of the 'real' MFD.[19]
In order to investigate this a bit more in-depth, Table 7.4 shows the maximum average density at the maximum accumulation. Interestingly, this value is quite constant and between 25 – 30 veh/km. The only exception is formed by the arterials of network 4 – 5, which are between 10,8 – 13,7 veh/km. Again implying that adding additional arterials is not beneficial to the network.

Table 7.4: *Maximum average network density at maximum accumulation*

Network	Arterial	Subnetworks	Total network
1	25,5	29,0	29,6
2	26,7	26,9	29,2
3	25,6	24,4	27,1
4	13,7	26,6	22,7
5	10,8	25,7	22,9
6	13,1	29,6	25,8
7	11,7	27,7	23,3

Also interesting is the fact that the maximum average network density found seems to be close to the critical density. This would mean that the network would reach full gridlock if the average critical density is transgressed. Whether or not this is actually the case, is unclear, but it could be hypothesised that when the average critical density is transgressed, shockwaves start spreading through the network, causing the system to always become unbalanced, which given enough time, will inevitably end up in gridlock.

Optimal accumulation and average density (lower): The optimal accumulation and the optimal average density were expected to decrease, as a result of the increasing conflict ratio. Even though the outcome is correct, it seems that the effect cannot be attributed to differences in the conflict ratio, as has been shown for the performance.

Deviation of density (higher): It was hypothesised that the additional arterials would result in a higher deviation of density, as the central arterials should attract more traffic than the outside arterials. The amount of scatter found in the MFDs of network 4 – 7 shows that the deviation of density has strongly increased, as is confirmed by Figure 7.2 and Figure 7.3.

[19] Assuming a maximum density of 100 veh/km, the maximum accumulation should be around 10.000 vehicles for any of the arterials, showing that in none of the cases the maximum accumulation is even remotely reached

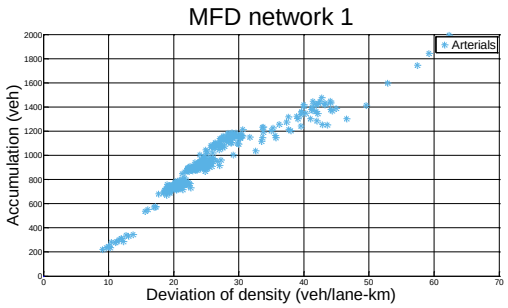


Figure 7.2: *Deviation of density vs. accumulation network 1*

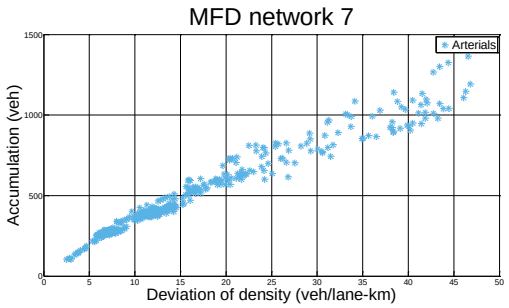


Figure 7.3: *Deviation of density vs. accumulation network 7*

Subnetworks

Maximum production (lower): It was assumed that due to the increase of the arterials, the length of the subnetworks should lower, resulting in a lower production. However, as it turns out, the actual difference in the size of the subnetworks is negligible, and should not influence the production. Nevertheless, it was also hypothesised that due to the increased amount of traffic signals, traffic is hindered more, resulting in a lower production. It is found that the production of the subnetwork has indeed decreased, by approximately 25%. But as no substantial difference between network 4 - 5 and 6 – 7 is found, it cannot be said with certainty that the decreased production can be fully attributed to the additional traffic signals, because if this should be the case, the maximum production of network 6 and 7 should be lower than for network 4 and 5

Maximum performance (lower): The maximum performance of the subnetworks is expected to decrease due to a reduction in the bottleneck capacity, caused by the additional intersections. The reduction found in the performance is indeed lower and roughly the same as the production at 25%. As with the production, it is not certain if this can be attributed to the additional intersections.

In order to test how the traffic signals could affect the performance of the subnetworks, the total amount of green time given to the traffic signals regulating the outflow from the subnetworks is calculated for the different signals. To this end, the green time of the signal with the shortest amount of green time per hour is determined for each intersection regulating the different subnetworks, after which the time is totalled over all subnetworks. The result of this is shown in Table 7.5, which also shows the difference with the average green time given in the base networks. The table shows that the total amount of green time given to all vehicles in the subnetworks is actually larger than for the base networks, meaning that traffic should actually be able to enter the arterial more easily.
However, it should be noted that traffic which initially had an origin and destination on either side of the subnetwork, was not affected by the traffic signals in the base networks. Due to the additional arterial, this traffic is affected which could lead to a decrease in the production.

Table 7.5: *Green time give to outflow of subnetworks*

Network	1	2	3	4	5	6	7
Green time (s)	16.162	16.011	15.895	21.177	20.761	22.108	22.679
Difference in green time (%)	-	-	-	132%	130%	138%	142%

Maximum accumulation (equal): Due to a decrease in the subnetwork length, the maximum accumulation was expected to decrease. However, the subnetwork length has not significantly changed, implying that the accumulation should be unaffected. Looking to the data, this indeed seems to be the case.

Optimal accumulation and average network density (equal): The optimal accumulation is assumed to become lower as a consequence of a decrease in the subnetwork length and an increasing amount of conflicts, due to the added intersections. However, no significant changes were found in either the optimal accumulation and average network density.

Deviation of density (higher): The effect of adding additional arterials on the distribution of traffic in the subnetwork was not hypothesised, as the reduced subnetwork size might benefit from the deviation, while the increased waiting time at signals might cause higher densities along the edge of the subnetwork, leading to larger differences in the density. Looking at the different MFDs, it can be noticed that the amount of scatter has increased, meaning that the deviation of density has increased.

Total network

Maximum production (lower): It was hypothesised that the production should become lower, as it was assumed that the total length of the network was decreased and that additional intersections would increase conflicts and therefore reduce capacity. It has been indeed found that the production has lowered. However, changes in network size are minimal, and the increase of the number of signals are insufficient to explain the reduction of 30 – 40%. No apparent reason for this reduction can be found.

Maximum performance (lower): As the production has lowered and the network length has not significantly changed, the maximum performance should be lower as well. However, the reduction in performance is even 5% higher than the reduction in production, which can be related to changes in the network size. This however does imply that the other 40% reduction are caused by factors other than changes to the network size.

Maximum accumulation (lower): Although the maximum accumulation of the subnetworks has remained almost unchanged, the reduction in the accumulation of the arterials has caused a decrease of the maximum accumulation of 10 – 15%.

Optimal accumulation and average network density (lower): The reduction of the optimal accumulation in the total network has been found to be approximately 40% for the networks with

one additional arterial and only 20% for the networks with two additional arterials. Although the first would seem to imply that adding arterials reduces the optimal accumulation, this is contradicted by the 6th and 7th network, in which the conflict ratio is higher, which was identified in the first place as a factor that should negatively impact the MFD. The only reason that can be found to explain this effect is that due to the single central arterial, traffic concentrated around this arterial, leading to clusters of congestion. This effect should be less apparent in the other networks. In the first three networks, traffic should be equally distributed over the different arterials, while the 6th and 7th network have more central arterials, causing a more even distribution of traffic.

Deviation of density (higher): Adding arterials to the network, causes the traffic to cluster more around the arterials, leading to a more uneven distribution of traffic over the network, adding to the reduction of production, performance, accumulation and network density.

Travel times

Although the MFD has been negatively affected by adding the additional arterials, the travel times have improved. However, the amount of traffic that the network could process was halved for most networks, other than the base networks, which should also have a decent impact on the travel times. Nevertheless, the average free flow travel time was found to be reduced from approximately 450 seconds for the first 3 networks, to 310 seconds in the other 4 networks, meaning a reduction in travel time of approximately 70%, and also the average travel time has decreased, although this ranges heavily between 5 and 40%, depending on which networks are compared to each other, as is shown in Table 7.6. Interestingly enough, adding 1 or 2 additional arterials does not seem to make a large difference in this respect.

Table 7.6: *Reduction of average travel times*

Network	4	5	6	7
1	11%	15%	5%	25%
2	27%	30%	22%	39%
3	22%	25%	17%	34%

7.3.3 Conclusion

Adding additional arterials to the network has not resulted in positive effects in the MFD, as almost all the parameters ended up lower than in the base networks. Only the accumulation of the subnetworks seems to have been unaffected.

The arterials themselves have been found to be the most affected by these changes, as their production reduced by more than 40%, and their performance even by 60%! Changes in the total length are only in the range of 10 – 15%, and cannot explain these immense differences. Also the reduction of the bottleneck capacity due to an increased amount of intersections cannot be the sole reason for this difference, as the conflict ratio on the arterial (lane length / number of intersections) only increased with 10 – 20%. Also the free flow speed has not significantly changed, as can be derived from the angle of the free flow branch of the various MFDs. It is therefore

assumed that the main benefactor to the decrease of the MFD parameters is to be found in the more uneven distribution of traffic over the network.

Nevertheless, the free flow travel times and average travel times have improved after the arterials have been added. Showing that the MFD does not convey any information regarding travel times and the performance of traffic itself. It should however be noted that the number of vehicles that could be processed by the network, without ending in gridlock was only half that of the original networks.

7.4 Asymmetric arterial placement

7.4.1 Hypothesis

It was hypothesised in paragraph 4.5.2, that the way in which the arterials were placed in the network, should not have a significant impact on the shape of the MFD and that only the deviation of density is assumed to increase.

7.4.2 Analysis of the results

Judging from the results in Table 7.1, the hypothesis seems to be valid and no substantial differences are found between network 4 and 5, and network 6 and 7. These differences are shown in Table 7.7.

Table 7.7: *Differences caused by asymmetric arterial placement*

Network	Production	Optimal Accumulation	Performance	Optimal Network density	Maximum accumulation
Arterials					
4 - 5	4,37%	25,22%	3,60%	26,23%	30,19%
6 - 7	3,77%	2,90%	2,35%	1,49%	11,63%
Combined subnetworks					
4 - 5	2,71%	8,68%	0,91%	5,28%	0,17%
6 - 7	4,47%	2,17%	3,98%	1,69%	7,46%
Total network					
4 - 5	2,25%	3,81%	0,71%	6,96%	2,83%
6 - 7	3,66%	1,88%	2,91%	2,58%	10,94%

As can be seen in the table, most differences are not larger than 5% and do not give cause to assume that asymmetric arterial placement has an effect on the MFD. The values that do jump out though, are the accumulation and density of the arterials of network 4 and network 5, which show that the optimal accumulation and network density are 25% higher and that the maximum accumulation is 30% higher. These differences are actually quite significant and cause for further investigation.

The length of the arterial of network 4 and 5 is 98,66, respectively 96,24 kilometre, resulting in a difference of 2,5% and thus not explaining the difference. Other factors involved in the accumulation are the free flow speed, capacity and signal settings. The free flow speed can be excluded, as this is kept the same for all arterial links. Apart from that, it can be derived from the MFDs, that the free flow network speed is roughly equal.
Also the difference in signal timings has been investigated, but the total amount of green given in both arterials was found to differ no more than 1,5%.
Furthermore, a comparison of the deviation of density is made, which is shown in Figure 7.4. From this diagram it can be found that the deviation of density of network 5 is higher at the same accumulation than in network 4, meaning that the traffic in network 5 is distributed more uneven over the network. Although it seems that these differences are not very large, at an average network density of 10 veh/lane-km, the difference in the deviation of density is 37 / 30 ≈ 25%, which seems to match the observed difference.

However, this observation does contradict the hypothesis that the deviation of density should be higher for the network with the asymmetrically placed arterial, as the figure below illustrates that actually the symmetrical network has a higher deviation of density. The underlying reason for this observation is unknown.

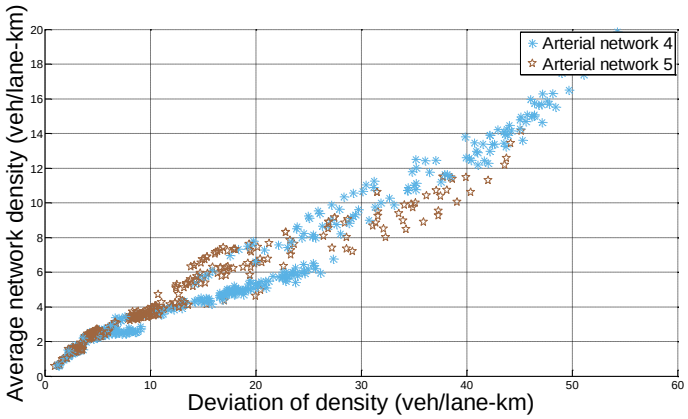


Figure 7.4: *Deviation of density of the arterials of network 4 and 5*

Although most of the differences in the MFD might not be significantly different between both networks, the differences in the average travel times (Table 7.2) do show different results. Especially between network 6 and 7 a significant difference is present, with the average travel time of network 7 being approximately 25% lower than for network 6. Apart from that, the number of vehicles that network 7 can process is also approximately 25% higher, which implies that the network is actually better designed.
Looking at the different traffic patterns of network 6 and 7, it was found that congestion mainly spreads from the intersection between the central arterial. As this intersection is placed in the centre of network 7, traffic distributes more evenly over the different subnetworks and takes longer to reach the outer arterials, keeping these longer in a state of free flow. In Figure 7.5 and Figure 7.6, the gridlock state of both networks is shown, showing that congestion in network 7 is much more contained within the inside of the network.

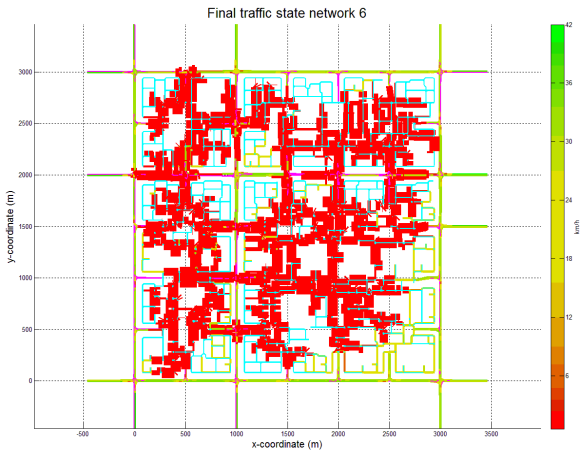


Figure 7.5: *Gridlock state network 6*

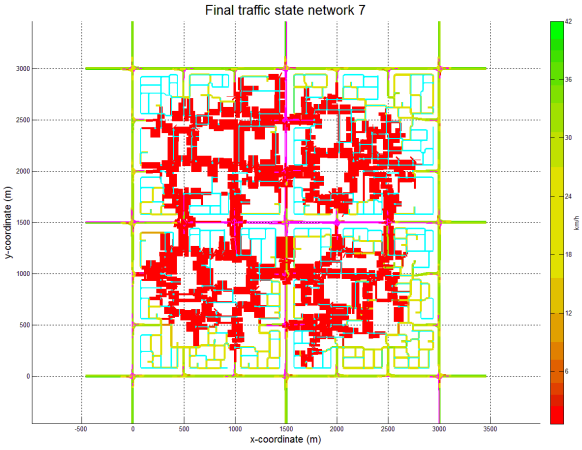


Figure 7.6: *Gridlock state network 7*

The fact that this difference does not seem to be reflected in the MFD is assumed to be caused by the fact that the maximum performance and production are derived from the first simulation run, in which the number of vehicles loaded onto both networks was equal at approximately 30.000, resulting in temporarily higher productions, as bottleneck capacities are not activated yet at the start of the simulation. This phenomenon is explained in paragraph 3.2.5 and is referred to as 'the rush', which will be analysed in more detail in section 7.5.

7.4.3 Conclusion

The asymmetric placement of arterials, does not seem to have a significant impact on the shape of the MFD. From this the conclusion is drawn that the exact physical layout of the arterials in a network does not seem to affect the MFD. This is in line with the hypothesis that the MFD is mainly derived from the characteristics of the individual underlying links. However, some differences in the accumulation for the arterials of the set of networks with a single central arterial were found, and are contributed to differences in the deviation of density. The fact that the deviation of density for the symmetrical network is higher, contradicts the previous hypothesis. The reason for this is not clear and is not confirmed by the other two networks.

Apart from changes in the MFD, the average travel time is investigated as well. It is found that the average travel time in network 4 and 5 is not significantly different. The difference in travel time for network 6 and 7 differs approximately 25%, implying better design of network 7. The reason that network 7 performs better is that due to the more centrally placed intersection between the arterials, it takes longer for traffic to spill-back on the outer arterial, keeping it longer in free flow than for network 6. This effect however is not reflected in the MFD itself.

7.5 Different traffic loading patterns

7.5.1 Hypothesis

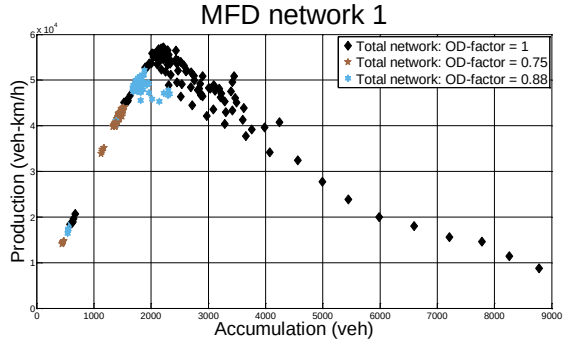
It has been found in paragraph 3.2.5 that different traffic loading patterns might create different MFDs. The reason for this is, that when traffic is loaded onto the network at a high rate, the maximum production and performance are derived from a state, in which the links in the network are at maximum capacity. However, the capacity of each link eventually is restricted by the

capacity of the first downstream bottleneck. When the capacity of these bottlenecks is reached, the maximum capacity of all the links affected by this bottleneck are reduced to this bottleneck capacity (or even less, as multiple links have to share the capacity of the same bottleneck), reducing the overall production. This phenomenon should be mostly witnessed during fast changing demand patterns, which is usual during rush hours. As such, we have dubbed this phenomenon 'the rush'. The hypothesis made, is that the maximum production and optimal accumulation should be roughly linearly related to the number of vehicles loaded onto the network per unit of time.

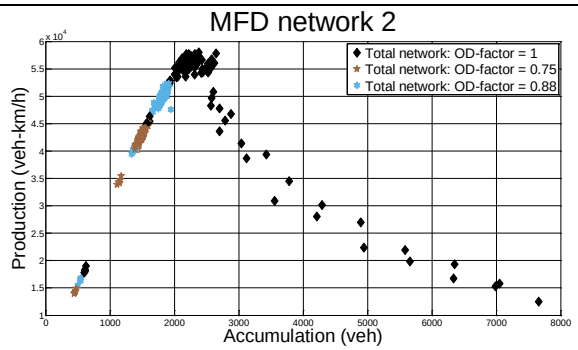
7.5.2 Analysis of the results

In Table 7.8 below, the MFDs of the total network, for the different loading patterns are given, including the maximum production, optimal accumulation and deviation of density per OD-factor. The OD-factor is a scaling factor, which is applied to the complete OD-matrix. This factor is decreased if the network ends up in gridlock (>90% of vehicles has not changed link during the last 5 minutes), and is increased if the network did not end up in gridlock (see step 4 of paragraph 5.4.1).

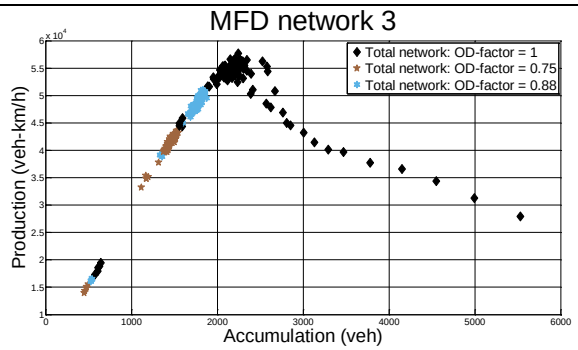
Table 7.8: *Overview of MFDs*



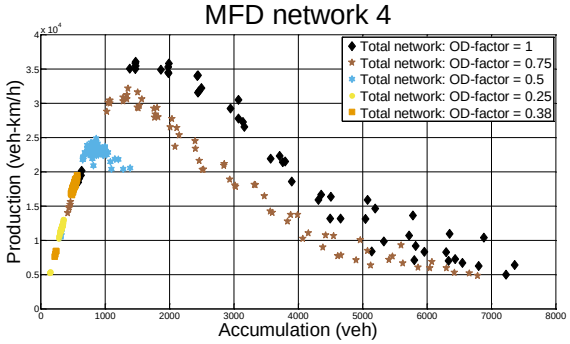
OD scaling factor	Max. prod.	Optimal accumu-lation	Dev. of density
1,00	57.167	2.214	12,29
0,75	44.088	1.541	8,43
0,88	52.115	1.886	10,27



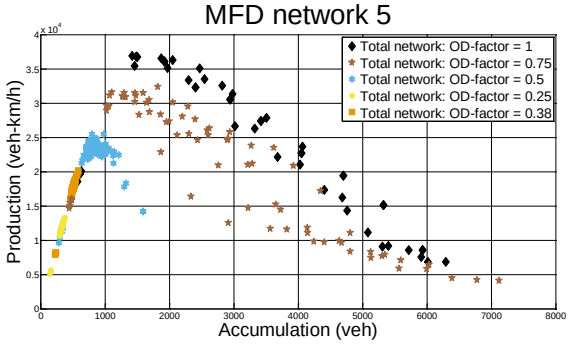
OD scaling factor	Max. prod.	Optimal accumu-lation	Dev. of density
1,00	58.038	2.377	14,82
0,75	45.234	1.594	9,50
0,88	52.827	1.954	11,00



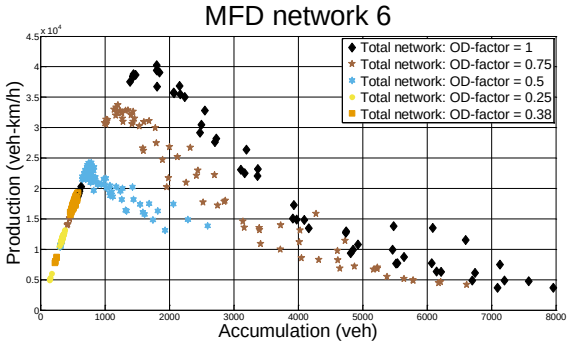
OD scaling factor	Max. prod.	Optimal accumu-lation	Dev. of density
1,00	57.661	2.242	12,21
0,75	43.681	1.544	9,33
0,88	51.079	1.832	10,43



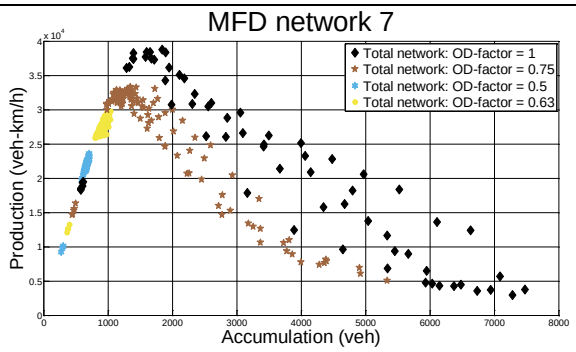
OD scaling factor	Max. prod.	Optimal accumu-lation	Dev. of density
1,00	36.100	1.477	14,00
0,75	32.172	1.358	13,42
0,50	24.877	861	8,15
0,25	13.006	358	1,89
0,38	19.510	566	3,31



OD scaling factor	Max. prod.	Optimal accumu-lation	Dev. of density
1,00	36.931	1.423	11,78
0,75	32.456	1.819	18,15
0,50	25.522	800	5,68
0,25	13.334	374	1,98
0,38	20.178	585	3,38



OD scaling factor	Max. prod.	Optimal accumu-lation	Dev. of density
1,00	40.233	1.805	15,46
0,75	33.701	1.200	9,23
0,50	24.179	786	5,97
0,25	13.156	379	2,01
0,38	19.319	583	3,13



OD scaling factor	Max. prod.	Optimal accumu-lation	Dev. of density
1,00	38.813	1.839	17,58
0,75	33.378	1.345	11,63
0,50	23.661	709	4,70
0,63	29.812	1.039	8,92

From the MFDs and the data presented above, it can be found that the production is indeed higher when the OD-factor and thus the number of vehicles loaded onto the network is higher, which should be expected, as long as the network does not end in congestion.

In the first three networks, that had little congestion throughout the different simulations, the production was found to be very constant over time, and no peak was found at the start of the simulation. Nevertheless, the production was found to drop slightly as time progressed in network 1, for an OD-factor of 1,00, ending up at the same production as for the OD-factor 0,88.

All other networks showed either a constant production over time, or a very high production at the start of the simulation, after which the production dropped severely, to end up in a state of gridlock, ending the simulation. These graphs are shown in Figure 7.7 to Figure 7.13. The figures illustrate, that for network 4 to 7, a high production is achieved at the start of the simulation, at the higher OD-factors, which cannot be sustained, resulting in gridlock.

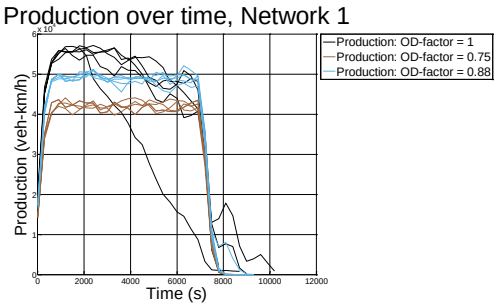


Figure 7.7: *Production over time, network 1*

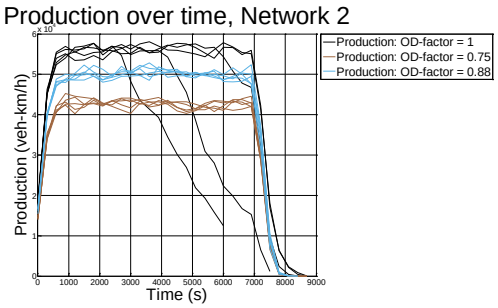


Figure 7.8: *Production over time, network 2*

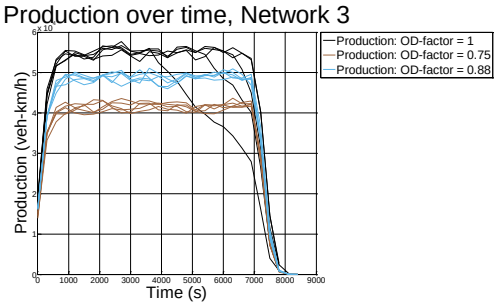


Figure 7.9: *Production over time, network 3*

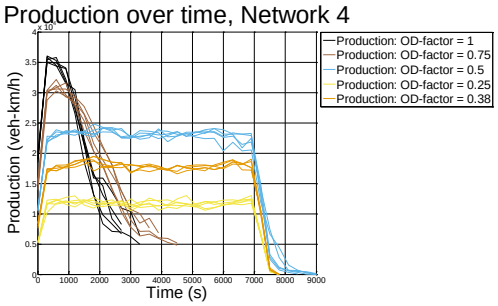


Figure 7.10: *Production over time, network 4*

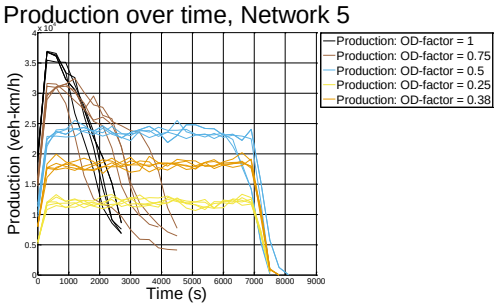


Figure 7.11: *Production over time, network 5*

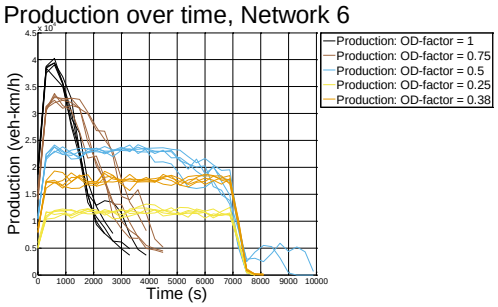


Figure 7.12: *Production over time, network 6*

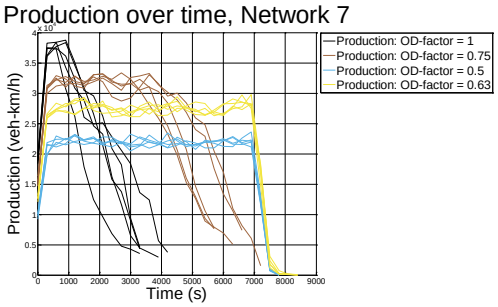


Figure 7.13: *Production over time, network 7*

From this, it is concluded that an effect, referred to as 'the rush' seems to exist. However, the assumption that this is directly related to the number of vehicles loaded onto the network is not apparent, as the relation between OD-factor and maximum production and optimal accumulation is not found to be constant as is shown in Table 7.9. What the table does show is that when the network is uncongested, that the OD/production-ratio is fairly constant, and that when the network becomes congested, the OD/production-ratio increases, meaning that the increase of the production is less than the increase in number of vehicles.
Nevertheless, the initial production is still higher than at the lower factors. Also, network 4 to 7 show that the production is still higher at an OD-factor of 1,00 than at 0,75, even though the network already becomes congested at an OD-factor of 0,75. This means that the maximum production is not capped by the point at which congestion sets in, implying that these states are indeed outside the boundaries of the steady-state MFD.
Table 7.8 also shows that the deviation of density at these higher OD-factors is larger, which means that the higher production achieved cannot be attributed to a lower deviation of density, making the conjecture that a phenomenon like 'the rush' exists, all the more plausible.

Table 7.9: *OD/production-ratio and OD/accumulation-ratio*

Network	OD-factor	Maximum production (veh-km/h)	Optimal accumulation (veh)	OD/ production-ratio ($x10^{-6}$)	OD/ accumulation-ratio ($x10^{-6}$)
1	1,00	57.167	2.214	17,5	451,7
	0,75	44.088	1.541	17,0	486,6
	0,88	52.115	1.886	16,9	466,6
2	1,00	58.038	2.377	17,2	420,7
	0,75	45.234	1.594	16,6	470,4
	0,88	52.827	1.954	16,7	450,4
3	1,00	57.661	2.242	17,3	446,0
	0,75	43.681	1.544	17,2	485,8
	0,88	51.079	1.832	17,2	480,3
4	1,00	36.100	1.477	27,7	677,1
	0,75	32.172	1.358	23,3	552,2
	0,50	24.877	861	20,1	580,8
	0,25	13.006	358	19,2	697,6
	0,38	19.510	566	19,5	671,8
5	1,00	36.931	1.423	27,1	702,9
	0,75	32.456	1.819	23,1	412,4
	0,50	25.522	800	19,6	625,0
	0,25	13.334	374	18,7	668,1
	0,38	20.178	585	18,8	649,3
6	1,00	40.233	1.805	24,9	554,1
	0,75	33.701	1.200	22,3	625,2
	0,50	24.179	786	20,7	636,4
	0,25	13.156	379	19,0	659,2
	0,38	19.319	583	19,7	651,8
7	1,00	38.813	1.839	25,8	543,7
	0,75	33.378	1.345	22,5	557,5
	0,50	23.661	709	21,1	705,2
	0,63	29.812	1.039	21,1	606,5

7.5.3 Conclusion

It has been found that the maximum production that can be achieved in a network is related to the number of vehicles loaded onto the network, which is logical, as the production is related to the average flow, which in turn is related to the number of vehicles in the network. So, for uncongested conditions a linear relationship is found. However, when the number of vehicles loaded onto the network is too high, the network becomes congested and the increase in the production is lower than the increase in the number of vehicles, meaning that the maximum production is not fully linear to the number of vehicles loaded onto the network, as was hypothesised.
Nevertheless, it has been found that the maximum production can still increase, even when the number of vehicles loaded onto the network is higher than the number that can be steadily sustained, as the network would already get into a state of gridlock when less vehicles are loaded.

These results therefore suggest that a phenomenon as 'the rush' clearly exists. What this finding entails, is that the quality of a network cannot be fully assessed from the maximum production or performance found in the MFD. Apart from that, basing a control strategy on the optimal accumulation found, might also prove to be ineffective, as these accumulations might inevitably lead to congestion.

7.6 Stochasticity in network design

7.6.1 Hypothesis

In order to investigate to what degree the MFD is affected by small changes to the network and stochastic variations in road patterns and signal timings, three similar networks have been created. In the network creation model, the street pattern of the subnetworks and the location of origins and destinations is created randomly. Also the number of trips between every OD-pair is randomly assigned. Due to these differences, the route pattern changes and the number of lanes assigned to the various links, caused by different demands, in the arterials differs for every network. Nevertheless, it has been hypothesised that the effects on the shape of the MFD might only be minimal, as the shape of the MFD should be mostly related to the capacity of the individual links. Although the production and accumulation could be different, because of difference in the network length, these are factored out in the PF-K MFD. From this it was hypothesised that the PF-K MFD for the three networks should be almost equal.

7.6.2 Analysis of the results

In order to compare the differences between the three networks, the different values of the MFD for the total network have been compared to each other graphically in Figure 7.14 to Figure 7.16 and numerically in Table 7.10. Judging from the MFDs, the three networks are almost equal and little difference in the shape is found, apart from the congested branch of the arterial MFD.

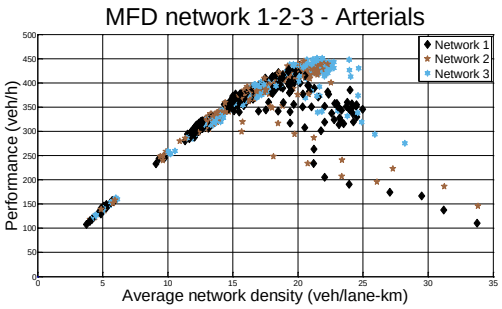


Figure 7.14: *Comparison of arterial MFDs network 1-2-3*

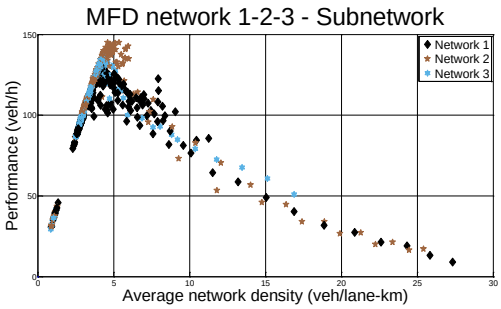


Figure 7.15: *Comparison of subnetwork MFDs network 1-2-3*

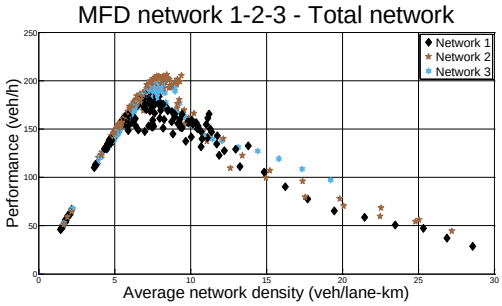


Figure 7.16: *Comparison of the total network MFDs network 1-2-3*

Looking at the comparison between the parameters of the different networks, the differences seem to be more substantial. Interestingly, the difference in the maximum production and the accumulation are quite similar, while the differences between the performance and the network density are higher.

Judging from the data of the arterials and subnetworks, the difference found can be mainly attributed to the subnetworks, in which the average network density in particular is substantially different. Looking at the data in Table 7.1, it can be found that the accumulation and the average network density of network 2 are far higher, causing these differences. However, no substantial differences can be found, explaining why network 2 performs this much better. The total outflow capacity of the signals is less than 1,7% higher, the number of trips originating from and/or destined for the subnetwork do not differ more than 1%, and the deviation of density (see Table 7.8) is not substantially different, meaning that none of these factors can explain this differences found.

Table 7.2 shows that the average travel time in network 2 is higher than in the other two networks. The average travel time of network 2 is 18% higher than in network 1, which could explain some of the differences, as it has been found that the accumulation is a function of the flow on a link and the average travel time on that link. The higher average travel time, can in turn be related to the average subnetwork speed, which is about 16% lower in network 2, than in the other networks. The cause of this however, still remains unclear and as such is attributed to differences in the layout of the subnetwork.

Table 7.10: *Cross-comparison of absolute differences due to stochastic network design*

Compared networks	Max. prod. (veh-km/h)	Opt. acc. (veh)	Perf. (veh/h)	Netw. dens. (veh/lane-km)	Length (km)
Total network					
1 - 2	1,50%	6,86%	9,75%	14,66%	8,04%
1 - 3	0,86%	1,27%	7,18%	7,57%	6,02%
2 - 3	0,65%	6,00%	2,85%	8,31%	1,87%
Arterials					
1 - 2	2,64%	2,27%	1,95%	1,59%	1,20%
1 - 3	3,55%	6,40%	3,44%	6,29%	0,59%
2 - 3	0,94%	4,22%	1,52%	4,78%	0,61%
Combined subnetworks					
1 - 2	0,68%	12,80%	9,93%	21,99%	6,50%
1 - 3	3,78%	6,49%	4,44%	1,95%	7,00%
2 - 3	3,08%	22,12%	6,09%	25,68%	0,47%

7.6.3 Conclusion

It was hypothesised that stochastic variations in the creation of networks derived from the same set of parameters, having the same general layout, should not significantly impact the shape of the MFDs of these networks. The MFDs derived from these networks confirm this hypothesis, as the MFDs have a strong overlap, as can be seen in Figure 7.14 to Figure 7.16.

However, when the different parameters are compared to each other, substantial differences between the accumulation and average network density of the subnetworks are found (>20%). Multiple factors have been investigated to explain this difference, such as network length, total green time of signals controlling outflow, average trip length and number of trips originating from, or destined for the subnetwork, but none of these factors, or even the combination of these factors can explain this difference.

This observation therefore leads to the conclusion that differences in the shape of the MFD can arise due to the stochasticity in network design and variability in traffic. These differences however, are not always clearly visible in the MFD itself.

7.7 Scalability of differently shaped subnetworks and perimeters

7.7.1 Hypothesis

In order to quantify the impact of different factors on the MFD, the first thing that should be done is to see whether or not, different network are actually scalable, or that the impact of the network dynamics is of such nature that this is not possible. In the previous section it was found that stochastic variations in the creation of networks can indeed lead to differences in the MFDs, that cannot be fully explained by any of the factors assumed to have an influence on the shape. To further investigate how stochasticity in network design might impact the MFD, the MFDs of different subnetworks and perimeters are investigated.

Under the conjecture that traffic conditions are roughly equal between different subnetworks and the ratio between network length and the total waiting time at signals is not substantially different, it is expected that the shape of the PF-K MFD should almost be the same.

7.7.2 Analysis of the results

Network 4

The first comparison that is made, is between the different MFDs of the subnetworks of network 4. Network 4 consists of two subnetworks, in which the first subnetwork is twice as large as the first one. The total length of the perimeter of the second network is 10% higher than that of the first subnetwork. If the shape of the MFD should rely on the size of a network, it should be expected that the resulting PF-K MFDs should have the same scale. The MFDs for the subnetwork, and the perimeter and the combination of subnetwork and perimeter are presented in Figure 7.17 to Figure 7.19. The data is presented in Table 7.11.

Results

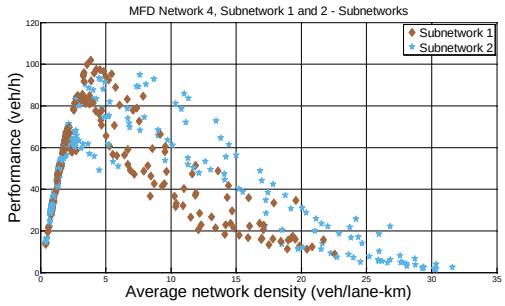


Figure 7.17: *Comparison of subnetwork MFDs network 4*

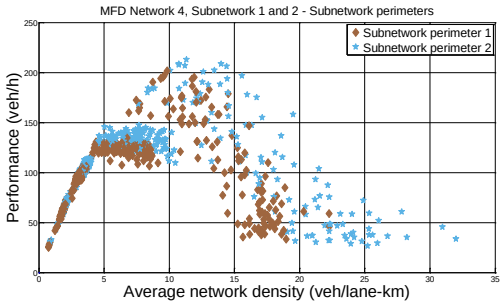


Figure 7.18: *Comparison of perimeter MFDs network 4*

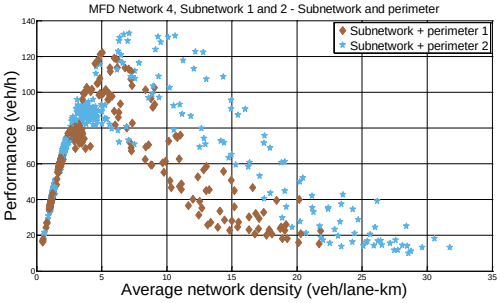


Figure 7.19: *Comparison of the combined subnetwork and perimeter MFD network 4*

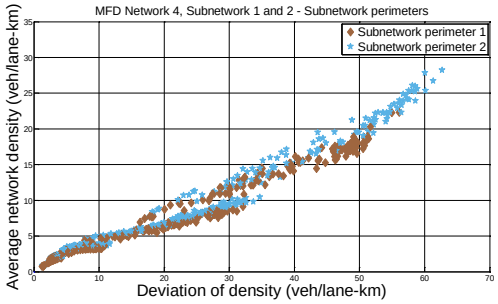


Figure 7.20: *Deviation of density of the perimeters of network 4*

Table 7.11: *Parameters for the subnetworks and perimeters of network 4*

OD-factor		1	0,75	0,5	0,25	0,38
Sub-network 1	Maximum production	16.531	14.895	11.307	5.962	8.733
	Optimal accumulation	624	538	357	166	249
	Maximum performance	102	92	70	37	54
	Optimal network density	3,8	3,3	2,2	1,0	1,5
Sub-network 2	Maximum production	7.118	6.795	5.352	2.888	4.118
	Optimal accumulation	571	255	161	79	121
	Maximum performance	95	91	71	38	55
	Optimal network density	7,6	3,4	2,1	1,0	1,6
Difference	Maximum production	2,32	2,19	2,11	2,06	2,12
	Optimal accumulation	1,09	2,11	2,21	2,11	2,06
	Maximum performance	1,07	1,01	0,98	0,95	0,98
	Optimal network density	0,50	0,97	1,02	0,98	0,95
Perimeter 1	Maximum production	9.253	8.053	6.214	3.374	4.877
	Optimal accumulation	452	502	306	99	167
	Maximum performance	202	176	135	74	106
	Optimal network density	9,8	10,9	6,7	2,2	3,6
Perimeter 2	Maximum production	8.287	7.363	5.734	3.138	4.526
	Optimal accumulation	441	438	299	95	160
	Maximum performance	213	189	147	81	116
	Optimal network density	11,3	11,2	7,7	2,4	4,1
Difference	Maximum production	1,12	1,09	1,08	1,08	1,08
	Optimal accumulation	1,02	1,15	1,02	1,04	1,05
	Maximum performance	0,95	0,93	0,92	0,91	0,91
	Optimal network density	0,87	0,97	0,87	0,88	0,89

General

The first thing that can be noticed is that the shape of the MFDs match quite nicely in the uncongested range, but differ in the congested range, in which the second subnetwork performs better. This is opposed to expectations, as this is the smaller subnetwork. Nevertheless, in

hindsight, this is actually quite logical, because the traffic can be distributed over more space in the second network, leading to lower average network density.
The shapes of the different MFDs do however match quite nicely and follow the same pattern.

Subnetworks

When comparing the two subnetworks to each other, it can be found that the production and accumulation of subnetwork 1 is roughly 2,1 times larger than subnetwork 1, which is equally very close to the difference of the area, but even closer to the difference in actual lane length 162,2/75,0 = 2,16. This is also shown by the performance and average network density, which are almost equal for the two subnetworks. The only substantial difference found is at an OD-factor of 1, in which both subnetworks perform almost equal. The reason for this is, that at this OD-factor, subnetwork 2 ends up in gridlock first, leading to higher densities. The fact that the accumulation ends up equal in both subnetworks is ascribed to coincidence.

Perimeters

When comparing the parameters of both perimeters, the maximum difference in the optimal accumulation and maximum production both are about 15%, which is close to the difference in lane length of the perimeters at 62,6/57,0 = 1,10.
Still it should be pointed out, that even though the data matches properly, the PF-K MFDs do not. A closer look at the deviation of density (Figure 7.20) does not reveal a significant difference to explain why the MFD of the perimeter of subnetwork 2 shows higher average densities. The only reasonable explanation for this is, that the perimeter of subnetwork 2 consists of a larger part out of the central arterial, which has the highest density of the different arterials. The effect of this higher density is therefore larger on the perimeter of subnetwork 2, than for the other perimeter.

Network 5

The next subnetworks that are compared are the subnetworks of network 5, which are both sized in the same way and only have minor differences in lane length of the subnetwork and perimeter. The MFDs for the subnetwork, and the perimeter and the combination of subnetwork and perimeter are presented in Figure 7.21 to Figure 7.23. The data is presented in Table 7.12.

Results

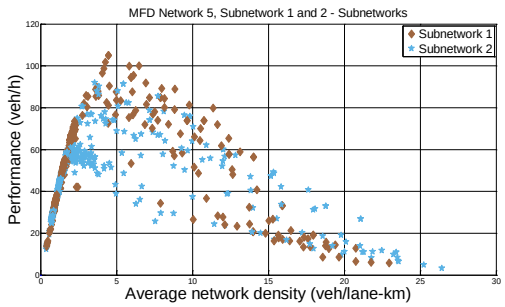


Figure 7.21: *Comparison of subnetwork MFDs network 5*

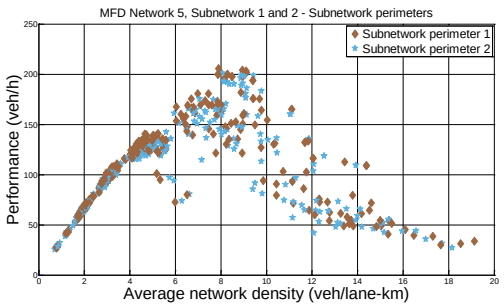


Figure 7.22: *Comparison of perimeter MFDs network 5*

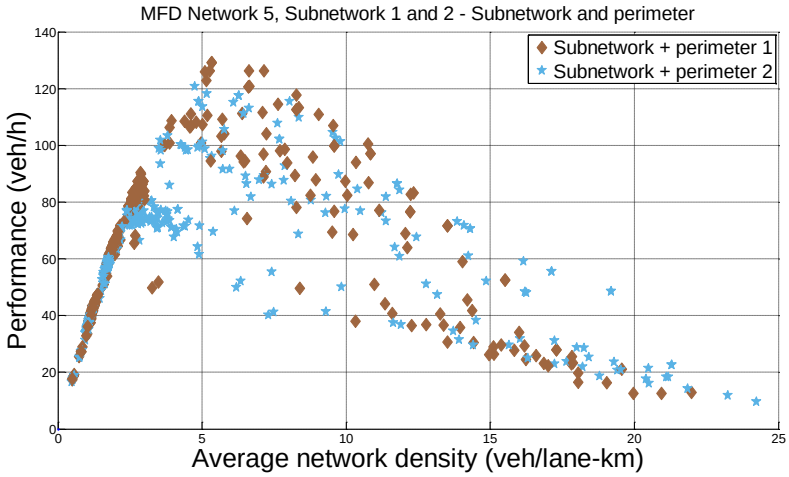


Figure 7.23: *Comparison of the combined subnetwork and perimeter MFD network 5*

Table 7.12: *Parameters for the subnetworks and perimeters of network 5*

OD-factor		1	0,75	0,5	0,25	0,38
Sub-network 1	Maximum production	13.260	11.598	9.275	5.005	7.400
	Optimal accumulation	564	875	282	140	218
	Maximum performance	105	92	73	40	59
	Optimal network density	4,5	6,9	2,2	1,1	1,7
Sub-network 2	Maximum production	11.029	9.677	7.880	3.957	5.945
	Optimal accumulation	425	611	391	111	172
	Maximum performance	92	81	66	33	50
	Optimal network density	3,5	5,1	3,3	0,9	1,4
Difference	Maximum production	1,20	1,20	1,18	1,26	1,24
	Optimal accumulation	1,33	1,43	0,72	1,26	1,27
	Maximum performance	1,14	1,14	1,12	1,20	1,18
	Optimal network density	1,26	1,36	0,69	1,20	1,20
Perimeter 1	Maximum production	8.844	7.777	6.048	3.247	4.991
	Optimal accumulation	340	321	225	96	162
	Maximum performance	206	181	141	75	116
	Optimal network density	7,9	7,5	5,2	2,2	3,8
Perimeter 2	Maximum production	8.633	7.529	6.118	3.209	4.767
	Optimal accumulation	347	303	240	98	154
	Maximum performance	201	175	142	75	111
	Optimal network density	8,1	7,1	5,6	2,3	3,6
Difference	Maximum production	1,02	1,03	0,99	1,01	1,05
	Optimal accumulation	0,98	1,06	0,94	0,98	1,05
	Maximum performance	1,02	1,03	0,99	1,01	1,05
	Optimal network density	0,98	1,06	0,94	0,98	1,05

General

When looking at the shape of the different MFDs, they are found to be almost equal for both subnetworks, as all the MFDs follow the same pattern and scatter is present over the same region. Simply judging from the MFDs, it can be concluded that both subnetworks and perimeters perform equal.

The major difference that can be found is that subnetwork 2 is a little bit easier to become congested. This can be seen at a performance of 60 veh/h, in which the first subnetwork remains in free flow for most of the time, but subnetwork 2 shows a little bit more points below the linear branch of the MFD.

Subnetworks

In order to make a good comparison between both subnetwork, the MFD parameters are compared in Table 7.12. From the table it can be found that the production of subnetwork 1 is 15 – 25% higher than subnetwork 2. Although the total lane length of subnetwork 1 is about 5% larger, it should not be the cause of the difference in performance. Also the differences in the physical layout of the subnetworks seems to be negligible.

Another explanation could be found in the routes that vehicles are taking and the way in which the network is loaded, but when looking at the traffic patterns, still no substantial difference can be found. However, when the OD-matrix is investigated, it is found that the number of vehicles originating from or destined for subnetwork 2 is 11 – 13% higher. Which seems to be a reasonable explanation for the difference, as all other factors seem to be equal. This also matches the performance, which is found to be strongly related to the number of vehicles loaded onto the network (as proven in section 7.5). The optimal accumulation and average network density are found to deviate strongly, resulting in a substantial dip at an OD-factor of 0,50. The reason for this dip is not completely clear, but it can be found that the accumulation and network density are highly volatile, as they increase from an OD-factor of 1,00 to 0,75 and then start decreasing. Interestingly, this shows that there seems to be an optimum in the optimal accumulation.

Perimeters

The resulting values illustrate that the perimeter for both subnetworks perform almost equal at any OD-factor, supporting the hypothesis that the MFDs of different subnetworks and perimeters should be scalable.

Network 6

Network 6 consists of 4 different subnetworks, with varying shapes and sizes. An overview of the different network parameters can be found in Table 6.1 and Table 6.2. The resulting MFDs (PF-K and PD-A) of the different subnetworks and perimeters are presented in Figure 7.24 to Figure 7.27. In order to make any statement on the relation between the different subnetworks, the accumulation and production of subnetwork 1, 3 and 4 are all scaled against subnetwork 2, which is the largest subnetwork in the network. This format differs from the previous results, but is done in order to keep the size of Table 7.13 within reasonable limits.

Results

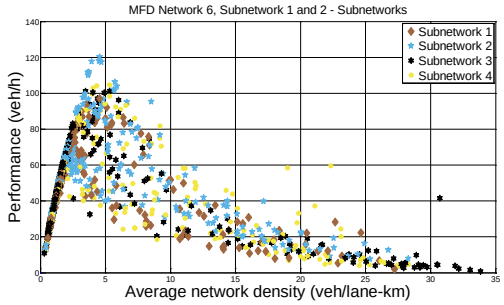


Figure 7.24: *Comparison of subnetwork MFDs network 6 (PF-K)*

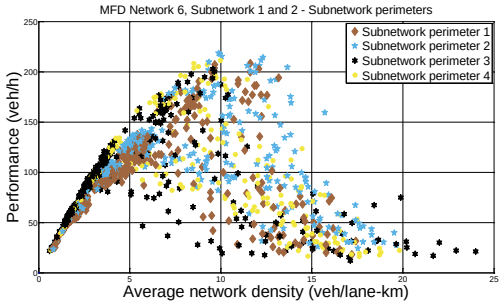


Figure 7.25: *Comparison of perimeter MFDs network 6 (PF-K)*

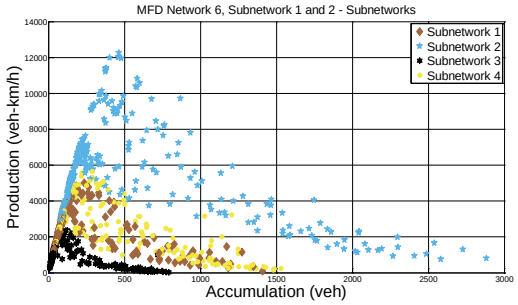


Figure 7.26: *Comparison of subnetwork MFDs network 6 (PD-A)*

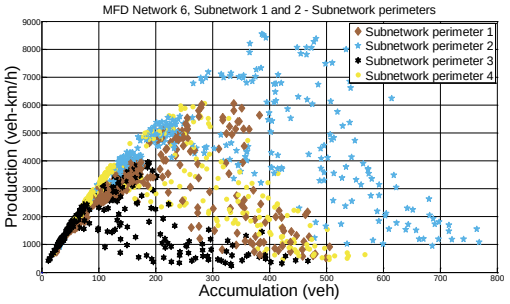


Figure 7.27: *Comparison of perimeter MFDs network 6 (PD-A)*

General

When looking at the different MFDs, it can be noticed that the shapes of the MFDs of various subnetworks and perimeters are roughly the same. The shapes of the different MFDs however, does vary quite strong. Whereas the MFD of the subnetworks has a very long tail (congested branch) and a short region in which it is in free flow, the MFD of the perimeters is much more compact and has a very short tail in relation to the free flow region, meaning that when congestion has set in, the perimeters quickly end up in gridlock. On the other hand, the average density at which the transition from the free flow state to the congested state is made, is roughly twice as high for the perimeters than for the subnetworks, which means that the perimeters are better in processing traffic. This is also shown by the fact that the performance of the arterials can also become twice as high as the subnetworks.

The shape of the resulting MFD for both the subnetwork and perimeter (see Appendix C), does resemble that of the subnetworks the most, which is logical, as the total length of the subnetwork is much higher than that of the perimeter, making its impact larger.

Table 7.13: *Parameters for the subnetworks and perimeters of network 6, compared to the subnetwork and perimeter of subnetwork 2*

OD-factor		1	0,75	0,5	0,25	0,38
Sub-network 1	Maximum production	0,41	0,42	0,42	0,44	0,43
	Optimal accumulation	0,51	0,40	0,39	0,43	0,42
	Maximum performance	0,81	0,83	0,83	0,86	0,84
	Optimal network density	1,00	0,78	0,76	0,84	0,84
Sub-network 3	Maximum production	0,19	0,20	0,20	0,21	0,20
	Optimal accumulation	0,27	0,19	0,19	0,20	0,19
	Maximum performance	0,84	0,89	0,89	0,91	0,87
	Optimal network density	1,17	0,85	0,82	0,88	0,84
Sub-network 4	Maximum production	0,46	0,48	0,48	0,47	0,48
	Optimal accumulation	0,62	0,52	0,46	0,47	0,47
	Maximum performance	0,87	0,91	0,91	0,89	0,91
	Optimal network density	1,17	0,98	0,87	0,88	0,89
Perimeter 1	Maximum production	0,71	0,73	0,71	0,70	0,72
	Optimal accumulation	0,87	0,66	0,65	0,75	0,75
	Maximum performance	0,96	0,98	0,96	0,95	0,97
	Optimal network density	1,17	0,88	0,87	1,01	1,01
Perimeter 3	Maximum production	0,46	0,47	0,46	0,44	0,44
	Optimal accumulation	0,48	0,33	0,33	0,41	0,35
	Maximum performance	0,93	0,94	0,92	0,89	0,89
	Optimal network density	0,97	0,67	0,67	0,82	0,70
Perimeter 4	Maximum production	0,71	0,70	0,69	0,67	0,70
	Optimal accumulation	0,74	0,60	0,52	0,60	0,58
	Maximum performance	0,97	0,96	0,94	0,91	0,95
	Optimal network density	1,01	0,82	0,71	0,82	0,79

Subnetworks

The MFD of the subnetworks shows that spill-back can start at very low densities. The main reason for this is that the intersections in the subnetwork are not controlled, leading to complex interactions between different vehicles and streams, which decreases the over-all performance of the subnetworks. Looking at the MFDs, it is also found that the maximum production and optimal accumulation of each of the subnetworks in relation to the perimeter differs quite strongly. Subnetwork 2 has a higher production than its perimeter, whereas subnetwork 2 and 4 are roughly equal to their perimeter, even though they have 35 – 50% more road length. Subnetwork 3 even has a lower production than its perimeter, but as they are both almost equal in length, this is expected, as traffic is processed less effectively.

These results imply that the way in which a control strategy, such as subnetwork perimeter control, should be highly dependent on the MFD of both the subnetwork and the perimeter and cannot be

generalised. In paragraph 3.1.4 it was assumed that a valid control strategy should rely on keeping the arterials/perimeter at optimal accumulation during free flow states and that when congestion sets in, vehicles should be stored in the subnetworks. Looking at the different MFDs, this does not seem to be completely accurate, as it is found that for the subnetwork and perimeter of subnetwork 3, the subnetwork should be kept at optimal accumulation, while the perimeter should be used to 'store' excess traffic. Nevertheless, as the perimeter of subnetwork 3 shares the same roads with subnetwork 1 and 4, storing vehicles in the perimeter might have an adverse effect on the MFDs of the other subnetworks and arterials.

Looking at the differences presented in Table 7.13, it can be found that most of the subnetworks are scalable based on their length, as the difference in performance and average network density is within 20% for all the subnetworks. Any remaining differences were hypothesised to be related to the conflict ratio, i.e. the ratio between the total (sub)network length and the total waiting time. The conflict ratio of subnetworks 1, 3 and 4 in relation to subnetwork 2 are 1.31, 1.45 and 1.26 respectively. However, compensating the performance and average network densities with these ratios, results in basically the same differences for subnetwork 1 and 4, and results in a difference of over 30% for subnetwork 3, as is illustrated in Table 7.14. As this does not seem to decrease the difference between the subnetworks, the hypothesis that the conflict ratio should influence the shape of the MFD is therefore rejected. Nonetheless, it should be noted that the difference in performance between network 1 and 4 (which are equally shaped), is now almost equal. This then implies that this ratio does have an influence on the shape of the MFD, but that this relationship is not linear.

Table 7.14: *Subnetwork performance difference compensated by conflict ratio*

Subnetwork	Conflict ratio	Compensated performance				
1	1,31	1,04	1,07	1,07	1,11	1,09
3	1,45	1,27	1,35	1,34	1,38	1,31
4	1,26	1,06	1,11	1,11	1,09	1,11

Perimeters

From the PF-K MFD for the perimeters was derived that the shape was generally the same. However, looking closer at the MFD, it can be found that the steepness of the angle of the free flow branch, i.e. the average network speed does differ for the various perimeters. The lowest speed is found on perimeter 1, followed by perimeter 2. Perimeter 3 and 4 perform almost equal. As a consequence of the lower speeds on the perimeters, the average density of these perimeters is higher than for the other perimeters, resulting in higher performances at the same average density in the congested region.

The worst performing perimeter is that of subnetwork 3, which is also the smallest perimeter. The reason for this, as found before in paragraph 7.4.2 and shown in Figure 7.5, that the central arterial attracts the most traffic and congestion is mainly centred around this intersection. Due to the relative closeness to this intersection, subnetwork 3 and its perimeter are affected by this the most, which is also reflected in the values for the average network density.

Looking at the differences in the performance of the perimeters in relation to the perimeter of subnetwork 2, than the maximum difference found is 11%. These results then support the hypothesis that PF-K MFDs should be roughly equal (under homogeneous traffic conditions).

The scalability of the optimal accumulation and average network density however seems to be less, as these differences are much larger and found to be even as much as 35%. The reason as to why this is less scalable is not completely clear. One of the explanations could be that the accumulation is related to the average travel times, which is assumed to increase exponentially as the flow increases (as illustrated with the BPR-function in equation (3.6)). As the exponent in this equation is very strong, small difference in the flow on individual links can have a high impact on the travel times and therefore have a strong impact on the accumulation. This also means that the accumulation is very sensitive to the distribution of traffic over the network.
In order to illustrate the above, imagine we would have a small network with 4 links and a capacity of 2.000 veh/h, each governed by equation (3.6), with $\beta_{BPR} = 4$. If all of the links have a flow of 1.600 veh/h, the ratio $\left(\frac{q_i}{C_i}\right)^{\beta_{BPR}}$ for the links would be 0,41 and the total increase in travel time (and thus accumulation) of 1,64 at a performance of 1.600 veh/h. Now if three of the links would have a flow of 1500 veh/h and one would be 1900 veh/h, the performance would still be 1.600 veh/h. The travel time however, has increased with a factor of 1,76.

Network 7

Results

The last network that has been simulated contains 4 subnetworks, which are all equally sized, with an area of 2,25 km^2 and a perimeter with a length of 6,0 km. Due to the fact that the subnetworks and perimeters are all equally shaped, it is expected that the shape of the resulting MFDs should be almost the same. The MFDs can be found in Figure 7.28 and Figure 7.29. As with the previous network, the different subnetworks are compared to one another, in which subnetwork 1 is used as the base subnetwork. The results of this comparison are presented in Table 7.15.

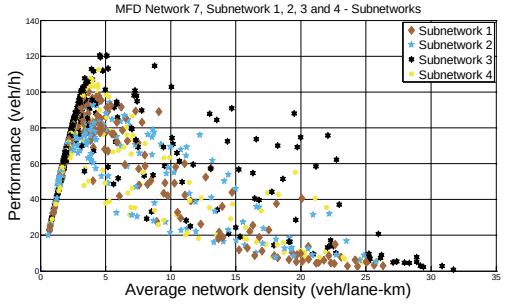


Figure 7.28: *Comparison of subnetwork MFDs network 7 (PF-K)*

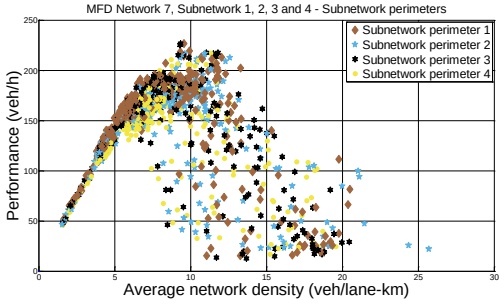


Figure 7.29: *Comparison of perimeter MFDs network 7 (PF-K)*

Table 7.15: *Parameters for the subnetworks and perimeters of network 7, compared to the subnetwork and perimeter of subnetwork 1*

OD-factor		1	0,75	0,5	0,63
Sub-network 1	Maximum production	1,07	1,00	1,00	1,00
	Optimal accumulation	1,07	1,23	1,00	1,07
	Maximum performance	1,00	0,93	0,94	0,94
	Optimal network density	0,99	1,15	0,93	1,00
Sub-network 3	Maximum production	1,04	1,07	1,02	0,96
	Optimal accumulation	1,07	2,41	1,05	1,04
	Maximum performance	1,20	1,23	1,18	1,11
	Optimal network density	1,23	2,77	1,21	1,20
Sub-network 4	Maximum production	1,12	0,99	1,04	0,98
	Optimal accumulation	1,20	1,05	1,05	1,05
	Maximum performance	1,13	0,99	1,05	0,99
	Optimal network density	1,21	1,06	1,06	1,05
Perimeter 1	Maximum production	0,96	0,98	0,94	0,94
	Optimal accumulation	1,01	1,04	0,88	1,12
	Maximum performance	0,96	0,98	0,93	0,94
	Optimal network density	1,01	1,04	0,88	1,12
Perimeter 3	Maximum production	0,99	0,98	0,94	0,96
	Optimal accumulation	0,97	0,77	0,97	1,14
	Maximum performance	1,00	0,98	0,95	0,97
	Optimal network density	0,98	0,77	0,98	1,14
Perimeter 4	Maximum production	0,95	0,93	0,90	0,91
	Optimal accumulation	1,18	0,74	0,98	1,30
	Maximum performance	0,96	0,93	0,90	0,91
	Optimal network density	1,19	0,74	0,98	1,31

General

Although the different MFDs show some different scatter, the general shape is almost the same for all 4 subnetworks and their respective perimeters, as should be expected based on the previous experiences. The only exception to this is subnetwork 3, which performs much better than the networks in the congested region.

Subnetworks

From Table 7.15 it can be found that in general the subnetworks are equal to each other. However, in some cases substantial differences in the optimal accumulation are found. However, these differences are mostly caused because the absolute number of vehicles are very low and small differences in these number can have a substantial impact on the relative difference.

Nevertheless, the value found for the optimal subnetwork accumulation of subnetwork 3, is quite substantial at 200% and cannot be overlooked, as this network seems to perform much better than the others.
When looking at the MFD of this subnetwork, it can be seen that this value does still fit within reasonable boundaries of the MFD and cannot be regarded as a faulty data point. Although a lower deviation of density should be expected to be the reason behind this, this is actually not the case, as the deviation is of subnetwork 3 is the highest of all of the subnetworks. One of the other reasons is that the total amount of lane length is lower in this subnetwork than in the other subnetworks, this is actually the case, but this only amounts to 15% of the difference. Also the amount of traffic originating from or destined for this subnetwork and the total amount of outflow at the boundaries are roughly the same. A re-examination of the simulation itself did not show specific differences that can reasonably explain these differences. As such it seems that this phenomenon cannot be reasonably explained and should have something to do with the specific layout of this subnetwork and the way traffic is processed, which is somehow different.

Perimeters

For the perimeters it can be found that the MFDs look a lot like each other and that the differences in production and performance are no more than 10% for all perimeters. The differences in the optimal accumulation and average network density however, are somewhat larger. It is assumed that these differences are caused by the sensitivity of the accumulation and network density to the distribution of traffic over the network, as has been shown for network 6. This should then result in a lower deviation of density of subnetwork 3 and 4. Looking at Figure 7.30, it does seem that this is actually not the case, as the deviation of density seems to be roughly equal for all perimeters. However, some substantial differences around a network density of 10 veh/lane-km are found. Especially the deviation of density for perimeter 3 and 4 can become quite high, up to 34, respectively 40 veh/lane-km, whereas the deviation of perimeter 1 is at 28 veh/lane-km. Only one point for perimeter 1 is found which has a high deviation and belongs to an OD-factor of 0,63. It is also at this deviation that the relative optimal density of perimeter 4 is found to be much higher than for perimeter 1.

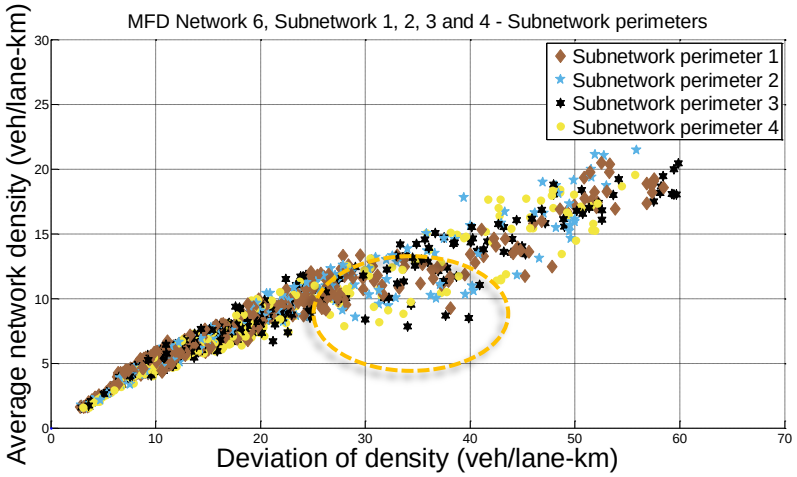


Figure 7.30: *Deviation of density of subnetwork perimeters of network 7*

7.7.3 Cross-comparison and scalability of results

When comparing the MFDs from the different subnetworks and perimeters to each other, it is found that the shape of the MFDs are very much alike.

Subnetworks

The shape of the subnetwork MFDs is for all subnetworks roughly the same. The congested branch is very steep, while the tail of the MFD (congested branch) is very long and somewhat curved inward, seeming to run asymptotically to the performance. Also the transition from the congested to the uncongested region is very sharp at high performances, and a little bit more fluent at lower performances. The maximum performance for most subnetworks is found between 100 and 120 veh/h, which is almost always at an average network density of approximately 5 veh/lane-km, regardless of the size of the subnetwork. The maximum average network density has been found to be around 25 – 30 veh/lane-km, which is interestingly close to the critical link density. The reason for this could very well be that when the network is at critical density, any small disturbance would inevitably lead to gridlock, as long as the number of vehicles in the network is kept equal.

The scatter in the subnetwork MFD often starts at half the maximum performance. Scatter is mostly found within a concave wedged-shaped region, bounded by the maximum performance, half the maximum performance and the critical density, as roughly sketched out in Figure 7.31.

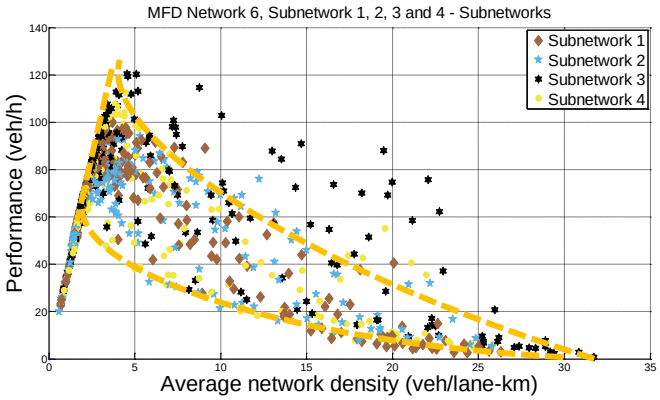


Figure 7.31: *Concave wedge-shaped area in which most scatter is found*

Perimeters

The MFDs for the different perimeters also show a strong resemblance to each other. These MFDs are differently shaped than the subnetwork MFDs, as they are generally a little bit more rounded and can achieve higher performances. Especially the congested area is highly scattered, often creating a 'shower'-like shape, as shown in the perimeter MFD of network 7 (Figure 7.29). The maximum performance is often found at a performance of 200 veh/h, at an average network density of the 10 veh/lane-km. The maximum average density achieved is approximately 20 veh/lane-km, but occasionally reaches 25 – 30 veh/lane-km if the deviation of density can be kept low enough.

These findings however do not seem to be universal, as the perimeters (arterials) of network 1, 2 and 3 can reach a performance 450 veh/h and average network density of 22.5 veh/lane-km, which is double that of the other perimeters. It should be pointed out that the shape of the MFD is strongly influenced by the free flow speed of the various links. The speed limit on the links of the different networks has been kept equal, which probably has led to the strong resemblance between the different MFDs. Introducing a higher speed limit, should cause the slope of the free flow branch of the MFD to become steeper, presumably resulting in a higher performance at a lower average network density.
Nevertheless, it seems that consistent perimeter MFDs can be created, if the networks from which they are derived resemble one another.

7.7.4 Conclusion

It has been hypothesised that the shape of subnetwork and perimeter MFDs should almost be the same in case the traffic conditions in the underlying network are homogeneous enough. Comparing the MFDs of 4 different networks, it has been found that the shape of different subnetworks and perimeters indeed strongly resemble each other, supporting the hypothesis.
Nevertheless, in some cases substantial differences were found, which could sometimes be attributed to signal settings, or differences in the number of trips originating or destined for a subnetwork, but were generally hard to explain.

It has also been found that the performance is much more scalable than the accumulation and the network density, and differences of up to 50% are found. These differences are attributed to the high sensitivity that the accumulation and average network density have to differences in the distribution of traffic. The reasoning behind this being that the accumulation depends on travel times, which increase exponentially at higher flows. Therefore an uneven distribution of traffic should lead to higher average travel times and thus higher accumulations. An example of this is shown on page 130.

It was further hypothesised that differences in the shape of the MFD of subnetworks, other than the total length could be attributed to the difference in the conflict ratio, i.e. the number of controlling intersections/outflow green time in relation to the network length. After adapting the performance for two similar networks by this ratio, the differences in performance were almost cancelled out. When this was applied to a much smaller network, the difference in performance only increased, thus disproving the hypothesis.

Also in some cases, substantial differences in performance or network density are observed, that cannot be related to any difference between the subnetworks or perimeters. In these cases the number of trips, average trip length, amount of green time of controlling signals, deviation of density, number of conflicts, et cetera, have been compared, but no substantial differences could be found explaining these differences. For these cases it is assumed that the differences are simply caused by the stochastic nature of traffic.

Furthermore it has been found that the perimeter might sometimes have a higher performance and a higher accumulation than the subnetwork it surrounds. The implication of these results is that a control strategy should not always aim at keeping the arterial/perimeter at optimal accumulation during free flow, while storing traffic in the subnetworks, which was hypothesised earlier. In this case, controlling the inflow to the subnetwork should be the best control strategy.

7.8 Conclusion

7.8.1 General

A number of aspects are tested in order to find out in what way the shape of the MFD was affected by changes in the network structure and which factors influence this shape. In the subsequent paragraphs, a summary of the various conclusions is made. For a more extensive conclusion the reader is referred to the conclusion of that specific section.

7.8.2 Adding additional arterials

Adding additional arterials to the network has not resulted in positive effects in the MFD, as almost all the parameters end up lower than in the base networks. Only the accumulation of the subnetworks seems to have been unaffected.

The arterials are affected the most, and in one case, a reduction in the performance of 60% is observed. Also the total number of vehicles that these networks can process is half that of the base network. The main explanation is that adding additional arterials in the centre of the network draws more traffic to the central arterials than to the outer arterials, causing a more uneven distribution of traffic over the network. The networks with two additional arterials however seem to be less affected than when only one arterial is added.

The average travel times in the networks with the additional arterials is lower, leading to a better result from the individuals perspective. This property however is not reflected in the MFD itself.

7.8.3 Asymmetric arterial placement

The asymmetric placement of arterials, does not seem to have a significant impact on the shape of the MFD. From this the conclusion is drawn that the exact physical layout of the arterials in a network does not seem to affect the MFD. This is in line with the hypothesis that the MFD is mainly derived from the characteristics of the individual underlying links.

However, the average travel time for network 6 and 7 differs approximately 25%, implying better design of network 7. The reason that network 7 performs better is that due to the more centrally placed intersection between the arterials, it takes longer for traffic to spill-back on the outer arterial, keeping it longer in free flow than for network 6. No significant differences in the average travel time are found for network 4 and 5. However, for these two networks, a substantial difference in the accumulation is found. Although this difference can be explained by the difference in the deviation of density between these two networks, it is not clear why the a-symmetrical network performs better than the symmetrical one.

7.8.4 Different traffic loading patterns

It has been found that the maximum production that can be achieved in a network is related to the number of vehicles loaded onto the network, but is only linear as long as the network does not

become congested. The production however can still increase if the number of vehicles in the network is higher than this optimum. The result of this is that production and accumulations are found that cannot be steadily sustained on the network.

These results support the hypothesis, that a phenomenon as 'the rush' clearly exists. The consequence of this effect, is that the quality of a network cannot be fully assessed from the maximum production or performance found in the MFD. Apart from that, basing a control strategy on the optimal accumulation found, might also prove to be ineffective, as these accumulations might inevitably lead to congestion.

7.8.5 Stochasticity in network design

Stochastic variations in the design of similar networks, can cause differences in the shape of the MFD. Substantial differences were found in the performance, accumulation and average network density. These differences could not be related to any specific factor and are assumed to be related to differences in the layout of the network and the stochastic nature of traffic.

7.8.6 Scalability of differently shaped subnetworks and perimeters

The shape of MFDs of differently shaped subnetworks and perimeters strongly resemble each other, when traffic conditions are homogeneous in the networks compared. Differences found, could in most cases be related to network length, signal settings, or the number of trips made in the subnetwork or perimeter.
Some substantial differences are found that cannot be explained by any factor and are attributed to the stochastic nature of traffic, or are assumed to be caused by 'the rush'.
It was also found that the performance is much more scalable than the accumulation, which can vary heavily. These variations of the accumulation and average network density are attributed to the uneven distribution of traffic over the network.

8 Effect of signal settings on the macroscopic fundamental diagram

In order to answer the question how the MFD of the subnetwork and perimeter is affected when changing the signal settings, the question has been broken down in two parts, which have both been simulated separately. After discussing the way in which the simulation results are presented in section 8.1, section 8.2 investigates how the shape of the MFD changes, when the timing of the signals regulating traffic from the perimeter into the subnetwork is changed. In section 8.3 it is investigated what happens to the MFD if the timing of signals regulating traffic from the subnetwork into the perimeter is changed. In both these sections, these effects are investigated per group of networks, i.e. the networks that are similar to one another.
After the effects of both controlling the inflow and the outflow have been obtained, section 8.4 will discuss which of these control strategies performs best. In section 8.5 the effect of different signal timings on the optimal accumulation will be investigated, in order to answer the question if the accumulation can be used as an input for control strategies such as perimeter control. Section 8.6 will close this chapter with a conclusion.

8.1 Simulation results

8.1.1 Presentation of simulation results

In order to investigate how the shape of the MFD of a subnetwork and its perimeter is affected by signal changes, multiple simulations have been run for every subnetwork and perimeter. As the number of MFDs resulting from these simulations is quite substantial, only the PF-K MFDs for each subnetwork and perimeter are presented in this report, in Appendix B. Both the MFDs for controlled subnetwork inflow and controlled subnetwork outflow are added. Other MFDs are added digitally as Appendix C.

8.1.2 Performance indicators

General

As has been found before, a MFD cannot convey information regarding average travel times and delays. However, in order to also assess the impact and the effectiveness of the different strategies, these parameters are actually very important. To this end, assessment tables have been created for each subnetwork and perimeter. Apart from showing the production, performance, accumulation and network density, these tables also include the average travel time, free flow travel, average delays and the number of vehicles that (1) had to be loaded onto the network, (2) have actually been loaded onto the network and (3) completed their trip. An example of such a table is shown in Table 8.1 which presents the maximum state and the optimum state. Due to their size, these tables are not added in this report, but can be accessed digitally in Appendix C.

Table 8.1: *Example of an assessment table*

Network 1 – Subnetwork 1 (outflow)											
State		Maximum					Optimum				
Green	Green (veh/h/signal) / control scenario	100	200	300	400	500	100	200	300	400	500
	Total green (veh/h)	4.400	8.800	13.200	17.600	22.000	4.400	8.800	13.200	17.600	22.000
	Green used (veh/h)	4.772	4.772	5.950	7.020	8.223	5.052	5.052	5.401	7.338	7.515
	OD - factor (-)	1,00	1,00	1,50	1,50	1,50	0,88	0,88	1,00	1,25	1,38
Sub-network	Production (veh-km/h)	27.265	27.265	38.917	40.413	40.271	25.851	25.851	27.160	34.800	38.330
	Accumulation (veh)	893	893	1.427	1.322	1.321	766	766	874	1.082	1.234
	Performance (veh/h)	110	110	157	163	162	104	104	109	140	154
	Network density (veh/km)	3,60	3,60	5,75	5,33	5,32	3,09	3,09	3,52	4,36	4,97
Peri-meter	Production (veh-km/h)	13.697	13.697	18.600	19.959	20.098	12.893	12.893	13.816	17.149	18.853
	Accumulation (veh)	457	457	795	810	775	400	400	422	537	686
	Performance (veh/h)	267	267	362	388	391	251	251	269	334	367
	Network density (veh/km)	8,90	8,90	15,48	15,76	15,07	7,79	7,79	8,22	10,45	13,35
Total	Production (veh-km/h)	40.446	40.446	57.422	59.713	60.029	38.681	38.681	40.651	51.731	57.086
	Accumulation (veh)	1.254	1.254	2.198	2.072	2.047	1.145	1.145	1.222	1.652	1.936
	Performance (veh/h)	135	135	192	199	200	129	129	136	173	191
	Network density (veh/km)	4,19	4,19	7,34	6,92	6,83	3,82	3,82	4,08	5,52	6,46
Vehicles	Average travel time (s)	381,9	381,9	474,5	465,9	454,9	370,4	370,4	383,7	384,6	438,6
	Free flow travel time (s)	340,0	340,0	350,0	350,0	350,0	340,0	340,0	340,0	340,0	350,0
	Average delay (s)	41,9	41,9	124,5	115,9	104,9	30,4	30,4	43,7	44,6	88,6
	Vehicles to be loaded (veh)	24.404	24.404	36.311	36.435	36.954	22.896	22.896	24.410	31.252	34.911
	Vehicles loaded (veh)	23.331	23.331	21.214	28.181	36.651	22.858	22.858	24.410	31.252	34.911
	Vehicles finished (veh)	17.566	17.566	13.613	21.186	34.702	21.393	21.393	24.410	31.252	34.911

Maximum state

The maximum state is the simulation, in which the highest production/performance is found. This is calculated separately for the subnetwork, perimeter and the combination of subnetwork and perimeter.

As it is found that in most cases the maximum production/performance for both the subnetwork and perimeter is achieved at the same OD-factor (which is also almost always the maximum OD-factor), these results have not been separated. Only for subnetwork 7.4, the maximum production was achieved at different OD-factors. In this case, the OD-factors are presented for each part of the network separately.

Optimum state

It has been found before that the 'the rush' can have quite an impact on the shape of the MFD and results in points that are found outside the boundaries of the steady-state MFD, i.e. the MFD containing states that can all be sustained over a longer period of time. These high productions and accumulations however, cannot be maintained and as a result, the simulation ended almost always in gridlock. From Table 8.1 it can be found that in the simulations that have the maximum performances, not all of the vehicles could be fully processed, and were stuck in the network due to gridlock.

In order to assess the effectiveness of each of the control scenarios properly, an optimal state is calculated as well, which is the OD-factor at which the most vehicles could actually finish their trip.

8.2 The effect of controlling subnetwork inflow

8.2.1 Network 1, 2 and 3

The shape of the MFD

The first networks that will be investigated are the three base networks. All of these networks consist of a single subnetwork and perimeter. The MFDs of the subnetworks, perimeters and the subnetwork and perimeter combined can be found in Appendix B. Below the MFDs of the combined subnetwork and perimeter are given.

From the MFDs it can be found that the performance of the subnetwork and perimeter does depend on the timing of the signals between them. Also a strong relationship between the subnetwork and its perimeter seems to exist, as both react in the same way to the signal changes, i.e. if their production would be ranked according to the controlled flow, the subnetwork and perimeter would be ranked in the same way.
The shape of each of the MFDs is different, as has been noticed before. The MFD of the subnetwork often has a short free flow branch and a long tail, with a sharp transition from free flow to congestion. The MFD of the perimeter is more rounded and can achieve a higher performance and average network density than the subnetwork.

For the subnetworks it is found that the maximum performance is in the range of 150 – 180 veh/h, with production between 26.000 and 37.000 veh-km/h. The optimal network density is approximately 5 veh/lane-km. In the optimal state however, the maximum performance is between 100 – 125 veh/h, and a production of 22.000 – 31.000 veh-km/h. The average network density is often lower and varies between 3 – 4 veh/h.

Again, the perimeter outperforms the subnetworks substantially, with maximum performances ranging from 250 to 400 veh/h, and optimal performances of 250 – 280 veh/h. The optimal accumulation of the perimeters on the other hand, is roughly half that of the subnetworks, ranging between 13.000 and 20.000 veh-km/h.

Performance ratio

This relation between the performance of the subnetwork and the perimeter has been tested and it is found that the ratio between the performance of the subnetwork and the perimeter is highly consistent throughout the different control scenarios. This ratio, which we will refer to as the 'performance ratio' is shown in Figure 8.1.

From the figure it can be derived that the ratio is almost unchanged at a control rate of 200 – 400 veh/h/signal. Only at a control rate of 100 veh/h/signal, the relative performance of the perimeter is somewhat higher, indicating that it does benefit from the restrictions imposed on the signals. Interestingly, this should be expected to work the other way around, as the restriction is imposed on the signals of the perimeter, giving more green time to the subnetwork, which should improve

its performance. What this indicates is that both the subnetwork and the perimeter suffer from a bad timing of signals. From this it can be concluded that simply restricting the inflow to a subnetwork, does not always improve the performance of the network.

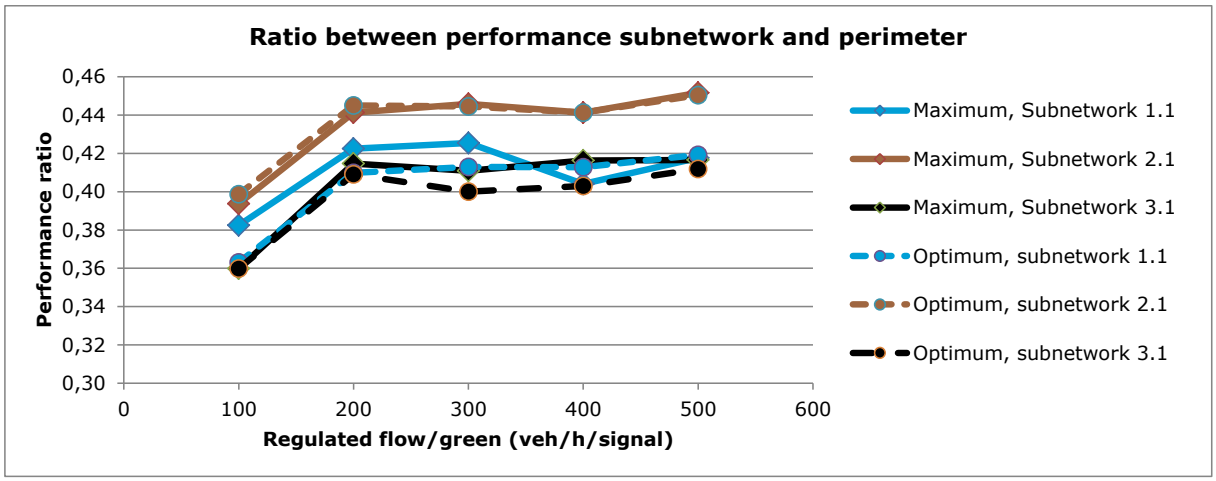


Figure 8.1: *Performance ratio subnetwork/perimeter, network 1, 2 and 3*

Imposing a restriction on the signals of 100 veh/h/signal, has an interesting side effect. As can be seen in Figure 8.2, the performance of the perimeter can be kept constant over a very large density range. The same has also been found for the other two networks. This shows that by adapting the traffic signals, a specific performance can be maintained. Whether this is possible for multiple performances, or that in this case a 'sweet spot' is hit, is not clear. This effect is however only found for the base networks and as such might not always be possible for every network. The reason that this does work for these networks can be attributed to the fact that these networks only have a single subnetwork. As almost all intersections are connected to this subnetwork, the resulting signal schemes and cycle times are fairly consistent over the whole network, creating more homogeneous traffic conditions.

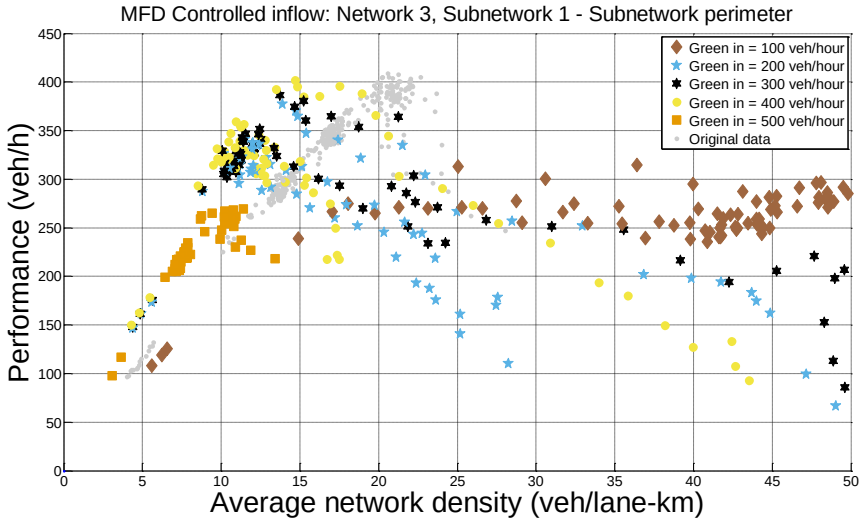


Figure 8.2: *MFD of subnetwork perimeter network 3 – Controlled inflow*

Optimal control

The number of vehicles that can finish their trip under the optimum state, ranges from 20% - 100% more than under the maximum state. Also the rank of the effectiveness of the different strategies, i.e. the number of vehicles finishing their trip, changes. None of the MFD parameters found was ranked in the same way as the number of vehicles completing their trip, illustrating that no relation can be formed between the shape of the MFD and the quality with which the traffic is processed in the network.

The control scenario in which the most vehicles finished, is the same for all three networks and is at 100 veh/h/signal. This is achieved for both the maximum and the optimal state, as is illustrated by Table 8.2.

Table 8.2: *Number of vehicles completing trips*

Green/ control scenario	Subnetwork 1.1		Subnetwork 3.1		Subnetwork 3.1	
	Max	Opt.	Max	Opt.	Max	Opt.
100	22.594	26.577	22.036	28.617	30.349	30.349
200	13.844	24.410	14.407	24.204	15.235	24.308
300	11.296	24.410	17.224	23.238	17.794	28.941
400	19.439	22.896	20.326	20.326	19.519	28.941
500	17.491	19.871	13.537	19.421	18.799	22.387

However, using this control scenario has a serious downside, as can be observed from Figure 8.3, which shows the average travel time, free flow travel time and average delay for each network and control scenario. From these figures it can be found that controlling the network in such a way that the most vehicles can complete their trip, results in a significant increase in the travel time, which can be more than double that of other control scenarios. Only the first control scenario results in high average travel times and delays, while these times level out for all of the other scenarios.

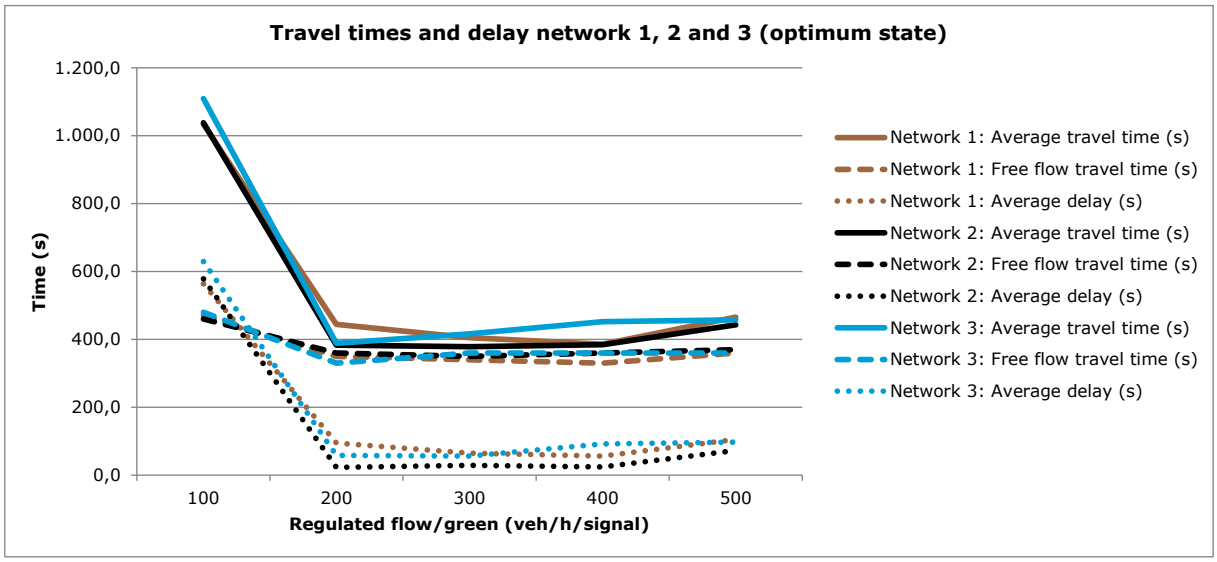


Figure 8.3: *Travel times and delay for network 1, 2 and 3 (optimum state)*

Judging from Table 8.2 and Figure 8.3, the most optimal control scenario would be 200 – 300 veh/h for network 1, 200 veh/h for network 2 and 300 – 400 veh/h for network 3. Interestingly, this does actually match the maximum performance found in the MFDs of each of these networks.

8.2.2 Network 4 and 5

The shape of the MFD

The second set of networks that will be investigated are the two networks consisting of two subnetworks each. The MFDs of the subnetworks and perimeters can be found in Appendix B.

The first thing that is noticed when looking at the MFDs, is that the amount of scatter in the diagram has increased enormously and that the results of the different control scenarios can all be found scattered throughout the diagram. This is opposed to the MFDs of the first three networks, in which the MFDs of the different scenarios are more clustered.

When the MFDs of the subnetwork and perimeter are analysed individually (Appendix B), it can be found that both MFDs have rounded off more, due to a decrease in the performance and increase in scatter. It can also be observed, that for most of the subnetworks, the average network density that can be obtained is higher than that of the perimeter, which differs from the results of the base networks, discussed in paragraph 8.3.1.

Performance ratio

Nevertheless, when comparing the maximum performance of the subnetworks to their perimeters, the same strong relationship is found, as is illustrated by Figure 8.4.
In this figure it can be found that the MFD of the subnetwork reacts in the same manner to changes in the signal settings as the perimeter. This substantiates the hypothesis that the MFD of the subnetwork and perimeter are highly dependent on each other and that changes to the signal settings do not significantly affect this relation.

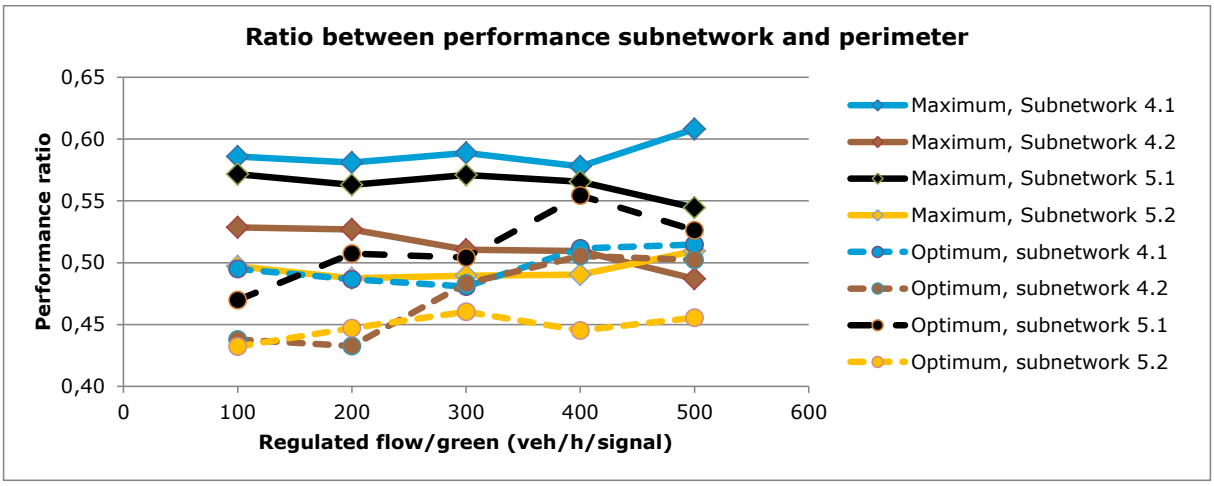


Figure 8.4: *Performance ratio subnetwork/perimeter for network 4 and 5*

That the ratio between the performance of the subnetwork and perimeter is very consistent, can also be attributed to the fact that the performance is quite consistent for the different control scenarios. The coefficient of variation (standard deviation divided by the mean) of the performance in the maximum state, is between 0,03 and 0,07 for all of the types of MFDs, for all networks. In the optimum state this is between 0,07 and 0,18.
From these results it can be derived that the maximum performance is not very sensitive to changes to the signal settings. Although this can be attributed to the rush for the maximum state, this is not the case for the optimum state. In the optimum state, all of the OD-factors found are either 0,13 or 0,25. This means that the total amount of vehicles destined for, or originating from the subnetwork had to be reduced to 0,13, respectively 0,25 times the original demand in order to reach a stable performance. Due to this low amount of vehicles, the network is mostly in free flow, resulting in mostly the same performances.

Optimal control

Paragraph 8.3.1 illustrated that in order to assess the most effective control method, the travel times and the number of vehicles finished provide decent information. This information is presented in Figure 8.5 and Table 8.3 respectively.

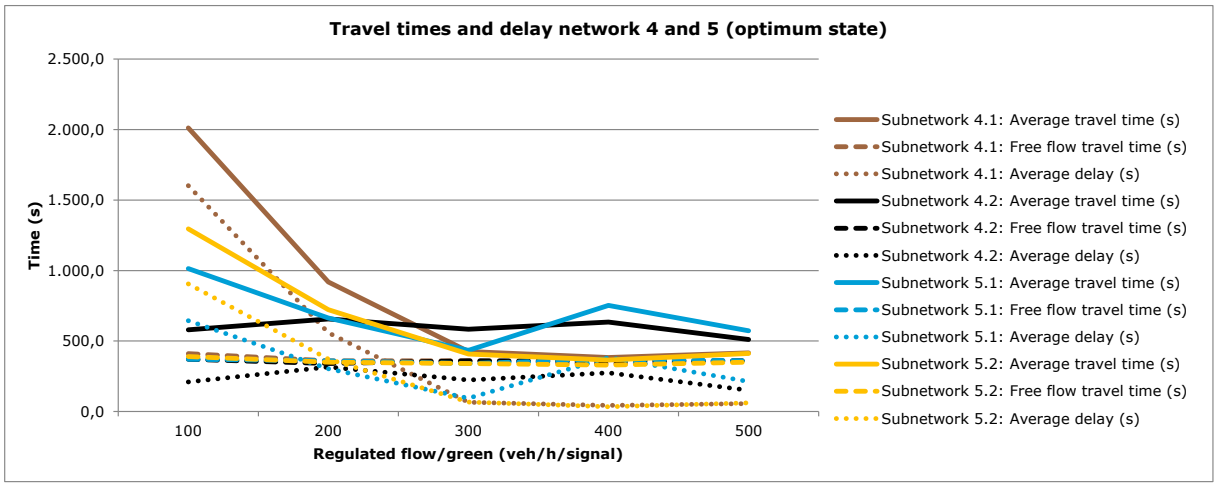


Figure 8.5: *Travel times and delay for network 4 and 5 (optimum state)*

From Figure 8.5, it can be found that the control scenario, resulting in the shortest travel times and smallest delays is at a control rate of 300 - 400 veh/h/signal. Table 8.3 supports these findings, as the number of vehicles reaching their destination is maximised in this way. For the previous networks, it was found that the optimal control scenario was actually the scenario with the highest performance. In this case, the highest performance for all subnetworks is achieved at 500 veh/h/signal, which invalidates the earlier hypothesis.

Table 8.3: *Number of vehicles completing trips in subnetworks of network 4 and 5*

Green/ control scenario	Subnetwork 4.1		Subnetwork 4.2		Subnetwork 5.1		Subnetwork 5.2	
	Max	Opt.	Max	Opt.	Max	Opt.	Max	Opt.
100	2.738	6.844	2.786	5.018	2.708	6.404	2.370	6.651
200	2.896	8.385	3.103	8.398	2.873	8.302	2.324	8.307
300	2.838	10.494	2.905	10.512	3.182	8.302	3.634	8.307
400	3.887	10.494	3.085	10.512	2.812	9.226	3.727	8.307
500	3.277	10.494	2.996	8.807	2.892	8.404	3.034	8.307

8.2.3 Network 6 and 7

As both network 6 and 7 consist of 4 subnetworks each, the results will not be fully presented (as has been done for the other networks). For the resulting MFDs and assessment tables, the reader is referred to Appendix B and Appendix C.

The shape of the MFD

The resulting MFDs of the subnetworks of network 6 and 7 show the same characteristics as those of network 4 and 5. They are highly scattered and the results of the different control scenarios can be found throughout the complete MFD. The optimum state is mostly found at 0,13 and 0,25 times the OD-matrix, and the coefficient of variation of the performance in the maximum state, is again found to be very small, between 0,03 and 0,12 for all of the MFDs. In the optimum state, the CoV is significantly higher and is between 0,03 and 0,70. The main contributors to this high CoV are subnetworks 6.2, 6.3, 7.3 and 7.4.

Regarding the above mentioned two subnetworks of network 6, these are the large and the small subnetwork respectively. The reason that the CoV of these two subnetworks is higher, is that both have one or more control scenarios, in which the highest number of vehicles finishing their trip is found at an OD-factor of 1,00. As we have seen before, the performance at these higher OD-factors is also higher, due to the rush. However, in both these networks, these states are suboptimal, as the number of vehicles finished is only slightly higher than at a factor of 0,13. The result is, that in these states, over 10.000 vehicles were either in gridlock, or could not be loaded onto the network at all, because of heavy spill-back blocking the origins and destinations. In most of the cases, only 6.000 – 7.000 vehicles can reach their destination.

The same line of reasoning applies to subnetwork 7.3 and 7.4. However, the question in this case is, why these two subnetworks do suffer from this problem, while the other two do not. Whereas the difference in size is probably the main contributor in network 6, this cannot be the case for network 7, as all subnetworks are sized equally. This is emphasised even more by the fact that in the simulations for these subnetworks, between 7.000 and 8.000 vehicles can reach their destination, while this is over 11.000 for the other two subnetworks. But as has been concluded before, these differences should most probably be attributed to the stochastic nature of traffic.

Performance ratio

In order to assess the relationship between the subnetwork and its perimeter and how this is affected by different signal settings, the performance ratio has been introduced. To this end, the performance ratio for the different subnetworks of network 6 and 7 are shown in Figure 8.6 and Figure 8.7 respectively. From both figures it can be found that the performance ratio is relatively constant for the different control scenarios, in the maximum state. In the optimum state, some more fluctuations can be found.

The first conclusion that can be drawn from this, is that the relationship between the shape of the MFD of the subnetwork and its perimeter is not strongly affected by different signal timings. The low CoV found earlier, also shows that the performance is relatively constant in most cases, meaning that the performance is not seriously affected by changes to the signal timings and is fairly constant.

The fluctuations found in the optimal state, are caused by the fact that the ratio between the maximum performance of the subnetwork and perimeter changes. The ratios obtained in the optimum state, are almost all found at low OD-factors (0,13/0,25). At these lower factors, the number of vehicles originating from, or destined for the subnetwork under control, are lowered. However, still other traffic uses the perimeter, as it is headed for another destination. The result of this is, that the performance of the perimeter decreases less than that of the subnetwork, lowering the performance ratio. As the performance ratio of the optimum state often is based on different OD-factors, these fluctuations arise.

From this, the second conclusion that can be drawn is that the performance of the subnetwork and its perimeter is affected more by the number of vehicles originating from, or destined for the subnetwork, than by differences in signal settings.

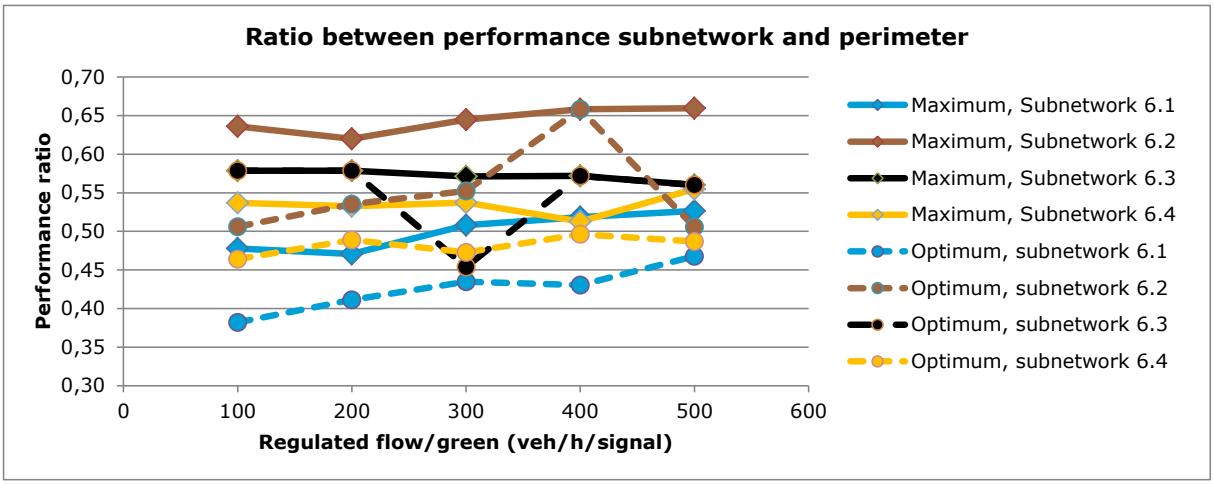


Figure 8.6: *Performance ratio subnetwork/perimeter for network 6*

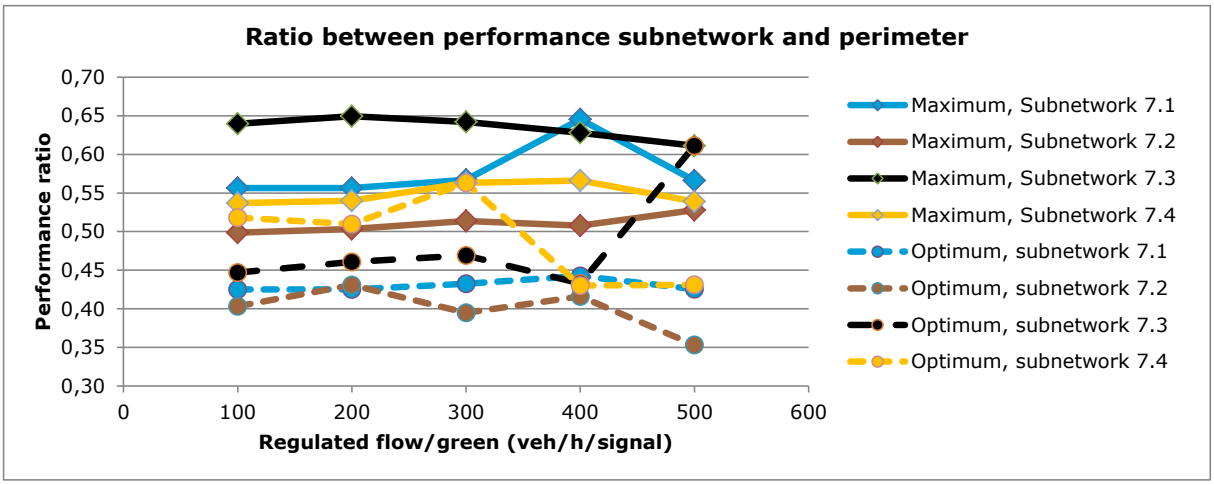


Figure 8.7: *Performance ratio subnetwork/perimeter for network 7*

8.2.4 Conclusion

The general shape of the MFDs of the subnetwork and its perimeter differ from each other, in which the MFD of the subnetwork has a short free flow branch and a long tail with a sharp transition, while the MFD of the perimeter is more rounded. For the networks with a single subnetwork, the shape of the MFD is well-defined and shows only small amounts of scatter. The MFDs for the other subnetworks and perimeters are highly scattered and the effect of different signal settings is not clearly visible in the MFD.

The ratio between the performance of the subnetwork and the performance of the subnetwork is found to be very consistent when different control scenarios are used. The performance however, can be effected by changing the signal timings. This relationship is not linear, and an optimal timing of the signals seems to exist, as a drop in the performance and network density is found, when too much green time is given.

8.3 The effect of controlling subnetwork outflow

Whereas the previous section discussed the effect of changes to the shape of the MFD as a consequence of changing the timing of signals regulating traffic from the perimeter to the subnetwork, this principle is reversed in this section, i.e. the effect of changing timing of signals regulating traffic from the subnetwork into the perimeter will be investigated.

As most of the basics of the analysis have already been covered in section 8.2, the analysis of the effect of controlling the subnetwork outflow will be more comprehensive.

8.3.1 Network 1, 2 and 3

The shape of the MFD

The shape of the MFDs of the different subnetworks and perimeters, do not seem to be substantially different from those obtained for the inflow. Still the MFD of the subnetwork shows a sharp transition from the free flow to the congested branch, while the MFD of the perimeter is more rounded. Also the maximum performance and optimal density for subnetwork and perimeter have not changed substantially.

The main difference is that the amount of scatter is less in these diagrams. This however can be contributed for some extent to the fact that the simulation for the control scenario of 100 and 200 veh/h/signal was identical. The reason for this is, that the signals controlling the outflow from the subnetwork, are all located at the same direction, i.e. the road leading out of the network. Because of this, these signals do not conflict with each other, leading to shorter cycle times than in the inflow scenario, in which all the signals regulating traffic into the subnetwork conflict with each other. Due to the low cycle times, the amount of green time need per cycle is very short and as such is set to the minimum cycle time ($t_{green,min} = 6s$ in our case). This results in the exact same signal schemes and the simulation results are exactly the same.[20]

Another important difference that can be found is that the highest performance for both subnetwork and perimeter is achieved at the control scenarios with the most amount of green time per signal. This is opposed to regulating the inflow, which showed a higher performance at low amount of green times.
Also the shape of the MFD for the different control scenarios is very consistent, making it possible to assess the effect of different signal settings on the MFD. Looking at the MFDs, it can be concluded that the performance is seriously affected by different settings, in which the scenarios with more green time, show an increase in performance.

Performance ratio

The performance ratio of each of the networks is extremely consistent and almost no difference between the maximum state and the optimum state is present, as is illustrated by Figure 8.8. These results confirm once more that a strong relationship between the subnetwork and perimeter is present. The performance ratios are also very consistent for different signal timings, which is in line with previous findings.

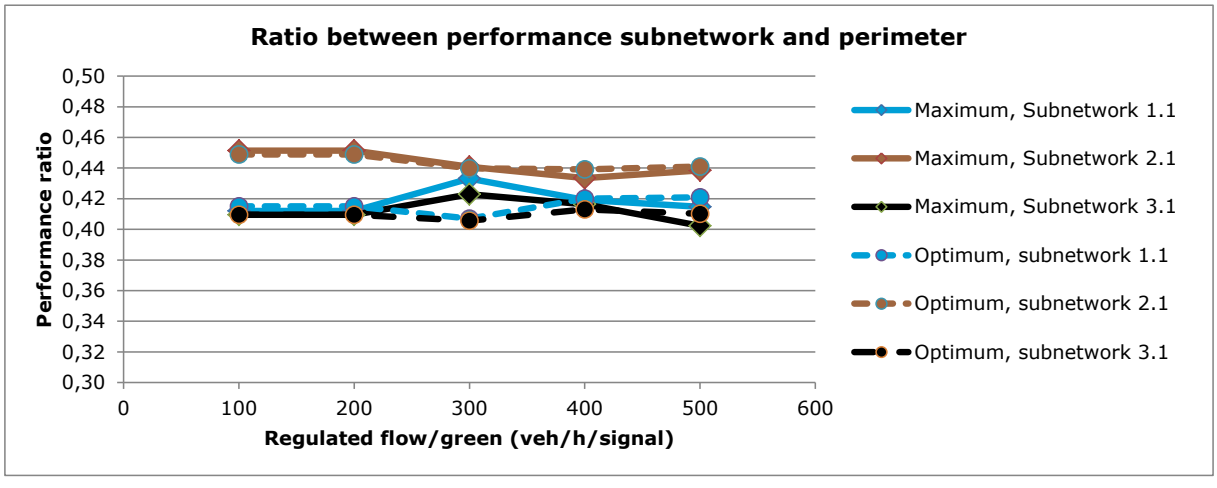


Figure 8.8: *Performance ratio subnetwork/perimeter for network 1, 2 and 3*

[20] This effect can be circumvented by using different random seeds in VISSIM. During these simulations, the same seed was used for every scenario, in order to keep the traffic loading pattern the same.

If these performance ratios are compared to the ratios in Figure 8.1, it is found that the performance ratios are also consistent between controlling the inflow and the outflow. So not only is the relationship of the MFD of subnetwork invariant to changes in signal timings, it also seems to be indifferent to the function (controlling inflow/outflow) of these signals.

Optimal control

As has been found before, the shape of the MFD and performances found, do not directly show which control scenario can process the traffic in the network the most effective.
In contrast with regulating the inflow, the travel times are all very short between 350 – 450 seconds. From Figure 8.9 it can be seen that the travel times are very consistent for the different scenarios. As such, the most appropriate control scenario is the one in which the most vehicles can reach their destination. In all of the three networks, this is the control scenario in which the signals are set to a flow of 500 veh/h. In network 1 and 3, a total of 35.000 vehicles could reach their destination and for network 2 this was 29.000. For all of the networks, this is higher than for controlled inflow.
The number of vehicles that can finish their trip increases when the amount of green time increases. This is in line with the maximum performance achieved for each of the control scenarios.

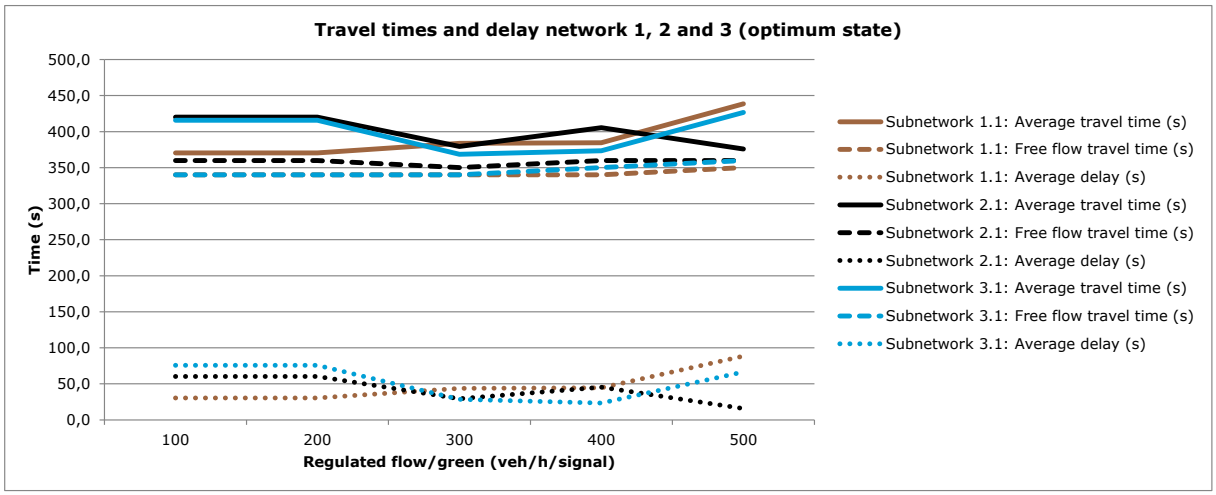


Figure 8.9: *Travel times and delay for network 1, 2 and 3 (optimum state)*

From these results it can be concluded that controlling the flow from the subnetwork to the perimeter is much more effective than vice versa. Not only are the travel times shorter and is the number of vehicles processed higher, the travel times are also very stable, meaning that a whole range of different restrictions can be placed on the signals controlling the subnetwork, without severely affecting the performance of either part. As such, this method is much more flexible than controlling subnetwork inflow. It should be noted that this is in line with insight 2 of Daganzo (2007): *"… system output is maximized when flow is at capacity only along road stretches with the greatest exit rates …".* As controlling the subnetwork outflow does not restrict the flow on the perimeter (the road stretches with the greatest exit rates) as much as controlling the subnetwork inflow, the perimeter can come closer to its capacity.

8.3.2 Network 4 and 5

The shape of the MFDs of the different subnetworks and perimeters of network 4 and 5 when controlling the subnetwork outflow, do not significantly differ from those of controlling the inflow. The MFD is still highly scattered and results of all the scenarios overlap. Also the maximum performance and optimal network density are not substantially different, which means that the performance ratios are unaffected as well.

For the optimum state, again very low OD-factors are found (0,13 – 0,25) at which the most number of vehicles could finish their trip. Opposed to earlier findings, the network output (number of vehicles that could finish their trip) is slightly higher when regulating the inflow, than when the outflow is regulated. The average travel times show strong fluctuations for different control scenarios and range between 400 and 1.500 seconds. No apparent trend for these travel times is to be found, as can be seen in Figure 8.10. As such it is concluded that from these results, no statement about the effectiveness of controlling the outflow can be made.

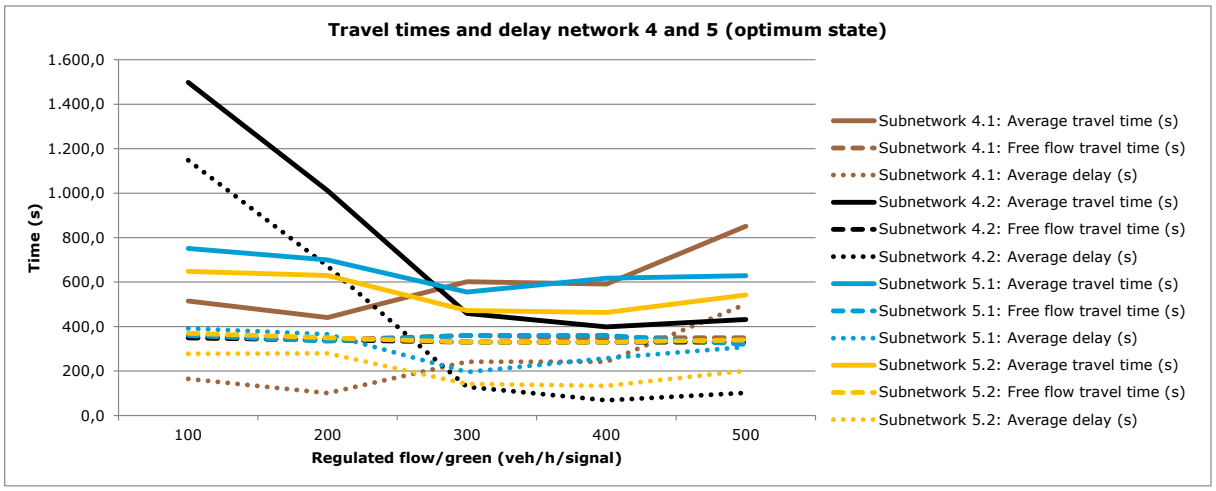


Figure 8.10: *Travel times and delay for network 4 and 5 (optimum state)*

8.3.3 Network 6 and 7

For most of the subnetworks and perimeters of network 6 and 7, no discernible differences between the MFDs of the inflow and outflow were found. As before, and also for the previous two networks, the MFDs are highly scattered, have a sharp transition for the subnetwork and are more rounded for the perimeters. However, a notable difference is found for the MFD of subnetwork 6.3 (see Figure 8.11), which is the smallest subnetwork of all.

As is shown by Figure 8.11, a very high performance is achieved when the outflow rate of the signals is set to 400 – 500 veh/h. This high performance is achieved in both the subnetwork and the perimeter. Apart from that, in these scenarios two times as much vehicles can finish their trip as under any other control scenario. The reason as to why these changes can have such a substantial impact on the network is not clear, especially as subnetwork 6.3 is very small.

Nevertheless, it should still be noted that in this case the MFD does give a clear indication that the network has improved.

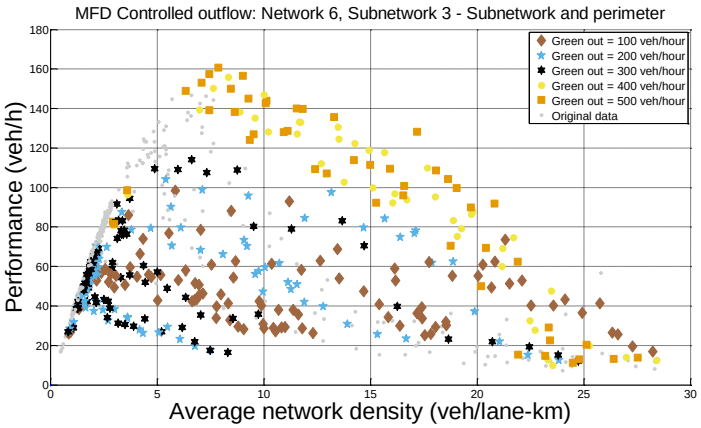


Figure 8.11: *MFD for controlled outflow of subnetwork 6.3 – subnetwork and perimeter*

Regarding the other subnetworks and perimeters, for the performance, performance ratio, travel times and vehicles finished, no clear relationship can be formed. For some of these subnetworks, the situation improves by using outflow, while others do not. This does not only vary per subnetwork, but also per control scenario, making it very hard to make any statement whether performance improves under inflow or under outflow restrictions. These results however, are in line with earlier findings that the MFD does not seems to be affected when the timing of signals are changed.

8.3.4 Conclusion

Controlling the subnetwork outflow seems to have the same effect on the MFD as controlling the inflow. The MFDs of subnetworks 1.1 to 3.1 are well-defined, while the other MFDs show a lot of scatter, due to the large differences of the amount of traffic loaded onto the network. As with the inflow, the ratio between the performance of the subnetwork and the perimeter is highly consistent, showing that a strong relationship between both of them exists. Differences in the maximum performance are found, but can be mainly attributed to differences in the number of vehicles loaded onto the network. The effect of different signal timings on the performance seems negligible.

8.4 Controlling inflow vs. controlling outflow

8.4.1 Inflow versus outflow

Throughout section 8.2.4, the effect of controlling the outflow has already been compared to the effect of controlling the inflow for some of the subnetworks. The main finding is that no general principle does apply for the different subnetworks and perimeters.

For the first three networks, controlling the outflow resulted in a somewhat higher output as controlling the inflow. The average travel times were also generally lower and highly consistent for

the different outflow scenarios. As a result, controlling the outflow is more preferable for these networks, than controlling the inflow.
For the subnetworks and perimeters of network 4 and 5 slightly better results are achieved when the subnetwork inflow is restricted. Although the differences are not very large, the output is a little bit more consistent over the different control scenarios. On the other hand, the average travel times and delays strongly fluctuate, making neither strategy better than the other.
Also for network 6 and 7, neither control strategies proves to be systematically better than the other. For some subnetworks, under some control scenarios, controlling the inflow performs better, while in other control scenarios controlling the outflow improves the output of the network. Notable exceptions are subnetwork 6.2 and 6.3, for which the network output was doubled when controlling the subnetwork outflow at a rate of 400 - 500 veh/h/signal.

8.4.2 Relation between the shape of the MFD and network output

It is investigated if a relation between the network size, perimeter length, number of signals, et cetera can be found. However, none of these parameters, or any combination of them, can explain why certain strategies work better than others. Furthermore the different MFD parameters (maximum production and performance, optimal accumulation and network density), have all been ranked and are compared to the ranking of the output under the different control scenarios. For none of these parameters, a consistent relationship is found, which applies to all of the subnetworks.
However, it is found that networks 1, 2 and 3 do have a consistent relation between all of the above mentioned parameters and the output of the network, in case the outflow is controlled. This implies that if the MFD is well-defined, i.e. does not have a lot of scatter, it can be used to assess the effectiveness of different control strategies.
With respect to the above, it should be noted that for networks 4 to 7, the OD-factor in the optimal state is very low (0,13 – 0,25), meaning that the number of vehicles originating from, or destined for the subnetwork under investigation is very low. The result of this is, that the amount of traffic using the perimeters, as they are headed for other subnetworks, is very large in comparison to the traffic heading for the subnetwork under investigation. This can strongly bias the results, as this leads to a very uneven distribution of traffic over the network.

8.4.3 Conclusion

The results from the simulations show strongly mixed results. For some subnetworks the network output is increased by controlling the inflow, while others perform better when the outflow is controlled. As such, no statement can be made which strategy performs better. Also the relation between the shape of the MFD and the effectiveness of each strategy has been investigated, but no relation could be found.

In line with earlier observations, the conclusion that can be drawn is, that a control strategy, such as perimeter control, cannot simply rely on either controlling the inflow or the outflow to any subnetwork, as the effectiveness of either strategy depends on an intricate combinations of factors. Based on the results obtained from the simulations, it is not clear which criteria are to be used, in order to choose between controlling the inflow or the outflow to the subnetwork.

The second conclusion is that is very hard to use the MFD to assess the effectiveness of a control strategy, as none of the parameters seems to have a strong relationship to the network output or travel times. Although a consistent relation between the output and the performance was found for networks 1 – 3, these results are too limited to state that well-defined MFDs can be used to assess the effectiveness of a control strategy.

8.5 The effect of signal settings on the accumulation

8.5.1 General

In paragraph 4.6.6 it has been hypothesised that increasing the amount of green time should result in a higher optimal accumulation for the controlled part of the network, whereas the accumulation of the uncontrolled part should stay the same. Knowing how the optimal accumulation is affected when making changes to the signal settings is very important, when using a control strategy in which signal timings are adapted, in order to keep the accumulation in a part of the network optimal. If the hypothesis would hold, the implication is that one cannot use the optimal accumulation as the control parameter, as this parameter changes when adaptations to the signal settings are made.

8.5.2 Analysis of the results

In order to test the hypothesis, the optimal accumulation for both the subnetwork and the perimeter are related to the amount of green time given. For this analysis, the results of the optimal simulation are used, because in this case the network output is the highest, meaning that the timing of the signals is best fitted to the traffic demand. The results are shown graphically in Figure 8.12 and Figure 8.13.[21] From these figures, it can be found that the accumulation does indeed change when signal timings are adapted, which in some cases can be very substantial.

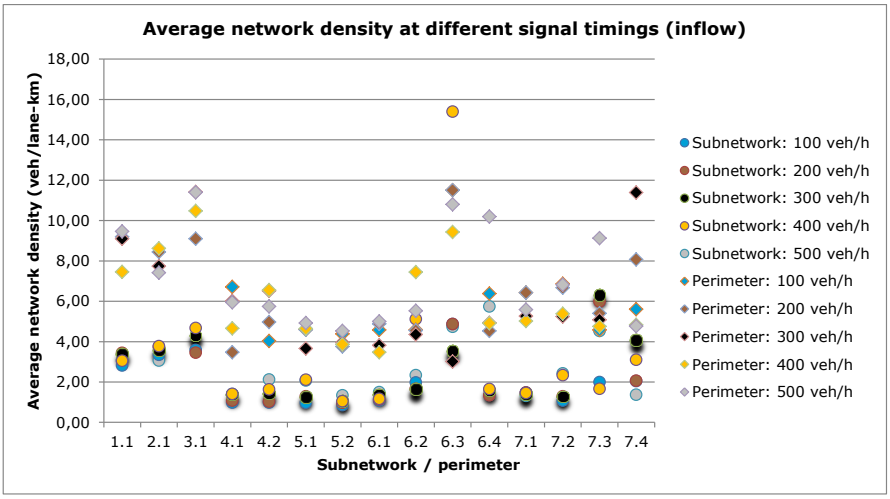


Figure 8.12: *Optimal network density at different signal timings (controlled inflow)*[22]

[21] It should be pointed out, that in these figures the average network density is used instead of the accumulation. This is done in order to keep the different results to scale.

[22] Data for the perimeter at 100 veh/h for subnetwork 1.1, 2.1 and 3.1 have been removed in order to keep the size of the graph in reasonable limits. These points are 23.65, 26,62 and 36.42 veh/lane-km respectively.

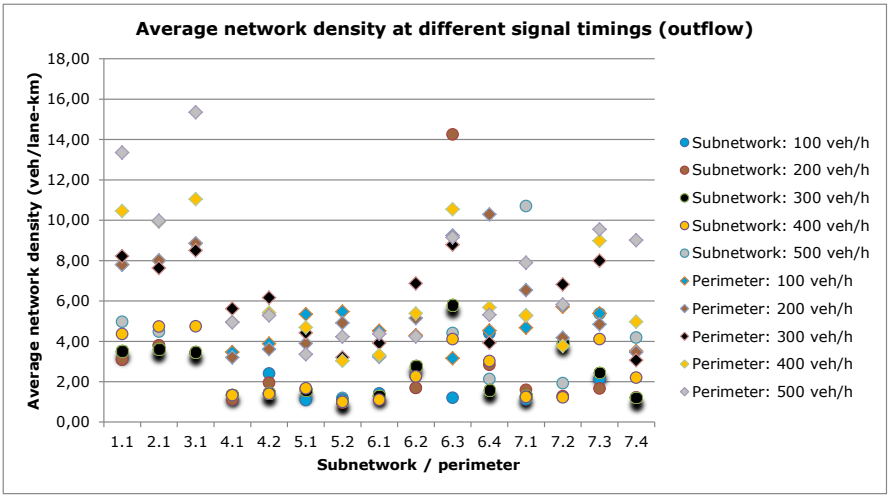


Figure 8.13: *Optimal network density at different signal timings (controlled outflow)*

Inflow

For the inflow it is found that the accumulation for most of the subnetworks is quite consistent for different signal timings. For the perimeter a strong increase is often seen when the inflow is set to 400-500 veh/h. The reason why the density in the perimeter increases at these higher control rates, is because in order to accommodate these high rates, the green time for the remaining signals is strongly reduced, thus reducing the overall flow on the perimeter and increasing the density. For the subnetworks of network 1 to 3, very high densities were found at a signal timing of 100 veh/h. The reason for this is that because of the low inflow to the subnetwork, spill-back occurs and the average speed drops, increasing the travel times, which leads to higher accumulations.

Outflow

When the outflow is controlled, the optimal average network density is also fairly consistent for the subnetwork, i.e. changes to the signal timings do not have a strong effect on the density and accumulation. For the perimeter, again a much more chaotic image is obtained, showing a wide variety of optimal network densities for the different control scenarios. Apart from the perimeters of network 4 and 5, every perimeter is strongly affected by this. However, no strong relationship between the density and the control scenario can be formed, as this strongly varies for the different subnetworks and perimeters. Even in network 7, which has 4 roughly the same shaped subnetworks, these results are far from equal.

8.5.3 Conclusion

The accumulation and/or average network density of subnetwork and perimeter are influenced by different signal timings. In most cases, the effect is stronger at the higher signal timings. The reason for this is, that the amount of green time for the uncontrolled turns is reduced in order to accommodate the controlled turns, which results in spill-back and gridlock in other parts of the network.

The optimal accumulation of the perimeter reacts stronger to changes in signal timings than the optimal accumulation of the subnetwork. This is because the optimal accumulation depends on the average travel times, which in turn are highly dependent on signal timings. Because the density of intersections on the perimeter is much higher than in the subnetwork, the optimal accumulation of the perimeter reacts stronger to changes in signal timings than the subnetwork does.

The results from this paragraph confirm our hypothesis that changing the signal timings can cause a substantial shift in the optimal accumulation. In turn, this validates the assumption that using the optimal accumulation as a control variable is not always the best strategy.
The above results do show that the optimal accumulation of the subnetwork is more consistent than the perimeter, making it easier to control the number of vehicles in the subnetwork. But still, in order to control the accumulation in the subnetwork, signals should be adapted in the perimeter as well, which can have highly adverse consequences.

8.6 Conclusion

8.6.1 Hypothesis

In paragraph 4.6.6 it has been hypothesised that increasing the amount of green time, when controlling the inflow, should result in a higher production, performance, optimal accumulation and optimal average network density. The maximum accumulation is assumed to remain unchanged and the deviation of density to become lower. The MFD of the subnetwork is assumed to remain unchanged, as no physical changes to the structure or the timing of the signals is made.
For controlled outflow, the effects are assumed to be inversed, i.e. the subnetwork improves, while the perimeter remains unchanged.
Furthermore it is expected that an optimal signal timing should exist, at which the performance of both the subnetwork and perimeter is maximised. After this optimum is transgressed, increasing the amount of green time should only lead to lower performances.

It has been found that for almost all of the subnetworks and perimeters, the maximum production, performance and optimal accumulation and average network density have decreased, and that the amount of scatter has increased substantially. Although for some networks an increase in these parameters is witnessed and network output is larger than when no control was exerted, no improvements for all networks were found and therefore the hypothesis must be rejected.

8.6.2 The effect of signal settings on the shape of the MFD

From the results of the simulations is concluded that the shape of the MFD is affected by changing the signal timings. However, a number of different effects have to be taken into account.
The *maximum performance* of a network does not directly improve by changing the signal timings, but because the signals have been changed, more traffic can be processed in the network, which in turn leads to a higher performance, but if the amount of traffic stays the same, no substantial differences in the performance will be found. On the other hand, bad signal timings should have a direct effect on the performance, as flow is seriously hampered. However, due to 'rush' this effect cannot be made completely clear.

The *optimal accumulation* does change when signal timings are adapted, even when the amount of traffic remains unchanged. The direct effect of improving signal settings is a higher average network speed, which improves travel times and results in a lower optimal accumulation. The indirect effect is, that more traffic can be processed by the network, resulting in even more substantial changes to the optimal accumulation.

8.6.3 The relationship between the subnetwork and its perimeter

The relation between the shape of the MFD of the subnetwork and its perimeter is found to be very strong. When signal timings and the amount of traffic destined for, or originating from each of the subnetworks are changed, the subnetwork and the perimeter react in the same way. This relation is expressed as the 'performance ratio', which is the maximum performance of the subnetwork divided by the maximum performance of the perimeter. This ratio is found to be almost constant under different signal timings. Only when the amount of traffic is severely reduced, this ratio decreases. However, the latter is only found for the networks with multiple subnetworks. The reason why this ratio decreases, is because the perimeter is used by traffic, other than traffic heading for the specific subnetwork.

8.6.4 The effect of controlling the inflow and outflow

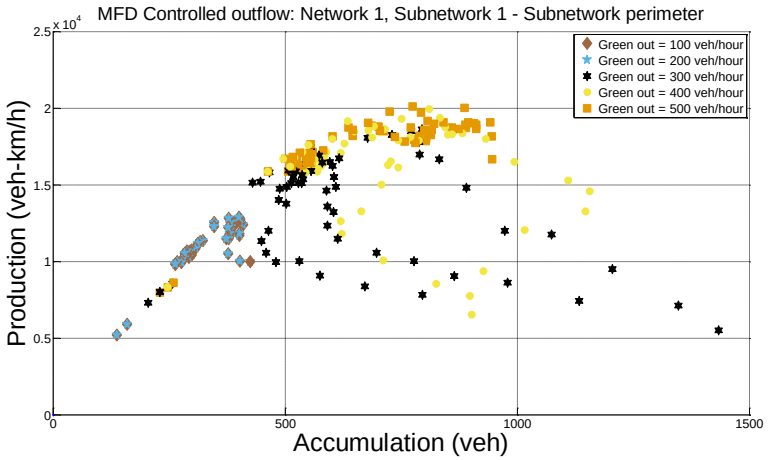
Controlling the inflow or outflow to the network does not seem to result in improvements in the network. Also, for most of the networks, no significant differences between controlling the inflow and the outflow are present. This implies that controlling either the inflow or the outflow to a subnetwork does not give the same results. The best strategy seems to be very specific for each situation and no general principles seem to apply.
Furthermore it has been found that it is very hard to use the MFD to assess the effectiveness of a control strategy, as none of the parameters seems to have a strong relationship to the network output or travel times. Although a consistent relation between the network output and the performance was found for networks 1 to 3, these results are too limited to state that well-defined MFDs can be used to assess the effectiveness of a control strategy. Nevertheless, these results are still very encouraging.

8.6.5 Using the optimal accumulation as control variable

The optimal accumulation of the perimeter reacts stronger to changes in signal timings than the optimal accumulation of the subnetwork. This is because the optimal accumulation depends on the average travel times, which in turn are highly dependent on signal timings. Because the density of intersections on the perimeter is much higher than in the subnetwork, the optimal accumulation of the perimeter reacts stronger to changes in signal timings, than the subnetwork does.

The MFDs of various networks show that changing the signal timings can cause a shift in the optimal accumulation. In turn, this validates the assumption that using the optimal accumulation as a control variable is not always the best strategy. A good example of this is given in Figure 8.14 below, for which the data is shown in Table 8.1. The optimal accumulation at control rate of 400 veh/h is over 75% higher than at a control rate of 100 veh/h!

It is concluded that using the optimal accumulation as a control variable is not viable without a-priori knowledge of the effect of these changes. This because the control variable (and the complete MFD for that matter), changes when adaptations to the signal timings are made.



Green (veh/h)	Optimal accumulation (veh)
100	457
200	457
300	795
400	810
500	775

Figure 8.14: *Example of differences in optimal accumulations under different signal settings*

Nevertheless, it has been found that the changes in the MFD of the subnetwork and perimeter are strongly correlated and that they react in the same way when changes to the signal timings are made. What this implies is that the optimal signal timing is the same for subnetwork and perimeter and that no trade-off between the two should have to be made. Simply finding out what signal settings result in the highest performance for both could result in substantial improvements in the network.

9 Practical implications

Based on the insights gained in the previous chapters, this chapter will discuss the practical implications of these results and how they can be used in everyday life. To this end, section 9.1 will start by discussing what data is needed in order to construct a decent MFD. In section 9.2 it will be discussed how the MFD should be interpreted and what information can be derived from it. In section 9.3 the different fields of application will be reviewed and how the MFD can be used in these cases, after which section 9.4 will discuss the implications of the results of this thesis for the MFD. In section 9.5 the big question will be answered, whether or not the MFD is practicably usable. Section 9.6 will end this chapter with the usual conclusion, summing up the main findings.

9.1 Obtaining MFD data in urban networks

9.1.1 Data requirements

In chapter 3.1, a review of different types of MFDs has been made, and the data from which these MFDs have been constructed. Generally speaking, the most important parameters are the flow (q) and the density (k). Using this data, one can construct a MFD for the network. However, the part of the network represented by each measurement is not taken into account, and as such, the resulting MFD does not always represent the network properly. Improvements can be made by weighing the data and taking the length of the network (l) into account which is represented by each measurement.

The main problem of this is, that it is very hard to find out which parts of the network can be properly represented by a single measurement, especially when the density of the detectors is very low. The main condition for the representation, is that the traffic conditions are homogeneous and equal to the measurement. However, if no data is available for other parts, this analysis cannot be made and can only rely on assumptions.

Another problem is that in reality the most detectors are present on the main arterials and not in the subnetworks, making the resulting MFD only representative for the arterials and not for the subnetwork. As we have seen, the length of the roads in all subnetworks combined is often far higher than for the arterials. Because of this, the MFD of the total network often has a stronger resemblance to the subnetwork MFD, than to the arterial MFD.
Basically this means that a MFD obtained from only arterial data is not representative for the total network.

9.1.2 Data sources

The data needed to construct a MFD can be derived from various sources. The most important source are the detectors placed in the network. Also data from detectors at traffic lights can be used, as they also provide the data needed.

However, this does still not solve the problem of the lack of subnetwork data, as the amount of detectors in these areas is generally very low. One method to obtained the data needed, is by using floating car data, like mobile phones and GPS tracking. The main problem with this is, that this data is only scarcely available and is not sufficient, as the average density and accumulation still cannot be measured, as not all of the vehicles are covered.

9.1.3 Additional data acquisition methods

As has been shown in paragraph 3.1.3, the accumulation can also be derived from the average flow/performance in a part of a network and the average travel time in that network (see equation (3.5)). The performance of a subnetwork can be measured by taking data from detectors of traffic signals along the perimeter. Travel times can be calculated by the little floating car data that is available. Another method is to temporarily set up a number of traffic cameras in a subnetwork and measure the average travel times over different sections. Whenever a vehicle 'disappears', it has reached its destination and its travel time between the perimeter and destination can be estimated. Apart from that, these cameras can also be used to estimate the average flow at certain trajectories. As the infrastructure of subnetworks does not change very often and is not strongly affected by traffic signals, the data obtained should be usable for a decent amount of time. A third method is to equip a single vehicle with GPS and keep it moving between the subnetwork and perimeter for one or more days and calculate the average travel times between the perimeter and various destinations within the subnetwork. Using these travel times, combining them with the flow measured at the signals at the boundary of the network and adding the total network length, derived from digital maps, a MFD for the subnetwork could possibly be constructed.

9.2 Interpretation of the MFD

9.2.1 Different types of MFDs

In paragraph 3.1.4 it has been shown that a couple of different types of MFDs can be constructed, depending on the parameter that is used on each of the axes. The most commonly used MFD for existing networks, relates the performance (PF) to the network density (K), in which the data is not weighted. This PF-K MFD is used most, because each detector point does not have to be linked to a part of the network, in order to obtain the production or accumulation. The data only has to be averaged. Apart from that, the PF-K MFD takes the length of the network out of the equation, making it possible to compare the results from different networks, or different parts of the same network to each other.

The main drawback of the PF-K MFD is that it does not convey any information on the optimal accumulation. The optimal accumulation is important to know, as it shows how much vehicles should be allowed in any part of the network, giving a stronger indication of the sensitivity of different parts of the network, to changes in demand.

9.2.2 What does the MFD tell us and what does it not

The MFD gives an indication of the quality with which the traffic in the network can be processed. It shows how much distance can be travelled in the network per unit of time, depending on the

number of vehicles in that network. The amount of scatter in the MFD gives an indication of how well the traffic is distributed over the network.

What the MFD does not convey, is any information on experienced travel times or delays, between origins and destinations. For example, if the production increases, it simply means that more distance can be travelled by all vehicles. Whether this means that vehicles can reach their destination more easily, or that they can drive two times around the city instead of once in the same amount of time cannot be properly derived. Even more so, we have seen that adding new infrastructure can decrease travel times, leading to a lower accumulation. It can also decrease detours, leading to a part of the network to experience lower flows, resulting in a lower production and/or performance.[23] Only when the amount of traffic increases, higher productions should be witnessed.
Using the MFD as a tool to evaluate the quality of a network, before and after certain measures have been taken, is therefore not always suitable, as additional trips have to be made in order to make these improvements visible.

9.3 Fields of application

The two most important uses for the MFD are using it to evaluate what happens on a network wide scale, when making changes to the network (infrastructure, road-works, dynamic traffic management, et cetera) and using it as an input for traffic control.

9.3.1 Evaluation of network changes

As has been shown in the previous section, using the MFD to evaluate how the network is affected when changes are made to the infrastructure, or the way it is controlled, is not straightforward and can only be done when a considerate amount of time has passed, and the original flows have been restored. By Qian (2009) it has already been advised to use the MFD as a supplement to the more conventional assessment methods.
Based on the results and insights obtained throughout this thesis, it is advised to not use the MFD at all as an evaluative tool. On the one hand because of the aforementioned problems with the MFD and on the other hand because the amount of scatter in most MFDs does not make it possible to reach a strong conclusion.
If however it should be possible to obtain well-defined MFDs with little to no scatter, all the way to the congested region (as in Figure 2.3), than it can be used as an evaluative tool, as this means that traffic is distributed evenly over the network and therefore all parts of the networks are equally used and all taken into account.

9.3.2 Traffic control

So far, the most promising traffic control strategy, that can use the MFD, seems to be perimeter control. When a full MFD can be constructed for an area, the optimal accumulation can be derived, which is the number of vehicles in the network that maximises the production or performance.

[23] It should be pointed out, that this not work the other way around, i.e. it cannot be said that if the performance decreases, the network has improved.

When the optimal accumulation is known, the number of vehicles in the network can be controlled precisely by holding back excess traffic along the border of the controlled network.

Another interesting traffic control measure using the MFD, is rerouting of traffic. If the optimal accumulation of different parts of the network is known, it can be assessed if the optimal accumulation is transgressed and if so, vehicles can be routed around the over-accumulated part of the network, by providing route information to vehicles.

As has been illustrated, the number of vehicles loaded onto the network can have a serious impact on the shape of the MFD and the maximum performance found. The optimal accumulation found in these situations can also be higher than the actual optimum. When the network is controlled based on this accumulation, it will inevitably end up in gridlock.
Apart from that, it has been found that making changes to the signal timings has a substantial impact on the shape of the MFD and can greatly reduce the optimal accumulation, which is the actual control target.
As such, it is recommended to not use the MFD for perimeter control, unless a strong relation between signal settings and the shape of the MFD can be obtained, or if the effect of changing the signal timings on the shape of the MFD can be properly predicted beforehand.

9.4 Implications for using the MFD in practice

Throughout this thesis, some conclusions have been drawn, which not only affect the way in which the MFD should be created and be used, but also could have some practical consequences. Based on the literature study, the assessment of the MFD and the results of the simulation, the following paragraphs set forth a number of implications of these results for using the MFD in practice.

9.4.1 Implications regarding the creation of the MFD

MFDs for subnetworks can possibly be constructed from perimeter data

In paragraph 9.1.1 it has been pointed out that data for subnetworks is often too scarce to construct a MFD of the subnetwork. However, it has been found that a strong and consistent relationship between the maximum performance of the subnetwork and the perimeter exists. If this ratio can somehow be determined, then the maximum performance of the subnetwork can perhaps be derived from that of its perimeter, for which more data is commonly available. Using the average speed in the subnetwork, the optimal accumulation can be calculated as well, especially as we have seen that the slope of the MFD of subnetworks is very linear in free flow.
As it also has been found, that the average density at which the subnetwork is in gridlock, is roughly between 20 – 25 veh/lane-km, the outer envelope of a subnetwork can theoretically be constructed.

The MFD cannot be simply created using network characteristics

MFDs for subnetworks and perimeters have been found to have a strong resemblance to one another. However, although a number of factors has been found to influence the shape of the MFD, none of these factors can really be quantified, as even (sub)networks which have roughly the same

characteristics can show different performances. These differences cannot be explained by any of the factors, resulting in the conclusion that the stochastic and dynamic nature of traffic still has a strong impact on the shape of the MFD, and that the MFD does not simply rely on network characteristics alone.

9.4.2 Implications regarding the use of the MFD

Be careful using unfiltered MFDs

As has been stated before, one should be very cautious using the MFD as it is found from data. As has been shown, overestimations of productions and accumulations can be made, due to differences in loading patterns. Data points created from these traffic states should be removed first, in order to obtain a MFD that can be properly used. Further research should study if these states can be found in real networks as well.

Use different types of MFDs to make an assessment of the network

A single type of MFD does not give full information of the network. The PF-K diagram does not convey how much vehicles can actually be stored in each part of the network, while the PD-A diagram does not show which part of the network performs best and should be kept at optimal accumulation under all circumstances.
Only when using both these MFDs, a decent insight in the different qualities of each part of the network can be obtained.

Do not use the MFD to make any statements on improvements from the individuals perspective

As it has been shown, no relation between the production, performance, accumulation or network density and the average travel times exists. As such, improvements in the maximum production or optimal accumulation, do not mean that the situation has improved for individual travellers.

Refrain from making statements on the performance of the total network, if only data from the arterials is present

It has been found that the general shape of the MFD of subnetworks and arterials strongly differ from each other. As the MFD of the total network is a sum of both these parts, the resulting MFD will show the strongest resemblance to the MFD based on the largest part of the network, which most often are the subnetworks. As such, the arterial MFD is actually less representative for the total network than the subnetwork MFD. Although most of the data can be gathered for the arterials, these measurements alone are insufficient to make any statement on the performance of the network as a whole. Therefore, network wide control strategies should not be based on these results as well.

Be reticent to use the MFD to support control strategies

As has been shown throughout this thesis, the MFD is affected by a number of factors, each changing the shape of the MFD in their own way. How each of these factors impacts the MFD is still not fully understood. Apart from that, some control strategies change the shape of the MFD itself, making it not possible to use this MFD as an input for these strategies, e.g. when the timing of

signals between a subnetwork and perimeter are changed in order to maintain a certain accumulation in one of these networks, the shape of the MFD will change and a different optimal accumulation is found.
Another problem is that the number of vehicles loaded onto the network can have a serious impact on the shape of the MFD and the maximum performance found. The optimal accumulation in these situations can be higher than the actual optimum. When the network is controlled based on this accumulation, it will inevitably end up in gridlock.

Be careful to use the MFD to support policy-making decisions
As has been stated in the previous point, the effect of different factors impacting the shape of the MFD is still not fully understood. Using the MFD as a basis for policy-making decisions, or to evaluate policies, is therefore advised against. This because the shape of the MFD can change in the same manner due to different factors. As such, changes to the shape of the MFD cannot always be attributed to specific measures that have been taken.

Take caution in using the MFD for evaluation studies
Using the MFD to evaluate how the network is affected when changes are made to the infrastructure, or the way it is controlled, is not straightforward and can only be done when a considerate amount of time has passed, and the original flows have been restored.
It is advised to not use the MFD at all as an evaluative tool. On the one hand because of the problems mentioned in paragraph 9.2.2 (lower production caused by adding infrastructure) and on the other hand because the amount of scatter in most MFDs does not make it possible to reach a strong conclusion. Having said that, well-defined MFDs with little to no scatter can be used as an evaluative tool.

9.4.3 Implications for network design and traffic control

Be careful in adding additional arterials or upgrading roads
For the networks in which additional arterials through the centre of the network were placed, a reduction in the total network performance was witnessed. An important reason for this is, that central arterials draw more traffic than the outer arterials, causing a more uneven distribution of traffic over the network. Another reason is that the number of possible routes is reduced, and more routes cross the intersections on the central arterials, which causes traffic to cluster around these intersections. This problem is most evident for traffic travelling between different subnetworks.

The obvious benefits of adding arterials is improving traffic flow going through the network and reducing travel times. Also the safety, noise, pollution and overall liveability of subnetworks/ neighbourhoods improves, which in itself should outweigh the decrease in traffic performance and accumulation of the network. Simply put, traffic jams are the consequence of our desire to live in a safe environment.

Nevertheless, it is something that can be kept in mind when designing new infrastructure.

Try to spread demand of vehicles in space and time

In line with the above, a good control strategy should aim to create a more even distribution of traffic over the network. If traffic is distributed evenly, the network can maintain in free flow at higher accumulations. By rerouting traffic, or slightly changing signal timings, more homogeneous conditions can be created. Also spreading traffic over time will result in a slower change of the traffic demand patterns and should result in more homogeneous traffic conditions, as peak demands are avoided, which as we have seen can be particularly nasty and often result in heavy congestion.

Changing signal timings can affect the network performance

It has been found that the MFD is sensitive to changes in signal timings, in which the accumulation is much more sensitive to changes than the performance. However, the MFD of both the subnetwork and the perimeter often reacts in the same manner to changes in the signal timings, and an optimal timing seems to exist, which maximises the performance for both the subnetwork and the perimeter. Finding these optimal timings can greatly improve the output of a network. On the other hand, this does not always improve the average travel time.

No single strategy for perimeter control should be used

When a strategy such as perimeter control is employed, one cannot simply use the accumulation found in the MFD, to control the inflow to, or outflow from the subnetwork. Whether the subnetwork or perimeter should be kept at optimal accumulation should depend both on the accumulation and the performance of either part of the network. The part of the network with the highest performance should be kept at optimal accumulation, while the part of the network with the highest maximum accumulation should be used to store traffic in case congestion sets in. Although mainly the perimeter has the highest performance and the subnetwork the highest accumulation, situations are found where the optimal accumulation of the subnetwork is lower than that of the perimeter, meaning that the subnetwork should be controlled instead.

9.5 Usability of the MFD

At this point, the reader might have concluded from the above that the MFD in its current form, is not the best traffic management tool available. Currently the different factors influencing the shape of the MFD are not fully understood and hard to quantify. Also the complete dataset needed to obtain MFDs is currently not available for complete urban networks.

Nevertheless, when a well-defined MFD can be achieved with low amounts of scatter, than the MFD can be usable and can provide very important information regarding the overall traffic operations in a network. Especially the optimal accumulation can be a very important parameter, as it is an important input factor for control strategies.

However, obtaining a well-defined MFD for heterogeneous networks seems to be nearly impossible, as the dynamics of the network still seem to have a strong impact on the shape and can still not be fully explained by taking the distribution of traffic over the network (deviation of density) into

account. Even in a controlled simulation environment, no well-defined MFDs were obtained. And as such it is highly questionable if this can be done for real networks.

Whether or not the MFD should be used in practice is left to the reader, but it is advised to be reticent in making strong statements based on the MFD.

9.6 Conclusion

When looking at the practical side of the MFD, the first hurdle is how the data should be obtained to create MFDs. Although detectors do gather data, this data is not representative for the complete network. This means that if one wants to make the MFD usable in an urban environment, additional data has to be gathered and simple and cheap data acquisition methods have to be created to obtain the massive amount of data needed. However, if a well-defined MFD can be obtained, it can form an invaluable tool for network wide control strategies and network wide evaluation studies.

Based on the literature study, the assessment of the MFD and the results of the simulation, the following implications for using the MFD in practice are set forth:

- Creation of the MFD:
  - MFDs for subnetworks can possibly be constructed from perimeter data, as a strong relation between the perimeter and subnetwork performance has been found;
  - The MFD cannot by simply created using network characteristics, as the stochastic and dynamic nature of traffic seem to have a strong impact on the shape of the MFD.
- Using the MFD:
  - Be careful using unfiltered MFDs, because fast changing demand patterns can create points in the MFD that cannot be steadily sustained and are not representative for the performance of the network;
  - Different types of MFDs should be used to make a decent assessment of the network;
  - The MFD cannot be used to make any statements on improvements in travel times and delays;
  - A MFD created from data obtained from the arterials is not representative for the complete network;
  - MFDs with a lot of scatter should not be used as input for control strategies, to support policy-making decisions, or to evaluate the effect of DTM strategies or changes to the network.
- Traffic control and network design:
  - Adding arterials in the network might have an adverse effect on the network performance;
  - Distributing traffic more evenly over the network, will keep the network in free flow for a larger accumulation range;
  - Changes to signal timings can severely affect the performance and optimal accumulation in a network. However, an optimal timing does seem to exist;
  - Perimeter control should not rely on a single strategy, i.e. only controlling inflow or outflow, but should be a mix of the two and take the performance and accumulation of each part of the network into account.

10 Conclusion and recommendations

From literature, it has not been found to what degree a relationship between the structure of a network (subnetwork size, perimeter length, number of intersections, signal settings) and the shape of the MFD exists. The effect of signal timings on the shape of the MFD of the controlled subnetwork and its perimeter have not been investigated in-depth as well. This thesis has set out to fill these voids and investigate whether or not these relationships exist.
In section 10.1 an answer to these questions will be formulated and the main findings supporting each answer will be presented. From these conclusions, a number of recommendations are drawn in section 10.2, regarding the use of the MFD, further research and improvements to the model.

10.1 Conclusions

In the subsequent two paragraphs, an answer to the research questions is formulated. To this end, first the research question is stated and the answer is formulated. Using the sub questions/ research aspects the answer to the research question is further substantiated.

10.1.1 A relationship between the shape of the MFD and network structure

Research question and answer

"Does a relationship exist between the shape of the macroscopic fundamental diagram and the structure of the underlying network and on which factors does this depend"

The structure of a network is not found to have a strong influence on the shape of the macroscopic fundamental diagram. Differences between MFDs are not caused by topological differences, but by the different characteristics of the links (length, speed, capacity) and intersections (controlled/uncontrolled) of the underlying network. The stochastic and dynamic nature of traffic and uneven distribution of traffic over the network have a strong influence on the shape of the MFD as well. Due to high variations in simulation results, none of the factors could be quantified.

Sub questions/research aspects

Does a general shape for the MFD exist or is this network specific?

- The MFD of uncontrolled/non-signalised parts of the network (subnetworks) exhibit an MFD which has a linear branch in the free flow region with little scatter and a sharp transition to the congested branch. The density range of the free flow branch is relatively short (up to approximately 10 veh/lane-km), whereas the congested branch is quite large (5 – 35 veh/lane-km);
- The MFD of the controlled/signalised parts of the network (arterials) is more rounded and has a somewhat smoother transition from the free flow to the congested state. The maximum performance and optimal density can be up to twice as large as the subnetwork, leading to the

conclusion that controlled parts of the network are better in processing traffic than their uncontrolled counterparts;

- The MFD of the total network is a combination of both the subnetworks and the arterials. As the subnetworks are generally longer, the total network MFD has a stronger resemblance to the subnetwork MFD than the arterial MFD. The balance between these elements in a network is assumed to strongly influence the shape of the MFD. However, the number of networks investigated in this thesis is too limited to fully support this statement.

Which factors influence the shape of the MFD and is it possible to quantify the effect of each of these factors?

- The network length, link capacities, free flow speed, signal settings and distribution of traffic are assumed to influence the shape of the MFD. In this study, none of these factors could be accurately attributed to the shape of the MFD, or be quantified for that matter;
- The stochastic and dynamic nature of traffic and its complex interactions has a strong impact on the shape of the MFD, as it lowers performance, increases accumulation and is an important cause for scatter in the diagram;
- Uneven distribution of traffic over the network reduces the performance and is the main cause for scatter in the diagram;
- A strong, almost linear relation between the accumulation and the deviation of density exists for most networks. No direct relation between the deviation of density and the performance seems to exist;
- MFDs derived from subnetworks and perimeters within the same network show a strong resemblance to each other and can be scaled according to the length of that network;
- The amount of traffic loaded onto the network has a strong influence on the shape of the MFD and can create points in the MFD that cannot be steadily sustained;
- Accumulation and network density are highly sensitive to differences in the distribution of traffic over the network, making the optimal accumulation hard to predict;
- The point in which the network is in gridlock seems to be close to the average critical link density. This implies that if the critical link density would change, the maximum accumulation for the network would change as well.

How is the MFD affected by changes in the network structure?

- The structure of the network, i.e. the shape and size of subnetworks and placement of arterials, does not have a strong influence on the shape of the MFD;
- Differences between MFDs of different networks are not because of certain topological differences between the networks, but are caused by differences in the underlying links of the network and the amount and location of bottlenecks in the network;
- Adding arterials in the network can have an adverse effect on the network performance, as traffic is drawn to central arterials, causing an uneven distribution of traffic and lowering demand;
- Networks with the same structure can have differently shaped MFDs;

10.1.2 A relationship between the shape of the MFD and signal settings

Research question and answer

"Is the macroscopic fundamental diagram of the subnetwork and its perimeter affected by different signal settings and does this affect its applicability for control strategies"

A strong relation between the shape of the macroscopic fundamental diagram of a subnetwork and its perimeter exists, in which both react in the same way to changes in traffic demand and signal settings. Changes in the signal settings can have a strong impact on the shape of the macroscopic fundamental diagram of both the subnetwork and its perimeter. Especially the optimal accumulation is highly sensitive to these changes. Due to this sensitiveness to signal changes, the macroscopic fundamental diagram is currently difficult to use as an input for control strategies.

Sub questions/research aspects

Does a relationship between the MFD of the subnetwork and its perimeter exist?

- The performance ratio, i.e. the ratio between the performance of a subnetwork and its perimeter, is highly consistent under different signal settings, as long as all traffic on the intersection can be served and spill-back is avoided;
- As subnetwork and perimeter react in the same manner to changes in the signal timings, an optimal timing should exist, in which the performance for both the subnetwork and perimeter is maximised. Finding this optimum can greatly improve the network output.

How are the MFD of the subnetwork and its perimeter affected by changes in the signal settings?

- Due to changes to the signal settings, the performance of the network can be improved, but only because more traffic can be served. If little to no congestion in the network is present, the MFD will not change. Only when the amount of traffic increases, improvements in the performance can be witnessed;
- Optimised signal settings do improve travel times and as a consequence the optimal accumulation in the perimeter decreases. This effect can be found even when the amount of traffic remains the same. This effect is amplified as the amount of traffic that can be processed by the network can be higher, which leads to higher traffic flows. These differences can be substantial and differences in the optimal accumulation at different signal settings of up to 75% have been found;
- The optimal accumulation of the perimeter reacts much stronger to changes in the signal settings than the subnetwork. This is because of the smaller lane length of the perimeter, but also because the intersection density is much higher on the perimeter;
- Certain signal settings have been found to be able to sustain a certain performance over a very large accumulation range;
- It is found that an optimal range for the timing of signals controlling the flow between subnetwork and perimeter seems to exist, in which the performance of both systems is maximised.

Can the MFD be used as an input for control strategies?

- As the shape of the MFD is highly sensitive to changes in the signal settings, the MFD is difficult to use for control strategies like perimeter control. Perimeter control aims to adapt the signal timings in order to maintain the optimal accumulation. However, if the signal timings are changed, the optimal accumulation changes, changing the objective function of the control algorithm itself. Only with a-priori knowledge of the effect of these changes, this is a viable strategy;
- For a number of networks it is found that controlling the outflow results in a higher performance and network output. From this it is concluded that the traffic situation can be improved by adapting the signal timings.
- As improvements in the performance and network output are achieved by fixing the flow to a certain demand, it can be concluded that adapting the signal timings to match the traffic demand does not maximise the performance.
- A control strategy should not focus on controlling either the inflow to, or outflow from a subnetwork. The best strategy is very specific for each situation and should depend on both the maximum performance and optimal accumulation of the subnetwork and perimeter.

10.2 Recommendations

10.2.1 Recommendations for using the MFD

The recommendations regarding the use of the MFD have been discussed at length in section 9.4. Below a short summary of these recommendations is given.

- Be careful using unfiltered MFDs, as fast changing demand patterns can create points in the MFD that cannot be steadily sustained and are not representative for the performance of the network;
- Different types of MFDs should be used to make a decent assessment of the network;
- The MFD cannot be used to make any statements on improvements in travel times and delays, as it only conveys information on the quality of the network;
- A MFD created from data obtained from the arterials is not representative for the complete network;
- MFDs with a lot of scatter should not be used as input for control strategies, to support policy-making decisions, or to evaluate the effect of DTM strategies or changes to the network.

10.2.2 Further research

Quantification of the impact of stochastic network design and dynamic traffic behaviour on the shape of the MFD

In this thesis, MFDs for a couple of networks and its subnetwork have been created. However, it is found that due to stochastic factors in network design and traffic behaviour, differently shaped MFDs are obtained for networks with the same layout. The implication of this is that the shape of the MFD is not invariant to the OD-matrix. In order to investigate how stochastic network design and dynamic traffic behaviour influence the shape of the MFD, it is advised to run a large amount

of simulations (possibly using the framework developed in this thesis), in order to obtain statistically significant results, to quantify the effect of these stochastic and dynamic components. Only when this insight has been gained, the question can be answered whether or not a well-defined MFD could be obtained for urban networks.

Investigating the shape of MFDs for highly heterogeneous networks

The networks investigated in this thesis are all still fairly homogeneous. The speed on all links has been kept equal and only two types of roads have been used. Apart from that, the subnetworks are all very simply sized. Investigating how the shape of the MFD changes if the heterogeneity is increased by using different speeds throughout the network and adding freeways can give more insight if it is likely that a well-defined MFD can be obtained for highly heterogeneous networks. The network creation model developed within this thesis, is capable of creating such networks.

Developing methods to obtain well-defined MFDs

Heterogeneously loaded networks are found to produce highly scattered MFDs, due to uneven distribution of traffic over the network. Although MFDs can be obtained for these networks, they cannot be properly used as input for control strategies, to evaluate the effects of changes to the network and to support policy-making decisions. Only well-defined MFDs can be properly used to accurately execute these tasks. As such, further research should focus on improving the shape of the MFD by removing scatter and faulty traffic states.

Obtaining MFDs for existing urban networks

The amount of MFDs obtained from data of existing urban networks is very scarce. Especially MFDs for heterogeneously loaded networks are almost non-existent. In order to quantify the severity of scatter in MFDs of urban networks, these MFDs should be obtained first.
It should be pointed out, that the shape of the MFD of the total network is strongly influenced by the MFDs of its subnetworks. Only using arterial data is therefore not representative for the network. As such, additional methods should be created to obtain this subnetwork data at a relatively low cost.

Investigating the short-term impact of changing signal timings

It has been found that changing the signal timings can have a strong impact on the optimal accumulation in a network. The implications of these results is that the MFD is hard to use as input for perimeter control strategies. However, the signal timings were fixed to a certain flow for the complete simulation. Further research should be done, in order to investigate how fast and under which conditions the optimal accumulation changes, after signal settings have been changed (during the simulation). The time needed for these changes to occur, can then be used to form the upper bound for the control horizon.

Investigating the applicability of MFDs in large-scale national models

Currently, the fundamental diagram is the basis for a lot of macroscopic simulation models, as it provides a simplified relation between flow and density. Drawing the parallel between the fundamental diagram and the MFD, it should be interesting to investigate if it is possible to use the MFD in large-scale national models. In these models the different cities can each be represented by

their own MFD (abstracted from initial simulations for the complete city). The output of each city to the national network, could then be simply calculated by the number of vehicles in it. As a result, each city can be replaced by only a couple of links, governed by the MFD, leading to massive improvements in calculation time. Especially for these large-scale applications the MFD should be highly suitable, but has not been investigated yet.

10.2.3 Suggestions for model improvements

Connecting the model to a macroscopic simulation model

Within this thesis a model has been created, capable of creating networks for the dynamic traffic simulation program VISSIM. However, this program also generates a link-node based network and as such it should be capable to connect the generated networks to a macroscopic simulation model. This adds to the flexibility of the created model. Apart from that, in retrospect it is assumed that roughly the same results could have been obtained using a macroscopic model, which in turn reduces the amount of calculation time significantly. This also makes it possible to compare to what degree the MFDs of a macroscopic and a microscopic simulation model differ from each other.

Adding demand-adaptive signals to the model

The model is capable of communicating with VISSIM and a module to control the signals and read traffic data online is already available. Coupling this data and adding an algorithm to change these traffic states can results in demand-adaptive traffic signals, which should be better tailored to the actual traffic demand at intersections. This also is more in line with reality than the current fixed time controllers.

Adding diagonal links

Currently, the networks can only be created in a grid-like structure. Although this works for fictive networks, this does make it difficult to recreate existing networks. Adding diagonal links can strongly improve the flexibility of the model.

Improving OD-matrix allocation

Currently the number of trips between origins and destinations are randomly assigned. Adding the functionality that a fixed number of trips between certain parts of the network have to be made, makes the model more suitable to investigate the effect of different OD-patterns, or to obtain a better match with an existing network.

Add compensations for other streams on intersections

Currently all the intersections are only designed for cars. No turns are added for cyclists, pedestrians or public transport. Although these streams could physically be incorporated in the model, a simple improvement can be made by prolonging the red time of the other turns in some way, as to compensate for these additional traffic streams.

Adding a warm-up period

The current model does not make use of a warm-up period. As such, the traffic is loaded onto the network very fast and as such 'rush'-effects arise. By adding a warm-up period, this effect can be mitigated and more accurate results/MFDs can be obtained.

Bibliography

Ardekani, S., Herman, R. (1987), *"Urban network-wide traffic variables and their relations"*, Transportation Science 21 (1), Pages 1–16. doi: 10.1287/trsc.21.1.1

Bliemer, M.C.J., Versteegt, H.H., Bovy, P.H.L. (2004), *"INDY: A new analytical multiclass dynamic traffic assignment model",* Prepared for the 83rd Annual meeting of the Transportation Research Board 2004

Boyacı, B., Geroliminis, N. (2010), *"Exploring the Effect of Variability of Urban Systems Characteristics in the Network Capacity"*, Proceedings of the 90th Annual Meeting of the Transportation Research Board, Washington D.C.

Buisson, C., Ladier, C. (2009), *"Exploring the Impact of the Homogeneity of Traffic Measurements on the Existence of Macroscopic Fundamental Diagrams",* Transportation Research Records, Journal of the Transportation Research Board. doi: 10.3141/2124-12

California Department of Transportation (2011), *"What is a traffic model",* [online] Available at: http://www.dot.ca.gov/hq/tsip/microsim/whatis_modeling.html [Accessed 15 June 2011]

Cassidy, M.J., Jang, K., Daganzo, C.F. (2011), *"Macroscopic Fundamental Diagrams for Freeway Networks: Theory and Observation",* Proceedings of the 90th Annual Meeting of the Transportation Research Board

CROW (2002), *"Handboek Wegontwerp wegen buiten de bebouwde kom: Gebiedsontsluitings-wegen",* Kenniscentrum voor verkeer, vervoer en infrastructuur, Publicatie 164c

Courbon, T., Leclercq, L. (2011), *"Cross-Comparison of Macroscopic Fundamental Diagram Estimation Methods",* Procedia Social and Behavioral Sciences 20, Pages 417-426. doi: 10.1016/j.sbspro.2011.08.048

Daganzo, C.F. (1997), *"Fundamentals of Transportation and Traffic Operations"*

Daganzo, C.F. (2005a), *"Improving city mobility through gridlock control: an approach and some ideas"*, Technical Report UCB-ITS-VWP-2005-1, U C Berkeley Center for Future Urban Transport, Inst. Trans. Studies, University of California, Berkeley, CA

Daganzo, C.F. (2005b), *"A variational formulation of kinematic waves: Basic theory and complex boundary conditions"*, Transportation Research Part B 39(2), Pages 187-196. doi: 10.1016/j.trb.2004.04.003

Daganzo, C.F. (2005c), *"A variational formulation of kinematic waves: Solution methods"*, Transportation Research Part B 39 (2), Pages 934-950. doi: 10.1016/j.trb.2004.04.003

Daganzo, C.F. (2007), *"Urban Gridlock: Macroscopic modeling and mitigation approaches",* Transportation Research Part B 41 (1), Pages 49-62. doi: 10.1016/j.trb.2006.03.001

Daganzo, C.F., Geroliminis, N. (2008), *"An analytical approximation for the macroscopic fundamental diagram of urban traffic",* Transportation Research Part B 42, Pages 771-781. doi: 10.1016/j.trb.2008.06.008

Daganzo, C.F., Gayah, V.V., Gonzales, E.J. (2011), *"Macroscopic relations of urban traffic variables: Bifurcations, multivaluedness and instability",* Transportation Research Part B 45, Pages 278-288. doi: 10.1016/j.trb.2010.06.006

Drake, J.S., Schöfer, J.L., May, A. (1967), *"A statistical Analysis of Speed Density Hypotheses",* Proceedings of the Third International Symposium on the Theory of Traffic Flow, Pages 112-117

Edie, L.C. (1963), *"Discussion of traffic stream measurements and definitions",* Proceedings of the 2nd International Symposium on the Theory of Traffic Flow, Pages 139-154

Gayah, V.V., Daganzo, C.F. (2011), *"Exploring the Effect of Turning Maneuvers and Route Choice on a Simple Network",* Transportation Research Record: Journal of the Transportation Research Board, no. 2249, Pages 15-19. doi: 10.3141/2249-03

Gayah, V.V., Daganzo, C.F. (2011), *"Clockwise hysteresis loops in the Macroscopic Fundamental Diagram: An effect of network instability",* Transportation Research Part B 45, Pages 643-655. doi: 10.1016/j.sbspro.2011.04.515

Geroliminis, N. and C.F. Daganzo (2007), *"Macroscopic modeling of traffic in cities",* Transportation Research Board 86th Annual Meeting, # 07-0413, Washington DC.

Geroliminis, N., Daganzo, C.F. (2008), *"Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings",* Transportation Research Part B 42, Pages 759-770. doi: 10.1016/j.trb.2008.02.002

Geroliminis, N., Sun, J. (2011a), *"Properties of a well-defined macroscopic fundamental diagram for urban traffic",* Transportation Research Part B 45, Pages 605-617. doi: 10.1016/j.trb.2010.11.004

Geroliminis, N., Sun, J. (2011b), *"Hysteresis Phenomena of a Macroscopic Fundamental Diagram in Freeway Networks",* Transportation Research Part A 45, Pages 966-979. doi: 10.1016/j.tra.2011.04.004

Geroliminis, N., Haddad, J., Ramezani, M. (2012), *"Optimal Perimeter Control for Two Urban Regions with Macroscopic Fundamental Diagrams: A Model Predictive Approach",* Transactions on Intelligent Transportation Systems and Intelligent Transportation Systems Magazine, Manuscript T-ITS-12-02-0055

Godfrey, J.W. (1969), *"The mechanism of a road network",* Traffic Engineering and Control, Volume 11, Pages 323-327

Greenshields, B.D. (1935), *"A Study of Traffic Capacity"*, Freeway Research Board Proceedings, Volume 14, Pages 448-477

Helbing, D. (2009), *"Derivation of a fundamental diagram for urban traffic flow",* The European Physical Journal B, Pages 229-241. doi: 10.1140/epjb/e2009-00093-7

Ji, Y., Geroliminis, N. (2011), *"Exploring Spatial Characteristics of Urban Transportation Networks",* 14th International IEEE Conference on Intelligent Transportation Systems. doi: 10.1109/ITSC.2011.6083062

Ji, Y., Geroliminis, N. (2011), *"Spatial and Temporal Analysis of Congestion in Urban Transportation Networks",* 90th Annual Meeting of the Transportation Research Board

Ji, Y., Daamen, W., Hoogendoorn, S.P., Hoogendoorn-Lanser, S., Qian, X. (2010), *"Investigating the Shape of the Macroscopic Fundamental Diagram Using Simulation Data",* Transportation Research Record: Journal of the Transportation Research Board, no. 2161, Pages 40-48. doi: 10.3141/2161-05

Joos, E. Deputy Director of Zurich Transport Authority (2000), *"Economy and Ecology are no Contradictions",* Ecoplan International [online] Available at: http://www.ecoplan.org/politics/general/zurich.htm [Accessed September 2011]

Keith, S.R. (1971), *"Mechanics, 3rd edition"*, Addison-Wesley, ISBN: 0-201-079392-7

Knoop, V.L., Hoogendoorn, S.P., van Lint, J.W.C. (2011a), *"Data requirements for Traffic Control on a Macroscopic Level",* Proceedings of 2nd International Workshop on Traffic Data Collection & its Standardisation

Knoop, V.L., Hoogendoorn, S.P., van Lint, J.W.C. (2011b), *"Route Advice Based on Subnetwork Accumulations: Control Based on the macroscopic fundamental diagram",* Submission to the 91st Annual Meeting of the Transportation Research Board

Kühne, R.D. (1987), *"Freeway speed distribution and acceleration noise: Calculations from a stochastic continuum theory and comparison with measurements",* in Gartner, N.H., Wilson, N.H.M. (Eds.), Transportation and Traffic Theory, Pages 119-137

Laval, J. (2010), *"The Effect Of Signal Timing And Network Irregularities In The Macroscopic Fundamental Diagram",* Georgia Institute of Technology

Mahmassani, H.S., Peeta, S. (1993), *"Network performance under system optimal and user equilibrium dynamic assignments: implications for advanced traveler information systems",* Transportation Research Record 1408, Pages 83–93

Mahmassani, H.S., Williams, J., Herman, R. (1987), *"Performance of urban traffic networks"*, Proceedings of the 10th International Symposium on Transportation and Traffic Theory, Pages 1-20

Mazloumian, A., Geroliminis, N., Helbing, D. (2010), *"The Spatial Variability of Vehicle Densities as Determinant of Urban Network Capacity",* Philosophical Transactions of the Royal Society A, 368, Pages 4627-4647. doi: 10.1098/rsta.2010.0099

Muller, T.H.J., Hegyi, A., Salomons, M., van Zuylen, H.J. (2011), *"Traffic Management and Control, CT4822-09: Approaches for urban and freeway traffic",* Delft University of Technology, Civil Engineering and Geosciences, Department of Transport & Planning

Olszewski, P., Fan, H.S.L., Tan, Y.W. (1995), *"Area-wide traffic speed-flow model for the Singapore CBD"*, Transportation Research Part A: Policy and Practice 29 (4), Pages 273–281. doi: 10.1016/0965-8564(94)00033-7

PTV Vision (2009), *"VISSIM 5.20: User Manual",* PTV AG, Karlsruhe

Qian, X. (2009), *"Application of Macroscopic Fundamental Diagram to Dynamic Traffic Management",* MSc Thesis Civil Engineering, Delft University of Technology, ITS Edulab, Delft

Schneider, M. (1959), *"Gravity Models and Trip Distribution Theory"*, Papers and Proceedings of the Regional Science Association, Volume 5, Issue 1, Pages 51–56. doi: 10.1111/j.1435-5597.1959.tb01665.x

Strating, M. (2010), *"Coordinated signal control for urban traffic networks by using MFD",* MSc Thesis Civil Engineering, Delft University of Technology, ITS Edulab, Delft

van Nes, R., Bovy, P.H.L. (2008), *"Course CT5802: Advanced transportation modelling and network design",* Delft University of Technology, Faculty of Civil Engineering and Geosciences, Department of Transport & Planning

Wu, X., Liu, H.X., Geroliminis, N. (2011), *"An empirical analysis on the arterial fundamental diagram",* Transportation Research Part B 45, Pages 255-266. doi: 10.1016/j.trb.2010.06.033

Yoshii, T., Yonezawa, Y., Kitamura, R. (2010), *"Evaluation of an Area Metering Control Method Using the Macroscopic Fundamental Diagram",* World Conference on Transport Research (2000), Pages 1-12

Zhang, H.M. (2003), *"Driver memory, traffic viscosity and a viscous vehicular traffic flow model",* Transportation Research Part B 37, Pages 27-41. doi: 10.1016/S0191-2615(01)00043-1

Zhang, L., Garoni, T.M., de Gier, J. (2011), *"A comparative study of Macroscopic Fundamental Diagrams of urban road networks governed by different traffic signal systems",* Transportation Research Part B, Preprint submission, 18 pp.

Appendix A: Network creation model

Introduction

In Figure A.1 below, an overview of the different components of the network creation and updating algorithm is given. In the subsequent chapters a detailed description of each of these steps is given. These steps form the paragraphs of every chapter. Within this paragraphs, every step is broken down further into the different algorithms as created and used in Matlab.
At the start of every step, a summary of the step is given in the introduction. Those who want to get a general understanding on how the algorithm works, should suffice by reading these introductions, or are referred to paragraph 5.3.2.

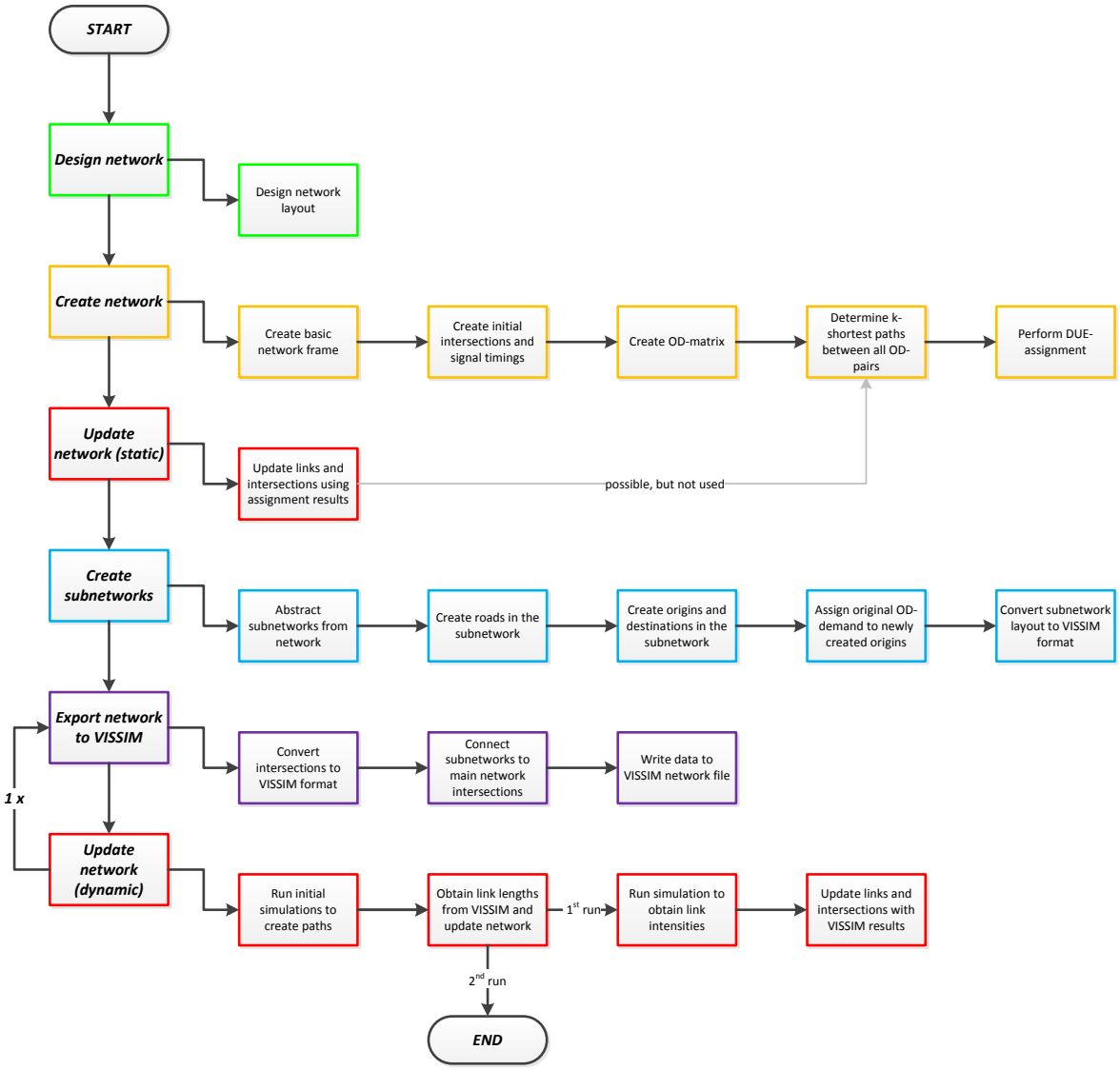


Figure A.1: *General components network creation and updating algorithm*

Step 1: Design network

In the first step a general design of the network is made in Excel. Using predefined values, an abstraction of a network is made. This network can then be imported into Matlab, which transforms it into a network that can be used in VISSIM.
Within this design, a function can be assigned to every intersection in the grid.

Designing the layout of the arterial is done in Excel. Using a predefined grid spacing and intersections spacing, the network can be designed by assigning certain values to every field within the grid. An example of this is shown in Figure A.2.
In creating the network, three different levels are distinguished. The first level are the freeways (shown in orange, values between 3 and 4), the second level is the arterial[24] (values between 1 and 2) and the third level is formed by the subnetworks (empty space).

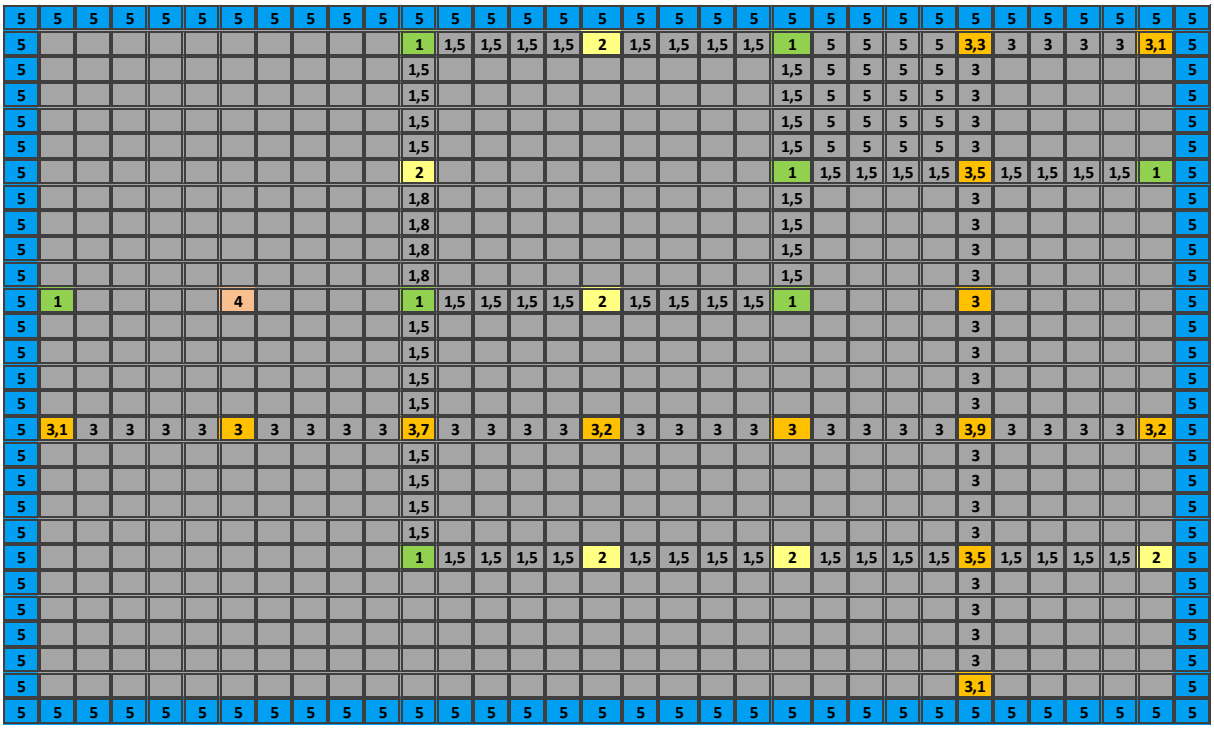


Figure A.2: *Network layout in Excel*

In the example shown above a grid spacing $d_{grid-grid} = 100m$ and an intersection spacing of $d_{int-int} = 500m$ is used. Within the model mainly the intersection spacing is used. The addition of the grid spacing is to (1) get a better indication of the final network and (2) to add links between the intersections indicating the maximum speed.
Intersections can only be placed in fields, in which the x-coordinate and the y-coordinate are equal to the intersection spacing (the coloured fields). In Table A.1 the definition of each of these values and its properties is explained in some more detail.

[24] Within the program code, arterials are referred to as 'arterial'

However, to fully understand how these values work, it is important to understand the way in which this format is processed within Matlab, which is described in step 2.1-a.

Table A.1: *Network design components*

Value	Definition of network component	Properties
1	Arterial intersection	• This intersection can connect to any other intersection; • This intersection can create a connection into any adjacent subnetwork; • This intersection can create an additional connection to the outside of the network (create new zone).
1,x	Arterial link	• x = speed / 10.
2	Subnetwork intersection	• This intersection can connect to arterial intersections; • This intersection can create a connection into any adjacent subnetwork.
3	Freeway link	• Links of the arterial cannot cross, unless specified otherwise.
3,1	End of freeway	• Defines the point where the freeway ends and a zone is to be created.
3,2	Freeway link, subnetwork connection may be made	• Does not hinder the creation of a connection of a subnetwork intersection into the subnetwork next to the freeway.
3,3	Freeway turn	• Point where freeway makes a turn.
3,5	Freeway link, underlying link may cross	• Link of the arterial may cross, but does not connect to the freeway.
3,7	Freeway off-/onramp	• Point where freeway connects with main intersections; • No connection with adjacent subnetwork intersections is made.
3,9	Freeway junction	• Intersection between two freeway links.
4	Subnetwork entry blocker	• Prevents the creation of a subnetwork exit/entry point.
5	Network boundary	• Boundary of the network; • Area in which no subnetwork can be created.
empty	Subnetwork area	• Area in which a subnetwork can be created.

Step 2: Create network

Within the second step, the previously designed network is imported in Matlab. Based on the function assigned to every intersection, that intersection looks in all four directions, if an adjacent intersection is present which it can connect to. With this, the basic nodes and links of the network are created.
In the next step all nodes are converted to intersections and their physical layout, including full signal schemes are created. Once the intersections are built, a fictional OD-matrix is constructed, after which a Deterministic User Equilibrium (DUE) assignment is performed to find the initial link loads, in order to update the network.

Step 2.1: Create basic network frame

Step 2.1-a: Load network design and create basic network (nwc_f01_gridload)

The algorithm starts by importing the network designed within Excel. After that, the network is scaled down to its essence, in which only the intersections are given. In Figure A.3 the scaled down version of the network in Figure A.2 is shown.

5	5	5	5	5	5	5	5	5
5	0	0	1	2	1	3,3	3,1	5
5	0	0	2	0	1	3,5	1	5
5	1	4	1	2	1	3	0	5
5	3,1	3,7	3,7	3,2	3	3,9	3,2	5
5	0	0	1	2	2	3,5	2	5
5	0	0	0	0	0	3,1	0	5
5	5	5	5	5	5	5	5	5

Figure A.3: *Scaled down essence of created network*

After the network has been scaled down, the different intersections are converted to basic nodes. Based on the function assigned to each intersection, every node 'looks around' in all four directions, to see if there is any adjacent node it can connect with (intersections can look through fields set to zero). If such a node is found, a link between the nodes is created.
In case of arterial (1) and subnetwork intersections (2), if an empty space is found, an additional node is created at some distance from the original node ($d_{node,grid-node,art} = 300\ m$) and links between the original and newly created nodes are established. These new nodes will then later function as an exit/entry point, with which the subnetworks can be connected to the arterial.
Another exception is formed by arterial intersections (1), or freeway ends (3,1 and 3,2). When one of these nodes is next to the boundary of the network, also a new node and links are created, which will function as origin-destination zones, also referred to as 'feeders'. In case of a freeway, the node is moved a lot further away from the original node ($d_{node,grid-node,fw} = 1.000\ m$). Reason for this is that when using the network in VISSIM, a sufficient amount of distance is available between the origin zone and the first downstream intersection. In this way, vehicles that origin from a freeway zone and have been loaded onto the network at the rightmost lane of the freeway link, have a reasonable amount of distance to move to the left lane in case they need that exit. The resulting network is shown in Figure A.4.

Step 2.1-b: Remove 'dangling' parts from the network (nwc_f02_Dijkstra strip)

After the network has been converted into basic links and nodes, the integrity of the network is checked. In the example given in Figure A.4, a small part of the network is not connected to the rest of the network. Because trips from and to this part cannot be completed, the inclusion of this

part is not relevant. Using the Dijkstra-algorithm, a path between all origins and destinations is created. After this has been done, the maximum number of destinations that can be reached from any origin is calculated. Any origin connected to less destinations is therefore located in the 'dangling' network and is removed. The resulting network is shown in Figure A.5.

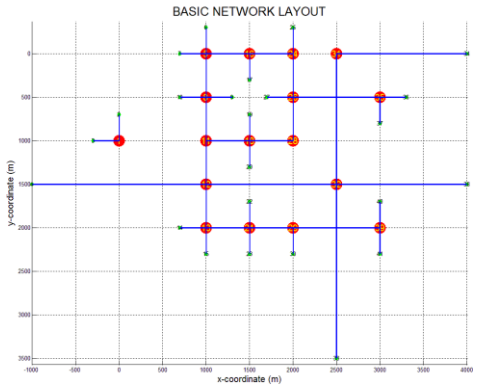


Figure A.4: *Network after importing it in Matlab*

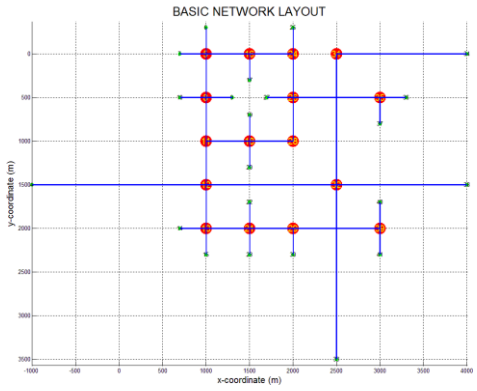


Figure A.5: *Network after removing the 'dangling' network parts*

Step 2.1-c: Turn nodes into intersections (nwc_f03_intersection_creator)

Now that the network has been turned into a fully useable set of nodes and links, it is time to turn the nodes into intersections, which can be used to regulate traffic.

At this stage, the node only knows in which direction it can connect to another intersection. Based on this information, it can be determined which turns are present. So in order to design the intersection, a single lane and an initial load of $q_{init,int,tn} = 100\ veh/h$ is assigned to each turn. Apart from the incoming lanes, also the outgoing lanes are created.

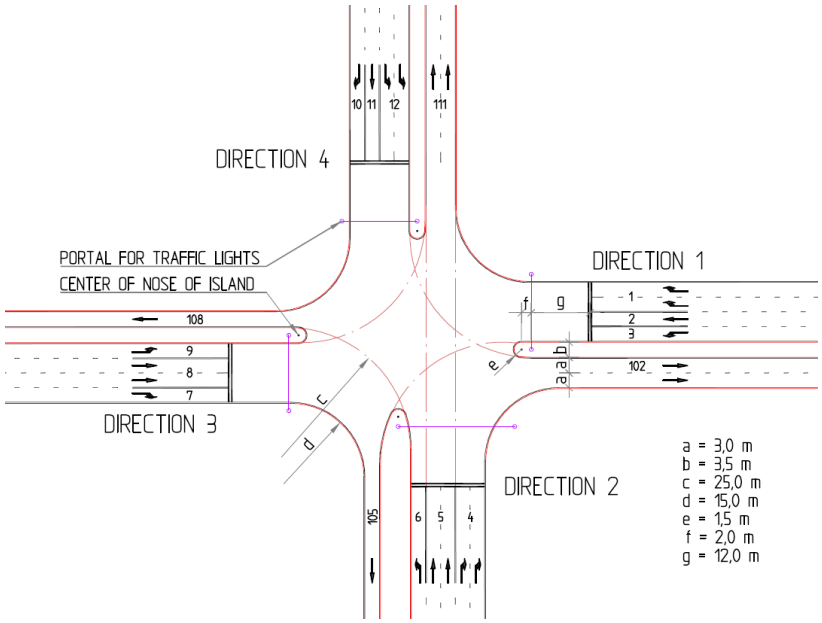


Figure A.6: *General intersection layout, terminology and numbering*[25]

[25] The numbering used for turns 1-12 is the standardised numbering as given in CROW (2002). The outgoing links are normally not numbered; the assigned numbering is solemnly used within this thesis.

In case the intersection is a freeway junction, or an off-/onramp, the through-moving freeway links are removed. Reason for this is that these links should not be included in the design of the intersection, as they do not conflict with the other turns. The other turns are created in the same way as normal intersections.

Step 2.2: Create initial intersections and signal timings

Step 2.2-a: Create physical layout of the intersection (sig_f01_intersection)

In step 2.1-c the turns that have to be incorporated within each intersection have been determined. The first step in creating the intersections and their timings, is to construct a physical layout for each intersection. Within this step, all lanes are created, turns are constructed and the location of traffic lights and stop bars is determined.

An important thing to note is that the intersection is designed only to accommodate normal traffic. Pedestrians, cyclists, and special lanes for public transport have not been added. Another thing to note is that every turn has its own lane(s), meaning that no combined turns are included.

Using the number of lanes defined in the previous step and the design variables below, the complete intersection can be constructed:

- Lane width ($w_{lanes} = 3{,}0m$);
- Minimal bank/island width ($w_{isl} = 3{,}5m$);
- Radius of left turn ($\varphi_{left,int} = 25\text{m}$);
- Radius of right turn ($\varphi_{right,int} = 15\text{m}$);
- Radius of nose of island ($\varphi_{isl} = 1{,}5\text{m}$);
- Distance between island nose and signal ($d_{isl-sig} = 2{,}0m$);
- Distance between signal and stop-line ($d_{sig-stop} = 12{,}0m$).

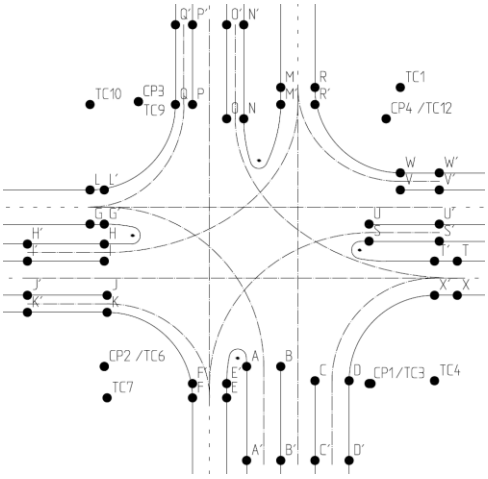


Figure A.7: *Coding of base points*

The first thing that has to be done, is to determine the way that the different turns are positioned within the intersection. The main assumption in this, is that the rightmost lane of the through-going turn is in line with the rightmost lane of the outgoing link. The number of lanes of the outgoing link is set equal to the incoming turn with the highest number of lanes. Based on this assumption, all base points of the intersection can be calculated. In the Matlab-script, each base point has been assigned a specific letter. The corresponding location in the intersection of each of these point is shown in Figure A.7.

The next step is to determine the location of the stop bar of every turn. The location of the stop bar is defined in CROW (2002), at a minimal distance of $d_{sig,low-stop} = 3{,}0\ m$ for lowly placed traffic lights and $d_{sig,high-stop} = 12{,}0\ m$ for highly placed traffic lights. The type of traffic lights to be used depends on the layout of the incoming turns. As a simplification, all distances within the model are set to $d_{sig,high-stop} = 12{,}0\ m.$ However, reducing this distance will decrease clearance times.

The problem with this however, is that there is no specific relation between the layout of the intersection and the location at which the signals are placed. As such, the actual location of the stop bar in relation to the intersection is somewhat arbitrary. Within the model, it is assumed that the support structure of the traffic lights (portal or pole) are placed within the central island, at a distance of $d_{port-isl} = 2{,}0m$ behind the centre of the nose point, as illustrated in Figure A.6.
The actual coordinates of this centre can be calculated using the intersection point of the radii of the incoming and outgoing left turn.

Now that the location of each stop bar is determined, it is possible to create the full layout of the intersection, as is shown in Figure A.8.

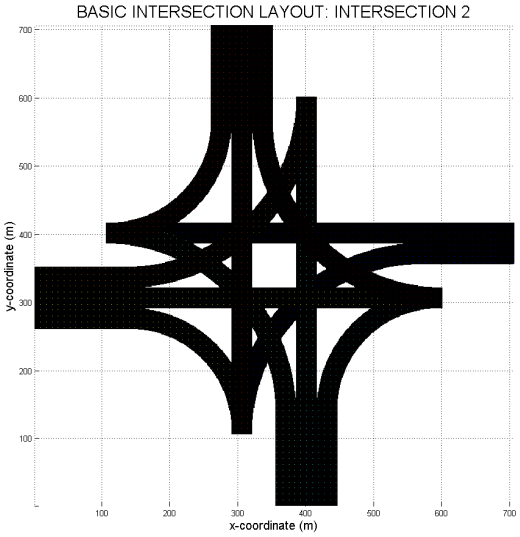


Figure A.8: *Example of automated design of an intersection*

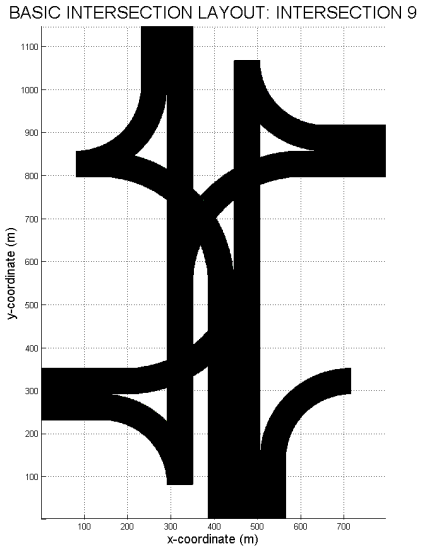


Figure A.9: *Example of automated design of an off-/onramp*

In the case that the intersection is a freeway junction or an off-/onramp, the width of the central bank is set equal the total width of all lanes plus a safety distance $d_{safe} = 10{,}0m$ on either side, creating enough space, to later add the freeway underneath. The result is shown in Figure A.9.

Step 2.2-b: Determine conflict areas and clearance times (sig_f02_conflict_generator)
In the previous step, the physical layout of each intersection has been calculated. Using this layout, the conflicting areas between the different turns can easily be determined. This is done by mapping every turn on a grid and then summing the values of these grids. At the point where turns conflict with each other, a different value is found. The results of this are given in Figure A.10.

Using the coordinates of the conflict area and the stop bar, the clearance times can be calculated. The clearance time is the time it takes for the conflict area to clear, before the next stream can enter this area.
When the distance between the conflict area and the first turn is very short and the distance between the conflict area and the second turn is very long, it can happen that traffic of the second turn has already been given the green light, before the last vehicle from the first turn has left the

area. It then goes to say, that when these clearance times are calculated very precise, the total throughput of an intersection can be maximised. Poorly calculated clearance times, on the other hand can lead to under-utilization of the intersection.

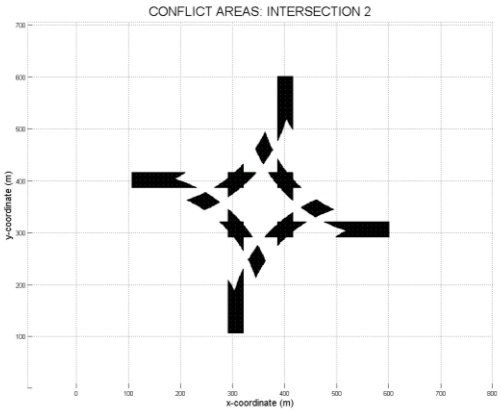


Figure A.10: *Conflict areas*

The method used to calculate the clearance times and determine the conflict groups, control structures and signal schemes has been adopted from Muller et al. (2011).

The clearance time $t_{clear}(g,h)$ is calculated as the time between the last vehicle from the first turn g leaving the conflict area $t_{exit}(g)$ and the first vehicle entering the conflict area $t_{entry}(h)$ from the second turn h, which is shown in equation (A.1) and Figure A.11.

$$t_{clear}(g,h) = t_{exit}(g) - t_{entry}(h) \tag{A.1}$$

When calculating clearance times, the first action is to calculate the critical conflict distance

$$d_{crit} = \frac{u_{max,int}^2}{2(acc-dec)}. \tag{A.2}$$

Using the critical conflict distance, the entry and exit time can now be calculated as

$$t_{entry}(h) = t_{react} + \frac{d_{entry}(h)}{u_{max,int}} + \frac{u_{max,int}}{2(acc-dec)} \quad \text{for} \quad d_{entry} > d_{crit}, \tag{A.3}$$

$$t_{entry}(h) = t_{react} + \sqrt{\frac{2d_{entry}(h)}{acc-dec}} \quad \text{for} \quad d_{entry} < d_{crit}, \tag{A.4}$$

$$t_{exit}(g) = \frac{d_{exit}(g)}{u_{exit}(g)}. \tag{A.5}$$

In which the entry distance is the Euclidian distance between the stop bar of the second turn and the start of the conflict zone. The exit distance is the Euclidian distance between the stop bar of the first turn and the end of the conflict zone cfz, increased by the average vehicle length

$$d_{entry}(g) = min_{XY \in cfz(g,h)}(|XY_{stop(g)}, XY_{cfz(g,h)}|) \quad \text{(A.6)}$$

$$d_{exit}(h) = max_{XY \in cfz(g,h)}(|XY_{stop(h)}, XY_{cfz(g,h)}|) + l_{veh} \quad \text{(A.7)}$$

The maximal driving speed for any controlled intersection is $u_{max,int} = 50\ km/h = 13{,}9\ m/s$. This is in accordance with the maximal speed limit of most urban intersections. The other default values used are derived from Muller et al. (2011) and are:

$acc = 2{,}0\ m/s^2$, $dec = -3{,}0\ m/s^2$,

$t_{react} = 1\ s$, $v_{exit} = 40\ km/h = 11{,}1\ m/s$,

$L_{veh} = 12\ m$.

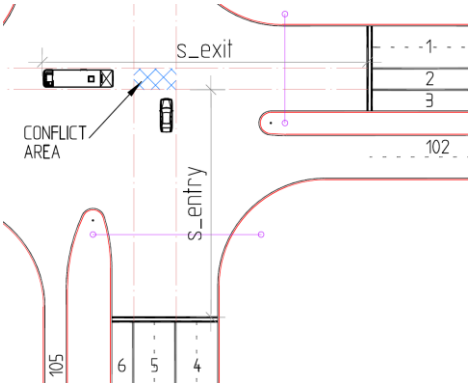


Figure A.11: *Clearance distances*

Step 2.2-c: Determine conflict groups (sig_f03_01_conflict_groups)

When the various conflicts have been found, it becomes possible to construct the different conflict groups. A conflict group is a set of signals which conflict with one another and cannot be operated at the same moment. These conflict groups form the basis of the final control structure of the intersection. As can be seen in Figure A.11, turns 2 and 5 conflict with each other. Turns 1 and 4 on the other hand do not conflict, meaning they can be given green at the same time.

The first step is to construct the 2-phase conflict groups, after which the 3-phase and 4-phase conflict groups are determined. This is done by finding the conflicts having mutual conflicts, for example {2,5} – {2,9} – {5,9} = {2,5,9}. For the intersection used within this example (Figure A.6) these conflicts are:

Table A.2: *2-phase conflict groups*

{1,5}	{2,5}	{3,5}	{4,8}	{5,8}	{6,8}	{7,11}	{8,11}	{9,11}
{1,9}	{2,6}	{3,6}	{4,12}	{5,9}	{6,9}		{8,12}	{9,12}
	{2,9}	{3,7}		{5,12}	{6,10}			
	{2,10}	{3,8}			{6,11}			
	{2,11}	{3,11}						
	{2,12}	{3,12}						

Table A.3: *3-phase conflict groups*

{1,5,9}	{2,5,9}	{3,5,8}	{4,8,12}	{5,8,12}	{6,8,11}
	{2,5,12}	{3,5,12}		{5,9,12}	{6,9,11}
	{2,6,9}	{3,6,8}			
	{2,6,10}	{3,6,11}			
	{2,6,11}	{3,7,11}			
	{2,9,11}	{3,8,11}			
	{2,9,12}	{3,8,12}			

Table A.4: *4-phase conflict groups*

{2,5,9,12}	{2,6,9,11}	{3,5,8,12}	{3,6,8,11}

The careful reader will notice that after obtaining the 4-phase conflict groups, not all 3-phase conflicts are accounted for, as they cannot be 'fitted' into a 4-phase conflict, such as {1,5,9}. The 4-phase conflict groups are then complemented with the remaining conflict groups from the 2nd and 3rd phase, creating a list with all conflict groups in the intersection, as shown in Table A.5.

Table A.5: *Final conflict groups*

{1,5,9}	{2,5,9,12}	{3,5,8,12}	{4,8,12}
	{2,6,9,11}	{3,6,8,11}	
	{2,6,10}	{3,7,11}	

Step 2.2-d: Create control structures (sig_f03_02_control_structures)

Using the final conflict groups, the different control structures can be created. A control structure is the sequence in which each conflict group is given green. These structures are any random sequence of the final conflict groups created in the previous step and are made by 'aligning' the different conflict group, as is illustrated in Table A.6. Every row in this sequence is called a 'stage'. The number of possible sequences in our example is 24. Based on the turns within the intersection, this can run up to 54 sequences[26].

Table A.6: *Example of a control sequence*

2	2	8	8	***1***	2	***7***	8
12	6	12	6	***1***	6	***7***	12
5	11	5	11	5	***10***	11	***4***
9	9	3	3	9	***10***	3	***4***

As can be seen in Table A.6, 3-phase conflict groups have certain signals, which are used within multiple stages. This is because only three stages are needed to give every turn within that conflict group green once. The 'open' stage for these groups can be filled with any signal of that group, as long as this does not conflict with the other signals.

Step 2.2-e: Create signal schemes (sig_f04_cycle_times)

After all the control structures have been created, the actual signal scheme can be made. Within a signal scheme the timing of the state change (green-yellow-red) of every signal is laid down.
The most important parameters for these schemes are:

- Traffic intensities (*step 2.1-c)*;
- Clearance times (*step 2.2-b)*;
- Control sequences (*step 2.2-d)*;
- Basic saturation flow ($q_{sat} = 1.800\, veh/h$);
- Minimum yellow time ($t_{yellow,min} = 3{,}0s$);
- Minimum green time ($t_{green,min} = 6{,}0s$);
- Maximum cycle time ($t_{cyc,max} = 120{,}0s$).

[26] In order to improve computation time, conflict groups and control structures are pre-calculated for any intersection configuration and stored in a separate structure, using its binary code as indicator. In our example, the configuration is 111111111111 = 4095.

As a first step, the saturation flow is adjusted to take specific geometric conditions of the intersection into account. Multiple factors influence the saturation flow, but within this thesis only adjustment factors for the lane width and the turn type are used. Again the values have been adopted from Muller et al. (2011).

$$\beta_{adj} = f_w * f_{RT} * f_{LT}, \tag{A.8}$$

In which:

$f_w = 0{,}94$ adjustment factor for lane width;

$f_{RT,1} = 0{,}85$ adjustment factor for right turn with single lane;

$f_{RT,2} = 0{,}75$ adjustment factor for right turn with multiple lanes;

$f_{LT,1} = 0{,}95$ adjustment factor for left turn with single lane;

$f_{LT,2} = 0{,}92$ adjustment factor for left turn with multiple lanes;

After determining the saturation flow adjustment factor, the cycle time t_{cyc} (the time needed to serve every turn) is calculated by:

$$t_{cyc} = \frac{3600}{n_{cyc}}, \tag{A.9}$$

$$n_{cyc} = \frac{t_{rest}}{t_{lost}}, \tag{A.10}$$

$$t_{rest} = 3600 - t_{green}, \tag{A.11}$$

$$t_{green} = \sum_{g \in cf} q_{int,tn}(g) \frac{3600}{q_{sat}\beta_{adj}(g)}, \tag{A.12}$$

$$t_{lost} = \sum_{g,h \in cf} \left(t_{yellow}(g) + t_{clear}(g,h) \right), \tag{A.13}$$

in which g is the first turn and h is the subsequent turn within the conflict group cf.

To clarify the above, Table A.8 illustrates this for the first conflict group of our example {2,12,5,9}. Although the intensities have been set to 100 in step 2.1-c, the intensities shown in Table A.7 are used, as they are better suited to illustrate the example.

Table A.7: *Traffic intensities and saturation flow per turn*

Turn	1	2	3	4	5	6	7	8	9	10	11	12
Intensity	200	400	300	200	400	300	200	400	300	200	400	300
Sat. adj. factor	0,799	0,94	0,893	0,799	0,94	0,893	0,799	0,94	0,893	0,799	0,94	0,893
Saturation flow	1438	1692	1607	1438	1692	1607	1438	1692	1607	1438	1692	1607

Table A.8: *Cycle time calculation*

g	**h**	**t_{yellow} (s)**	**t_{clear} (s)**	**q (veh/h)**	**t_{green} (s/h)**
2	12	3,0	-0,36	400	851
12	5	3,0	-0,72	300	672
5	9	3,0	1,15	400	851
9	2	3,0	-0,72	300	672
Total	**-**	**12,0**	**-0,65**	**1400**	**3046**

Using the above formulas, the initial cycle time for conflict group 1 is 14,6 seconds.

$$t_{lost} = 12{,}0 - 0{,}65 = 11{,}35\ s$$
$$t_{green} = 3046\ s/h$$
$$t_{rest} = 3600 - 3046 = 554\ s$$
$$n_{cyc} = \frac{554}{11{,}35} = 48{,}8\ \ cyc/h$$
$$t_{cyc} = \frac{3600}{48{,}8} = 73{,}8\ s$$

This calculation is made for each conflict group within the control structure, after which the initial minimum cycle time for the total control structure is set equal to the longest cycle time of any conflict group. In our example, this is conflict group 2, with a cycle time of also 73,8 seconds.
The next step is to create the timing for every signal. However, the cycle time calculated earlier only incorporated signals 2, 5, 9 and 12. The other signals will now have to be incorporated as well. The reason why the initial cycle time has been determined, is to speed up the computation time. In no situation, the cycle time can become less than 73,8 seconds, making this the point of departure for the calculation.
Based on the number of cycles per hour, the number of vehicles that have to be served per cycle and the green time needed per turn can be calculated. When this is done, at $t = 0$ the first stage is initialised, after which the subsequent signals from the next stages are appended, taking the green time, yellow time and clearance times into account. The start time of the subsequent signal h, depending on the previous signal g can be written as

$$t_{start}(h, g) = t_{start}(g) + t_{green}(g) + t_{yellow}(g) + t_{clear}(g, h). \quad \text{(A.14)}$$

As h can depend on multiple other signals, the maximum start time should be used

$$t_{start}(h) = max\big(t_{start}(g, h)\big) \text{ for all } g \text{ previous to } h. \quad \text{(A.15)}$$

For instance $t_{start,green}(2) = 19{,}5\ s$ and $t_{yellow}(2) = 3{,}0\ s$. The signals following 2 are 6 and 12 and

$$t_{start,green}(6{,}2) = t_{start,green}(6) + t_{green}(6) + t_{yellow}(6) + t_{clear}(2{,}6) = 0 + 19{,}5 + 3 + 1{,}15 = 23{,}65\ s,$$
$$t_{start,green}(12{,}2) = 0 + 19{,}5 + 3 - 0{,}37 = 22{,}13\ s.$$

However, the start of signal 6 and 12 also depends on signal 8, with

$$t_{clear}(8{,}6) = -0{,}41\ s \text{ and } t_{start,green}(6{,}8) = 22{,}09\ s,$$
$$t_{clear}(8{,}12) = 1{,}21\ s \text{ and } t_{start,green}(12{,}8) = 23{,}71\ s.$$

So

$$t_{start,green}(6) = max(23{,}65; 22{,}13) = 22{,}1\ s,$$
$$t_{start,green}(12) = max(22{,}13; 23{,}71) = 22{,}1\ s.$$

As it is not necessarily true that all signals of the first stage simultaneously start at $t = 0$ (for instance, in our case the optimal starting time for signals 1 and 7 is at 1,6 s), the process of adding the subsequent stages is repeated multiple times. After all stages have been added again, the cycle time is calculated and compared with the cycle time achieved in the previous iteration. This loop is ended when the current and the previous cycle time match.

After the signal timings have been obtained, the extended green time is calculated. Extended green is given when $t_{green}(g)$ has expired, but no conflicts with the turn g are present. A good example are signals 4 and 5. If $t_{green}(4) = 11{,}5s$ and $t_{green}(5) = 19{,}5s$, signal 4 can be set to green for the same amount as signal 5, as signal 4 does not have any conflicts as long as signal 5 is active (see Figure A.12).

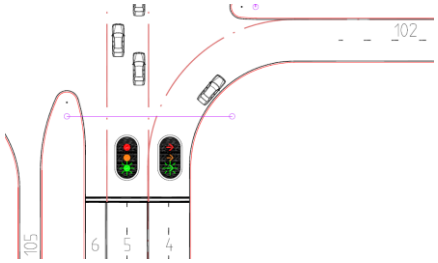


Figure A.12: *Extended green time as a consequence of a 'covered' turn*

After performing the above operations, the cycle time of our example structure is 75,3 seconds. As the cycle time has now changed, the number of cycles during an hour has changed from $\frac{3600}{73{,}8} = 48{,}8\ cyc/h \rightarrow \frac{3600}{75{,}3} = 47{,}8\ cyc/h$. This means that the number of vehicles arriving per cycle has now increased and $t_{green}(g)$ should be longer. $t_{lost}n_{cyc}$ on the other hand has decreased from $11{,}35 \cdot 48{,}8 = 554\ s/h \rightarrow 11{,}35 \cdot 47{,}8 = 542{,}5\ s/h$, meaning an increase in the effectiveness of the intersection. Using the obtained cycle time, the green time per turn is recalculated and the process described earlier is repeated. This iteration then results in a cycle time of 76,5 seconds. The loop is ended when the difference in the current and the previous iteration is smaller than 0,1 second, or the maximum cycle time is reached. In the latter case, the given control structure is discarded as it would take too long to serve all turns. Running the program for our example structure results in a cycle time of 82,7 seconds.

The resulting signal schemes have been validated by comparing the results, with the results generated by VRIGen (see Muller et al., 2011); a tool which also has been designed to generate signal schemes. Results of both methods matched within 0,2 seconds, thus validating this algorithm.

Step 2.2-f: Calculating lane lengths (sig_f05_lane_length)

When the signal schemes have been created, the scheme with the shortest cycle time is chosen (which by no incident is the one described above). Using this scheme, the number of vehicles arriving during the red phase of every signal sig in intersection int is calculated as

$$n_{veh,arr,red}(sig, int) = \frac{q(sig)\left(t_{cyc}(int) - t_{green}(sig)\right)}{3600}. \qquad (A.16)$$

An important factor in determining the lane length of each turn, is to avoid spill-back as a consequence of a lane getting blocked. An example of this is given in Figure A.13.

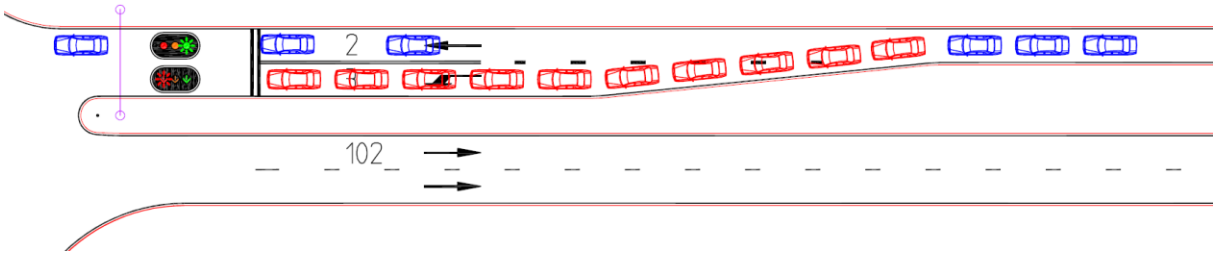


Figure A.13: *Spill-back caused by blocking back*

To avoid this problem, any lane has to be sufficiently long, to store the appropriate amount of vehicles. As traffic is probabilistic of nature, simply taking the average number of vehicles, as in (A.16) will not suffice, as in some cases more vehicles arrive and block other lanes.
Using a Poisson distribution, the maximum number of vehicles arriving during a red phase, with a probability of 95% is calculated by

$$P\left(x \leq N(sig); n_{veh,red}(sig)\right) = \sum_{k=0}^{N} \frac{e^{-n} n^k}{k!} \geq 0{,}95. \quad \text{(A.17)}$$

With the resulting number of vehicles, the length of each lane is calculated as

$$l_{lanes}(sig) = max\left(\frac{N(sig) l_{veh}}{n_{lanes}(sig)}, l_{lanes,min}\right). \quad \text{(A.18)}$$

The average vehicle length is set to $l_{veh} = 7{,}5\ m$, the minimum lane length is taken from CROW (2002) and is $l_{lanes,min,left} = l_{lanes,min,right} = 30{,}0m$ for a left and right turn and$l_{lanes,min,thr} = 40{,}0m$ for a through-going turn.
It should be noted however that this solution to the blocking-back problem only works for the left and right turn. Imagine the case in which the lane of the right turn has to accommodate 3 vehicles and the through-going turn 17. As the right turn lane is only 30 meters, vehicles that want to turn right and are located on the 5th or 6th position in the queue waiting for the through-going turn, cannot reach their lane and the intersection is not used optimally. This can be solved by setting the minimal length of the left and right turn lanes equal to the length of the through-going lane. This is done, but was discarded, as this causes certain problems in the further creation of the network.

Step 2.2-g: Create layout of the intersection (sig_f06_links_nodes)

Using the base coordinates determined in step 2.2-a and the length of the lanes of each turn, the full layout of the intersection can be created.
In the creation of the intersection it is assumed that the through-going turn is always connected to the incoming arterial link. Additional nodes are inserted into this link, to which the left and right turn can be connected. In case no through-going turn is present (in a T-intersection for example),

a fictional version of this lane is created, so the other turns can still connect to the arterial. The result for our example intersection is shown in Figure A.14.

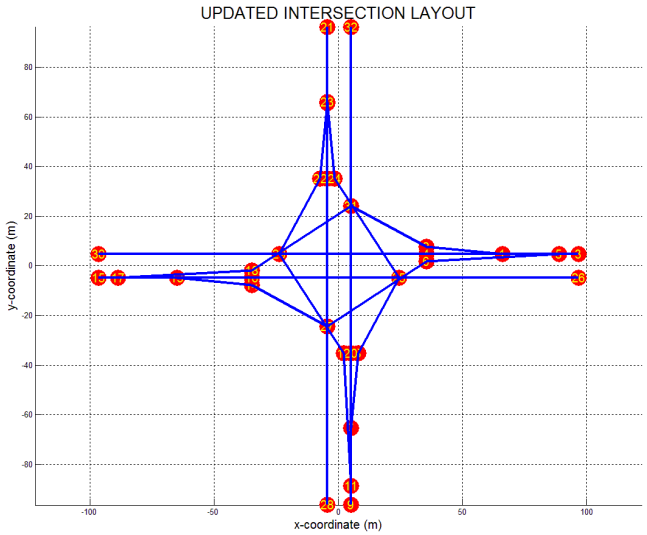


Figure A.14: *Layout of example intersection, after calculating signal scheme and lane lengths*

Step 2.2-h: Create full network (nwc_f04_build full network)

After the layout of all intersections has been calculated, every intersection is placed on its original location in the grid. Then the arterial links are added, interconnecting the intersections. The resulting network is shown in Figure A.15.

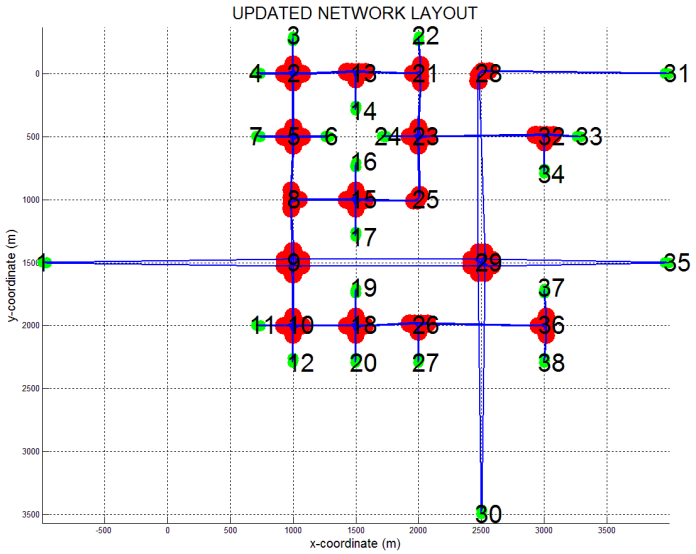


Figure A.15: *Network with all updated intersections*

Step 2.3: Create Origin-Destination-matrix (nwc_f05_create_OD)

As fictional networks are used, traffic flows in the network have to be created as well. This is simply done by generating a random number of trips between every origin and destination. The

origins and destinations in the network are all intersections that are only connected to the network at a single side, the so-called 'feeders' (shown in green in Figure A.15).
The number of trips assigned between an origin and a destination is drawn from a uniform random distribution. The minimum number of trip is set to $n_{trips,min} = 1$. The maximum number of trips is read from the Excel-sheet in which the total number of trips in the network $n_{trips,tot}$. The upper bound for the number of trips between every origin r and destination s is then calculated as $n_{trips,max,r,s} = n_{trips,tot} / \sum_{r,s \in Z} 1$. After all trips have been randomly assigned, the total OD-matrix is scaled, to match $n_{trips,tot}$.

The value of $n_{trips,tot}$ should be chosen in such a way that the resulting network looks 'reasonable'. When the number of trips is chosen too high, the resulting network contains links and/or intersections with 4 or more lanes. As this is almost never the case in an existing network, it can be concluded that the number of trips is not in accordance with the size or layout of the network. When on the other hand the number of trips is too low, no congestion arises and the need for perimeter control would be non-existent. The total number of trips is chosen manually, after which the network created by the model is checked on its realism. If needed, $n_{trips,tot}$ is changed to obtain a more realistic network. Developing and adding general rules-of-thumb between network size and a reasonable amount of trips could improve the model.
The next step is to create additional traffic on the freeways. The reason for this is, is that in most cases, a substantial amount of traffic on a freeway only passes a city and does not enter it. However, excluding this traffic in the model would not be realistic, as it does influence the way traffic is processed in the network. When no additional traffic would be present, vehicles can easily enter the freeway and leave the network in almost any case. With additional traffic, congestion at off-/onramps and spill-back onto the intersection is created.
The additional traffic is created by adding trips between the freeway feeders up to 70% of the estimated capacity of the freeway ($C_{fw,lanes,max} = 2400\, veh/h$).

Although VISSIM is a dynamic traffic simulator, the OD-matrix is not dynamic and only contains the number of trips. It is within VISSIM that the traffic is dynamically loaded onto the network, over a period $t_{sim,tot}$ using a random distribution.

Step 2.4: Determine k-shortest paths between all OD-pairs (nwc_f06_assign_traffic)
Once the network has been constructed, it becomes possible to create paths between all origins and destinations. As within the next step a deterministic user equilibrium (DUE) assignment is performed, multiple paths between all origins and destinations are needed. These paths are generated using a k-shortest-paths algorithm. This algorithm uses the Dijkstra-algorithm to find the shortest path, based on the free flow travel time.
On normal links the travel time for each link i is calculated as

$$tt_{f,i} = \frac{l_i}{u_{f,i}/3{,}6}. \qquad \text{(A.19)}$$

For links within the intersection, the average turning delay is incorporated as well. The free flow travel time on such a link i within intersection int is calculated by equation (A.21), which has been adopted from Muller et al. (2011).

$$tt_{f,i,int} = \frac{l_i}{u_{f,i}/3{,}6} + \frac{t_{red,i}^2}{2t_{cyc,int}[1-q_i/(q_{sat,i}\beta_{adj,i})]}. \quad \text{(A.20)}$$

When a path is found, the free flow travel times of the links of that path are increased ($tt_{i\in pt} = 1{,}50 \cdot tt_i$) and a new shortest path between the origin and destination is generated. The new path is accepted if the total overlapping length of the paths is less than a predefined value ($P_{ovl,pt} = 90\%$), adopted from Bliemer et al. (2004) , as presented in van Nes and Bovy (2008)). The number of paths per OD-pair can be set manually. Within this thesis $n_{pt} = 5$ is used. Although more paths would probably render better results, this value has been chosen, as this results in reasonable computational times. Apart from that, the DUE-assignment is only meant to create an initial network layout. The final network update is done using the actual simulation results from VISSIM, which should be more accurate. In Figure A.16 an example is given of the two available paths between origin 1 and destination 33. Due to the limited size of the network, no more than two paths are found.

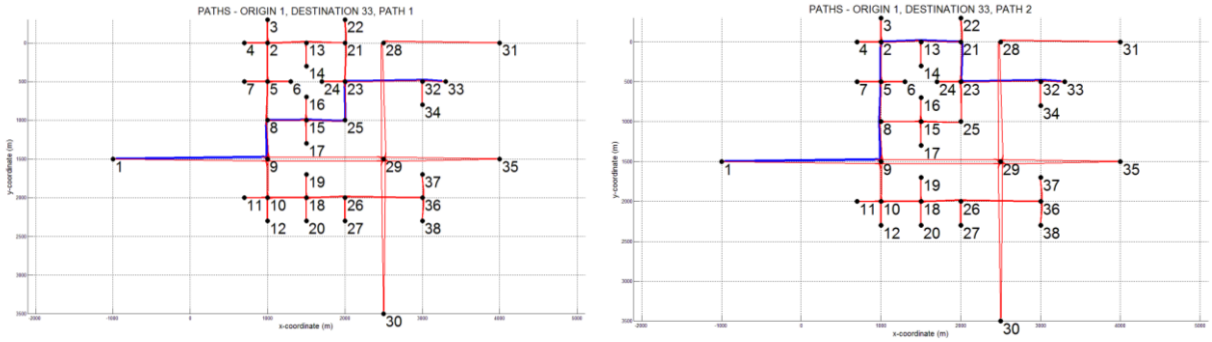


Figure A.16: *Two paths (in blue) between origin 1 and destination 33*
Travel time path 1: 355 seconds, travel time path 2: 689 seconds
Overlap: 69,7%

After obtaining the paths, a Deterministic User Equilibrium (DUE) assignment, using the Frank-Wolfe algorithm and the Method of Successive Averages (MSA) is performed.
The assignment is initialised by loading all trips between every OD-pair onto the first available path. Based on free flow travel time $tt_{f,i}$, the intensity q_i and the capacity C_i of every link i, the BPR function (A.21) is used to calculate the travel times on all links in the network.

$$tt_{DUE,i} = tt_{f,i}\left(1 + \alpha_{BPR}\left(\frac{q_i}{C_i n_{lanes,i}}\right)^{\beta_{BPR}}\right), \quad \text{(A.21)}$$

in which

$$q_i = \frac{n_{veh,i}}{t_{tot}/3.600}. \quad \text{(A.22)}$$

The parameters α_{BPR} and β_{BPR} are scaling factors and are set to $\alpha_{BPR} = 0.15$ and $\beta_{BPR} = 4$, as given in van Nes and Bovy (2008). These values however, are mainly used for normal freeway conditions and should therefore not be fully applicable to links within the urban network. However, no suitable parameters for the urban network were found in literature. Nevertheless, as is stated before, the assignment is only used to create an initial demand for all links, the actual network update is done using results from our VISSIM simulation.

The convergence criterion used is the duality gap DG, which calculates the difference between the total travel time tt_{tot}, at the current iteration itr

$$tt_{tot,itr} = \sum_{pt \in Z} q_{pt,itr}\, tt_{pt,itr} \tag{A.23}$$

in which

$$tt_{pt,itr} = \sum_{i \in pt} tt_{DUE,i,itr} \tag{A.24}$$

and the total travel time of all vehicles over the shortest path for all origins r and destinations s,

$$tt_{tot,f} = \sum_{r,s \in Z} min(D_{r,s}) q_{r,s} \tag{A.25}$$

in which $D_{r,s}$ is the set containing all $tt_{f,pt \in r,s}$.The duality gap is calculated by

$$DG_{itr} = \frac{tt_{tot,itr} - tt_{tot,f}}{tt_{tot,f}}, \tag{A.26}$$

with the stop criterion

$$DG_{itr-1} - DG_{itr} < DG_{prec}. \tag{A.27}$$

The stop criterion is set to $DG_{prec} = 10^{-4}$, which generally takes between 30 and 50 iterations to achieve. After the assignment has been completed, the final link loads are obtained and used to update the network.

Step 3: Update network (static)

Using the link intensities obtained in the previous step, the links and intersections are updated in order to match the network demand. The number of lanes of the arterial and freeway links are determined by dividing the intensity over the lane capacity. Intersections on the other hand are updated by adding lanes up to the point where the cycle time is less than the minimum allowed cycle time.

Step 3.1: Updating the network (nwc_f07_update network)

After the initial link intensities have been obtained by the DUE-assignment, the network undergoes its first update.

First off, the number of arterial links are updated, by dividing the link intensity q_i by the lane capacity $C_{lanes,art,i} = 2000\ veh/h$.

$$n_{lanes,art,i} = q_i / C_{lanes,art,i} \tag{A.28}$$

Next the number of lanes of freeways are updated. This is done in mostly the same fashion as the arterial links, with the difference that the lower bound of the number of lanes is set (in this case $n_{lanes,fw,min} = 3$). With $C_{lanes,fw} = 2400\ veh/h$, the number of lanes for a freeway is then

$$n_{lanes,fw,i} = max\left(q_i / C_{lanes,fw,i}; n_{lanes,fw,min}\right) \tag{A.29}$$

The first thing that is done for each intersection, is updating the traffic demand at each turn $q_{int,tn}$, using the link loads obtained from the DUE-assignment. The next thing that has to be done, is to determine whether the intersection should be signalised or not. The purpose of this is to prevent that simple intersections are signalised, which would not really match reality. Also the simulation speed in VISSIM decreases by adding additional signalised intersections, thus removing signals has a positive effect on the total simulation time. Whether an intersection should be signalised or not, can be determined by using the intensity criteria developed by Slop (see CROW, 2002)

$$\alpha_{SLOP} = \frac{q_{sub}}{q_I}\left(-1 + \sqrt{1 + \beta * (q_{main}/q_{sub})}\right), \tag{A.30}$$

In which

q_{main}	Intensity on the main direction of both ways (veh/h);
q_{sub}	Intensity on busiest crossroad (veh/h);
q_I	Parameter (see Table A.9);
β_{SLOP}	Parameter (see Table A.9).

Based on the value obtained for α_{SLOP}, a decision to signalise an intersection can be made. For four-legged intersections signals are desired if $\alpha_{SLOP} \geq 1{,}33$. If $1{,}00 < \alpha_{SLOP} < 1{,}33$, than signals are not directly necessary and other factors should be taken into consideration whether or not to use

them. As signals are very important in this model, intersections are signalised when $\alpha_{SLOP} \geq 1{,}00$. The same principle holds for three-legged intersections, only in this case signals are used when $\alpha_{SLOP} \geq 1{,}33$.

Table A.9: *Parameters for q_I and β_{SLOP}*

Number of lanes on main road in both directions	Number of lanes on crossroad	q_I (veh/h)		β_{SLOP}
		≤ 50 km/h	> 50 km/h	
1	1	300	210	2,4
> 1	1	300	210	2,0
1	> 1	400	280	3,2
> 1	> 1	400	280	2,7

In case an intersection is signalised, the updated link intensities are used to determine the new signal scheme. As the intensities can become very high, it can occur that it is not possible to find a new signal scheme with a cycle time shorter than the maximum allowed cycle time. If this is the case, an additional lane is added to the turn tn with the highest intensity per lane, i.e. $\max_{tn \in int}(q_{int,tn}/n_{lanes,int,tn})$. This process is repeated until a signal scheme with a cycle time shorter than the maximum allowed cycle time is found.

In Figure A.17 an updated version of intersection 2 is given, in which turns 1 and 12 now both have 2 lanes. The control structure used and the cycle time have also changed. Total cycle time of the intersection is now 63,0 seconds instead of the previous 82,7 seconds, so the intersection has been significantly improved by adding additional lanes. This however is not always the case. Especially with busier/larger intersections, cycle times are often close to the maximum allowed cycle time.

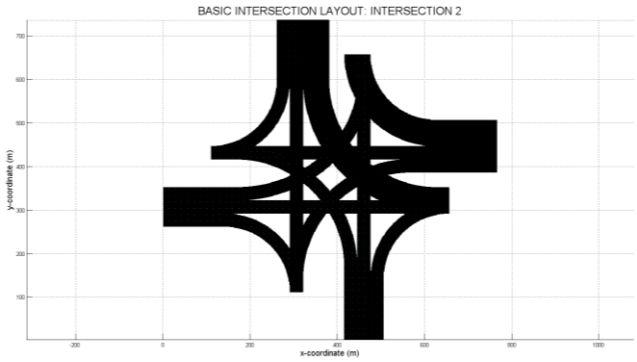


Figure A.17: *Intersection after update*

After all intersections and links have been updated, in order to match the demand in the network, a new set of links and nodes is created and the initial version of the network is complete. As only the number of links will have changed and only a few links will have been removed, the resulting layout closely resembles the previous layout, as is shown in Figure A.15.

After this step it could be possible to return to step 2.4 and recalculate the travel times based on the new network, perform a new DUE-assignment and re-update the network, until some convergence criterion has been met. For the sake of computational speed this has not been done, as the network is also updated using the simulation results of VISSIM.

Step 4: Create subnetworks

A next important step is to create subnetworks. The subnetworks are created as the areas between or next to the arterials. After the area of each subnetwork has been obtained, a random street pattern is inserted into this subnetwork. Then new origins and destinations are created in each subnetwork and trips are reallocated and spread out over the subnetwork. Finally the subnetwork layout is converted to a link-node structure and is prepared for its export to VISSIM.

Step 4.1: Abstract subnetworks from network (snc_f01_create_subgrid)

The basis for creating the subnetworks is provided by the original grid created in step 1. The first step is to construct a grid, from which the different subnetworks can be isolated. Based on the size of the network and a predefined width for the grid fields ($w_{x,grid} = w_{y,grid} = 20\ m$), a basic grid is created. As the subnetworks cannot cross the arterial links, the links are added to the grid. On either side of the links a 'safety zone' is created of $w_{safe} = 50\ m$ wide. The purpose of this zone is to create enough space for the arterials to fit into later on, without intersecting with the subnetwork links. Apart from that, this also matches real networks, as houses are often build some distance away from the main roads. Furthermore, at intersections, an 'indent' $l_{ind} = 250\ m$ is made into the subnetwork, which can accommodate the auxiliary lanes of that intersections. The last thing that is done, is to add the links penetrating the subnetworks, which can then later be used to connect to the arterial. The result of this is shown in Figure A.18.

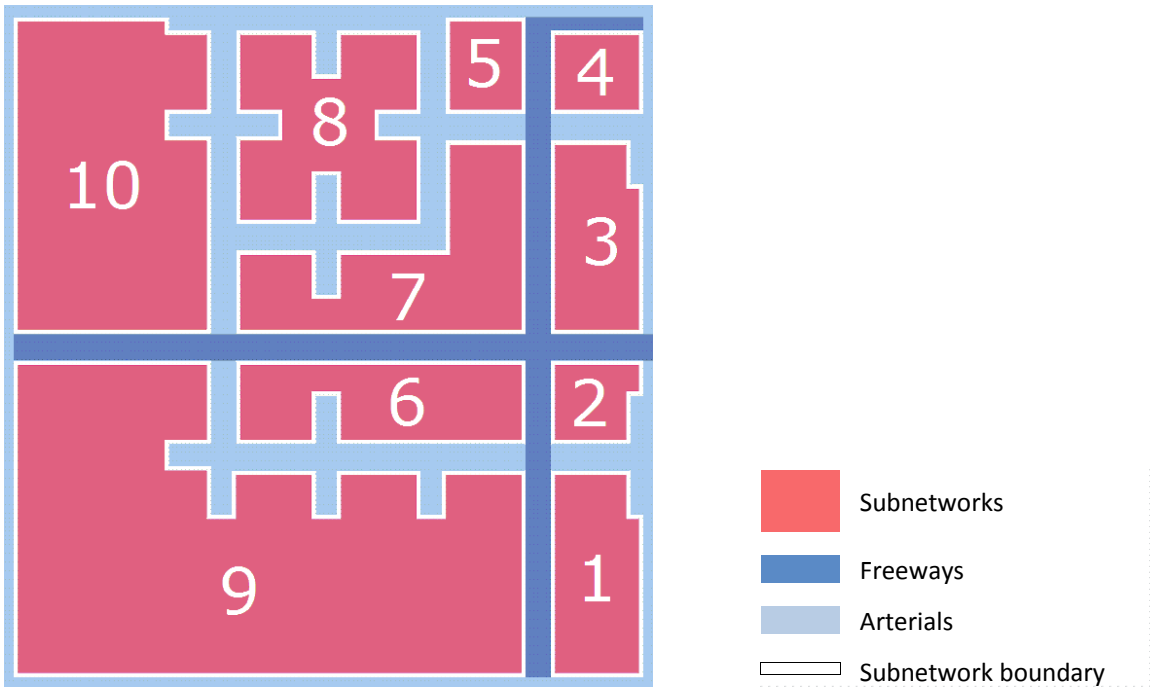


Figure A.18: *Basic subnetwork grid*

The areas shown in red are converted into subnetworks, the numbers shown in Figure A.18 are the numbers of the different subnetworks.

Step 4.2: Create roads in the subnetwork (snc_f02_create_subnetworks)

After the area for every subnetwork has been obtained, links within the subnetworks have to be generated. This is done by adding blocks into the subnetwork and adding streets around them. The size of the blocks are based on the size of different area types found in a common neighbourhood. Within this model, so far only components found in a normal residential neighbourhood are used, which are terraced houses, flats, parks and sport facilities. Additional types of areas can be added later on.

Based on measurements made in aerial photographs (Google Earth), upper and lower bounds for the length and width of each area type are measured. Also an estimation of the amount of space covered by these different area types P_A within a neighbourhood is made. Table A.10 shows the parameters that are used.

Table A.10: *Parameters used for subnetwork blocks*

Type of area	$w_{x,min}$ (m)	$w_{x,max}$ (m)	$w_{y,min}$ (m)	$w_{y,max}$ (m)	P_A (%)
Terraced houses	50	100	80	180	50
Flats	60	90	120	200	25
Parks	50	400	50	400	10
Sport facilities	150	400	200	500	15

Based on the space covered by each type of area b and the average size, the distribution ratio γ_b of each area type is calculated as

$$\gamma_{aux} = \frac{P_A}{100}\left(\left(\frac{w_{x,min,b} + w_{x,max,b}}{2}\right)\left(\frac{w_{y,min,b} + w_{y,max,b}}{2}\right)\right), \quad \text{(A.31)}$$

$$\gamma_b = \gamma_{aux} / \sum \gamma_{aux}. \quad \text{(A.32)}$$

Based on the distribution of the blocks, a type of block is chosen at random, after which its length and width are drawn from a uniform distribution between the minimum and maximum size of that type of area. Next an unused area within the subnetwork, with the same size as the block is chosen and the block is put onto the grid, after which a road is added around it. This process is repeated, until no more block can be fitted into the grid. The remaining areas are filled up as parks.

After all blocks have been allocated, roads with a parallel spacing $s_{i,sn-i,sn} \leq 50\ m$ are removed, as this does not generate a realistic network. Also bayonets and roundabouts are reconstructed to form normal intersections. Reason for this is that they create some problems in VISSIM, as vehicles easily get 'stuck' at these points, creating a complete gridlock in the subnetwork. The resulting subnetwork, for subnetwork 9 (as given in ***Figure A.18***) is shown in Figure A.19.

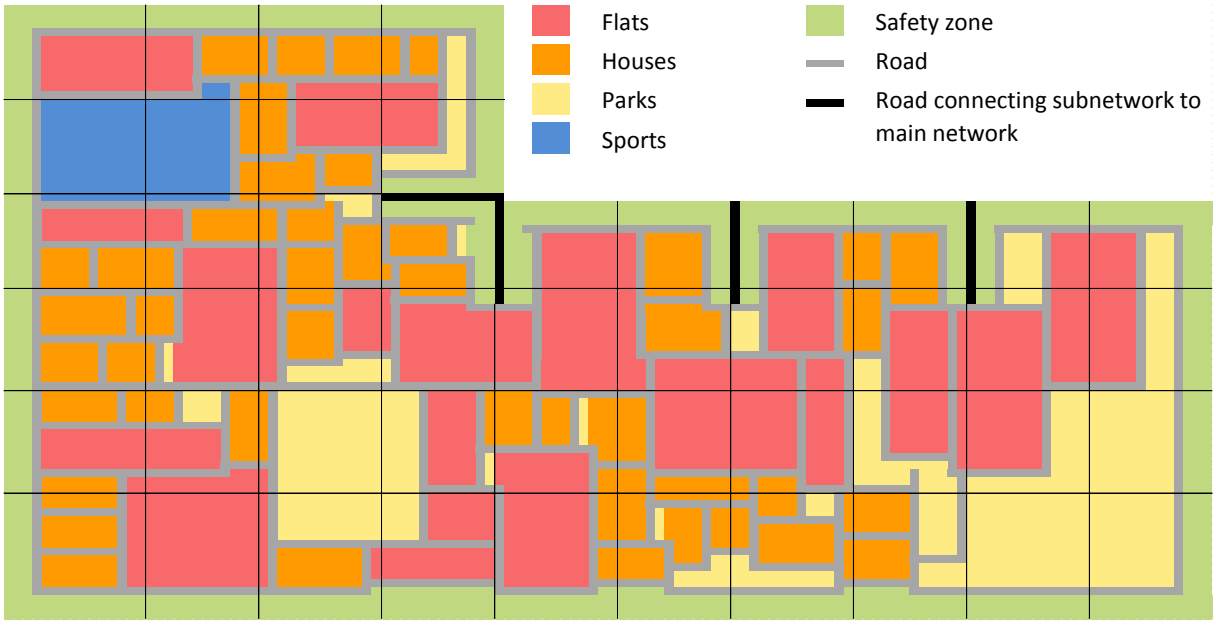


Figure A.19: *Subnetwork after blocks are inserted*

Step 4.3: Add feeders and modify OD-flow allocation (snc_f03_convert_subnetwork)

Step 4.3-a: Create origins and destinations in the subnetwork

After the street pattern within each subnetwork has been created, feeders (an intersection/link containing an origin and destination) are created within the subnetwork. As described earlier in step 2.3, the origins and destinations are located at the boundaries of the network and at the end of the subnetwork connectors (black roads in Figure A.19). However, to create traffic flow inside the different subnetworks, feeders have to be allocated within the subnetwork. Also a more diverse traffic flow pattern with many possible paths is obtained.

Allocating the feeders is done by dividing each subnetwork in multiple areas, with a size equal to half the intersection spacing (see gridlines in Figure A.19). Within each of these areas one feeder is created (if the available space allows it).

Step 4.3-b: Assign original OD-demand to newly created origins and destinations

When the feeders have been created, each feeder is added to the OD-matrix. The original feeders connecting to the subnetwork are removed. The flows to and from the original feeders are then distributed over the feeders in the subnetwork. The number of trips assigned to each feeder fn in the subnetwork sn is based on their relative distance to each original feeder fo connected to sn, is calculated as

$$n_{trips,fn} = \sum_{fo \in sn} n_{trips,fn} \left[\frac{|XY_{fn}, XY_{fo}|}{\sum_{fn \in sn} |XY_{fn}, XY_{fo}|} \right]. \quad \text{(A.33)}$$

All trips between the original intersections are removed and no internal trips are made. This is based on the assumption that travel distances within the subnetwork are sufficiently small, to make them by other means of transport, such as walking are cycling. As a consequence, the total number of trips made in the network decreases. To this end, the OD-matrix is again scaled to match $n_{trips,tot}$. Although this disrupts the created balance between the OD-matrix and the network

layout, this is compensated later on, as the network is updated with the simulation results generated by VISSIM.

Step 4.3-c: Convert subnetwork layout to link-node structure

After the feeders have been created, the grid structure is converted to a normal link-node structure. Opposed to the arterials, the subnetwork links are not updated. The main reason for this is, that a limited amount of infrastructure generates a large amount of different routes through the subnetworks and traffic is spread more evenly over the network (adding multiple lanes to roads in the subnetwork will cause traffic to concentrate along these roads in the subnetwork, creating an uneven distribution of traffic over the subnetwork). Also multi-lane roads within subnetworks are not very common. The result of this step is shown in Figure A.20.

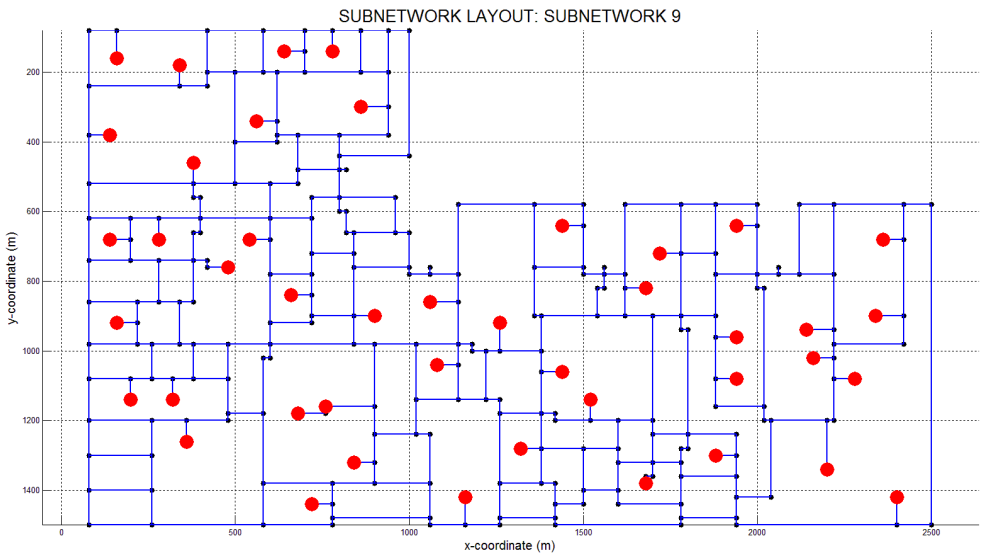


Figure A.20: *Subnetwork after adding feeders and conversion to links and nodes (nodes are feeders)*

Step 4.4: Convert subnetwork layout to VISSIM format (snc_f04_subnetwork_VISSIM)

When the subnetwork has been converted to a link-node structure, the subnetwork is prepared for its export to VISSIM. Links created in the previous step are bi-directional and have to be split in a separate link for each direction, as VISSIM does not support bi-directional links. Also the nodes are converted into intersections. As no signals are used in these intersections, these are fairly easy to make. As VISSIM uses so-called splines, which are curved links, each individual spline point has to be calculated. Using the equation for a normal circle, each of these points is calculated. The radii used are $\varphi_{left,sn} = 9{,}5$ and $\varphi_{right,sn} = 6{,}5$.

Step 5: Export network to VISSIM

In this step, all of the various components designed and created in the previous steps are exported to VISSIM. In order to do so, the layout of the intersections has to be redesigned, as VISSIM uses a link-connector structure instead of the link-node structure designed previously. In redesigning the intersections, the main task is redefining the way lanes from the network connect to the different lanes of each turn. When this has is done, all links and connectors in the intersection can be created. After the final intersections have been created, they are (where applicable) connected to the different subnetworks and the network is finally completed.
With the total network done, the different components are written to a number of text files, which enable visualization and simulation of the network in VISSIM.

Step 5.1: Convert intersections to VISSIM format (vex_f01_00_rewrite_network)

Step 5.1-a: Configure lanes (vex_f01_01_configure_lanes)

Before the network can be exported to VISSIM, its link-node structure has to be converted to the link-connector structure used in VISSIM. Within VISSIM, links are connected to each other by connectors, which are roughly similar to links. An important property of both the link and the connector is that they cannot be connected to their same type, e.g. links cannot connect to links and connectors cannot connect to connectors. Another important difference between the link and the connector is that a link is a straight line and that a connector can be curved.
It also has to be taken into account that connectors do not connect to the beginning or the end of the link, but have to connect to the link at a certain location. And in order to obtain a reasonably looking network, the coordinates of these connection points must be quite accurate, or otherwise vehicles will make strange moves.
And last, connectors can only connect an equal amount of lanes to each other, meaning that multiple connectors have to be used when the number of lanes at the downstream link has increased or decreased.

Within the link-node model, no distinction between the links is made and basically all links are linear, cutting curves short. To come to a proper representation of an intersection within VISSIM, the links and connectors have to be sequenced in the right order, so that the connectors are located at the points where curves are made (diverging points, merging points and corners).

In Figure A.21 the problem arising from this is clearly illustrated. As can be seen, the through-going turn [8] has been split into multiple links and nodes, to create connection points for the links of the left and right turn. When a 1-to-1 conversion to the link-connector situation is to be made, links [7], [8] and [9] have to become connectors, as they should be curved. This causes links [4], [5] and [6] to be links and [3] to be a connector. The problem now is that in this case, link [2] both has to be a link and a connector.

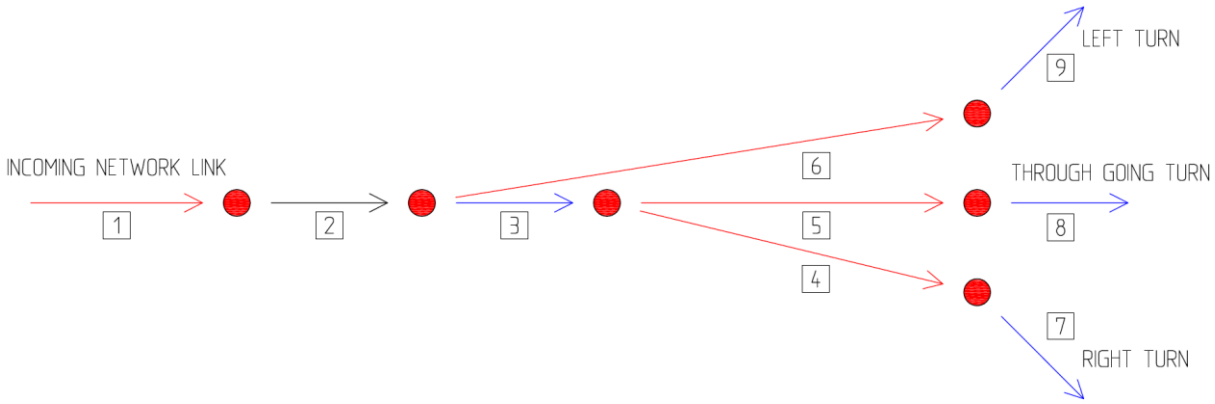


Figure A.21: *Example of original link-node structure at intersection*

Although the problem could be solved by adding an additional connector between [2] and [3], this does not fully solve the problem, as multiple other configurations occur. Apart from that, the resulting intersection would have large overlapping sections. Although VISSIM can handle this, the resulting layout is visually still incorrect. A better way to create this structure is shown in Figure A.22, which utilises the property of the connector, that it can be connected to any location in a link.

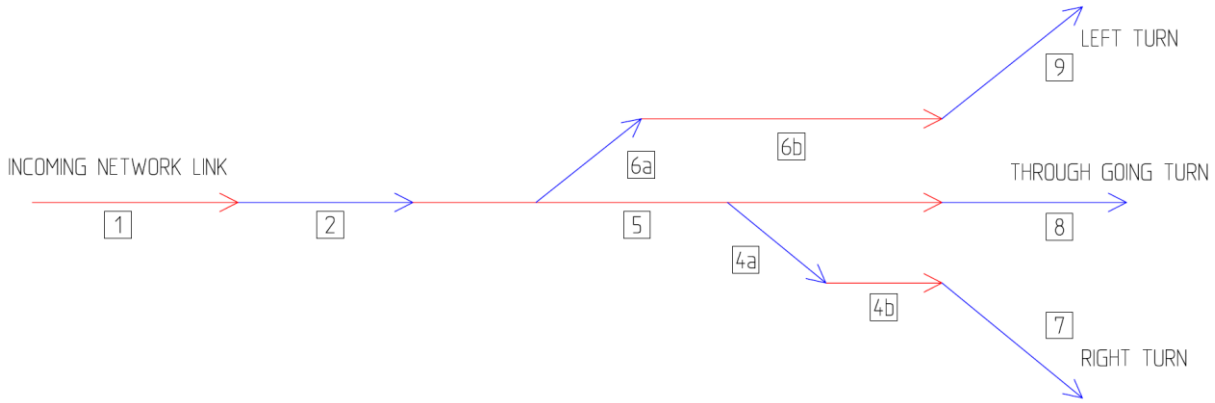


Figure A.22: *Example of link-connector structure in VISSIM*

The consequence of this is that the physical layout of the intersection has to be restructured, to work properly in VISSIM. However, as connectors can connect to a specific lane in the link, the way that the link of each turn is connected to the preceding link differs as well. If in the above example link [1] would consist of two lanes, and links [4b], [5] and [6b] would all have a single lane, than link [4b] and [5] could be connected to the right lane and [6b] to the left lane.

As multiple configuration types are possible, each has been pre-programmed to configure the incoming and outgoing lanes in a specific manner, as has been described in the cases below. A key assumption underlying all configurations is that the number of incoming lanes can never be greater than the number of lanes at the stop bar. This assumption seems to be justified, as the capacity of incoming lanes is higher than the capacity of lanes within the intersection. Within our model, this assumption is always true, as the $C_{lanes,art,i} = 2000\ veh/h$ and $q_{sat} = 1800\, veh/h$.

Case 1: Number of incoming lanes is equal to the number of lanes at the stop bar

Configuration of this structure is fairly straightforward, as each turn is connected to the incoming link on a 1-to-1 basis, as is shown in Figure A.23.

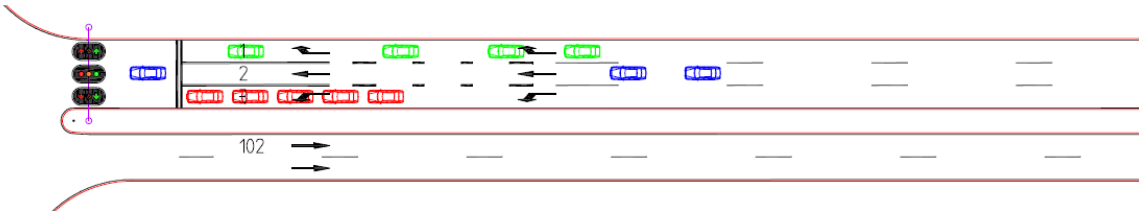


Figure A.23: *Configuration of case 1*

Case 2a: The through going lane is present and the number of through going and left turn lanes is equal to or larger than the number of incoming lanes

The basic principle behind this configuration is that it should always be tried to connect the rightmost incoming lane to the rightmost lane of the through going turn. Additional lanes are created on either side of the through going turn. In case the number of incoming lanes is also smaller than the number of through going lanes, additional lanes for the through going turn are made as well. The resulting configuration is shown in Figure A.24.

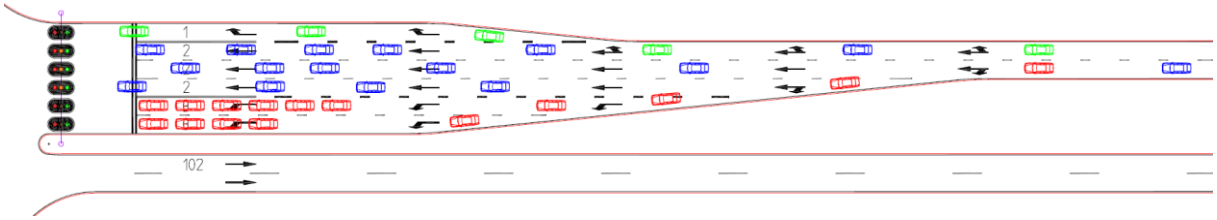


Figure A.24: *Configuration of case 2a*

Case 2b: The through going lane is present and the number of through going and left turn lanes is smaller than the number of incoming lanes

The basic principle behind this configuration is that the rightmost lane of the through going turn should be as close as possible to the rightmost incoming lane. Incoming lanes are assigned to the turning lanes from left to right. So first the lanes are connected to the leftmost lane of the left turn, up to the rightmost lane of the through going turn. Remaining incoming links are assigned to the right turn from left to right. Additional lanes are added to the right side of the right turn. An example of this configuration is shown in Figure A.25.

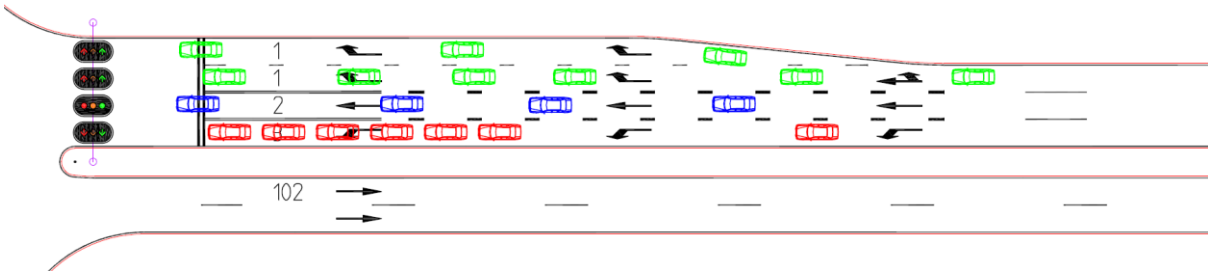


Figure A.25: *Configuration of case 2b*

Case 3: Only the left and right turns are present

In case only the left and the right turn are present (a T-intersection), the lanes are centred in such a way that the right half of the incoming lanes are assigned to the right turn and the left half of the incoming lanes to the left turn. In case the number of incoming lanes is unequal, the remaining lane is assigned to the turn with the highest intensity. The resulting configuration is shown in Figure A.26.

Case 4: Only the straight and one turning lane is present

In case only the left or the right turn is present, the incoming lanes are first assigned to the through going turn, from right to left. Any remaining lanes are then assigned to the left or right turn. If the number of incoming lanes is smaller than the number of through going lanes, then additional lanes for the through going turn are added. The resulting configuration is also shown in Figure A.26.

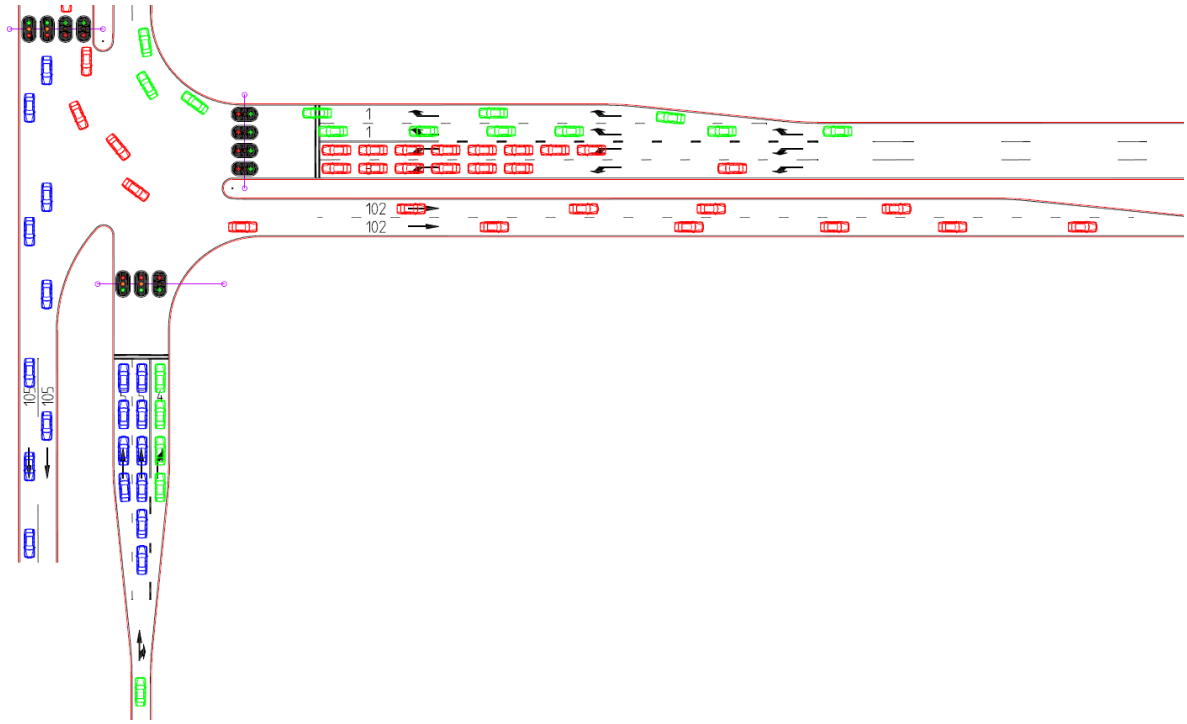


Figure A.26: *Configuration of case 3 (upper-right), case 4 (lower-left) and case 5 (upper right)*

Case 5: Configuring the outgoing lanes

The next step is to create the configuration of the outgoing lanes to the arterial links. As the capacity of the arterial links is higher than those within the intersection, it automatically follows that the number of outgoing lanes is always equal to or less than the number of lanes coming from the different turns. As a consequence, only a single configuration case exists. Basically the outgoing lanes are connected to the turn lanes from right to left. Any remaining turn lanes are connected to the leftmost lanes of the outgoing link. Again, the resulting configuration is shown in Figure A.26.

However, it is also checked if the number of turning lanes at the first upstream intersection is equal to or higher than the number of outgoing lanes of the current intersection. If so, the lanes are not merged. This is done to avoid a lot of blocking back in VISSIM, as a result of this merge. This also has something to do with the fact that merging and diverging behaviour in congested conditions can sometimes be quite poor in VISSIM.

Step 5.1-b: Recalculate base-coordinates of turns (vex_f01_02_calculate_coordinates)

Based on the new configuration, the coordinates of the base points of the different links in the intersection have to be calculated. Depending on the configuration of the lanes, the length of the link of certain turns has to be adapted, to enable the connectors to connect to them.

A good example of this is provided by Figure A.24. If for instance the original length of turn 2 has to be 55 meter and the length of the turn 3 has to be 40 meters, then there is no sufficient length to connect the connectors of turn 3 to the link of turn 2. As suggested by CROW (2002), the angle at which a new lane is added, is $\theta_{lanes,int} = {}^{1}/_{10}$. As we are using lane width of $w_{int,lanes} = 3{,}0\ m$, the additional length needed is 30 meters. In our example, the length of the link $l_{i,tn}$ at turn 2 must be at least $l_{i,2} = max(l_{i,3} + \frac{n_{lanes,3} w_{int,lanes}}{\alpha}; l_{i,2}) = max\left(40 + \frac{2*3{,}0}{1/10}; 55\right) = max(100; 55) = 100\ m$.

This calculation is performed for all turns and based on the value obtained and the original location of the stop bar, the new coordinates are calculated.

The length of the 'backlanes' (the lanes after the intersection has been passed) are calculated in mostly the same fashion. The start of the merging point should be located at a distance of at least $d_{int,stop-merge} = 150m$ behind the stop bar of the opposing direction (for instance, the merging point for turn 102 is 150 meters behind the stop bar of turn 8) according to CROW (2002). However, for T-intersections this cannot be calculated for all backlanes (see Figure A.26 as an example). To circumvent this, the assumption has been made that the average distance between the stop bar and the start of the backlane (which has been set equal to the location of the stop bar of the incoming turns of the same direction; turn 2 in our example) is $d_{int,stop-merge} \geq 60m$. This then implies that the merging point is located at a distance of 90 metre from the start of the backlane.

When all of these lengths have been calculated, the x- and y-coordinate of every point is determined within the relative coordinate-system of the intersection. Based on the coordinates of the start points of the outgoing links, the intersection is centred at the same location as the previously generated intersection. The reason that the intersection is matched with its original

position, is that the turn splines, created in step 2.2-a are in the same coordinate system as the new intersection.
The result of this step for intersection 2 is shown in Figure A.27.

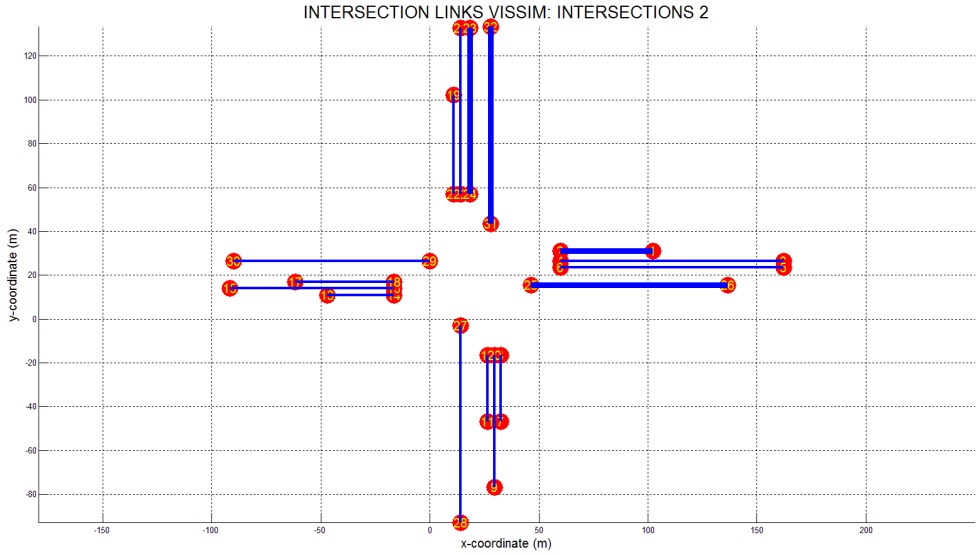


Figure A.27: *Newly created links for VISSIMification of intersections*

After the length of all the intersection links has been determined, the length of the arterial links is recalculated, as they have changed due to the elongation of the intersection links.

Step 5.1-c: Calculate connector and link lengths (vex_f02_01_calculate_connectors)
Figure A.27 clearly illustrates that the intersections are far from completed at this stage. Based on the configuration of the incoming and outgoing lanes and the base the nodes, links and connectors making up the intersection can all be calculated.

The first step is to turn the base points into nodes and links, as is already shown in Figure A.27. Based on these nodes, the connectors connecting the incoming network links to the intersection are made. The basis of this is formed by the configuration made in step 5.1-a. For each connector it is calculated which link it comes from and goes to, and at which distance and to what lanes it is connected to these links. This process is repeated for the connectors within the intersection (the actual turns) and the connectors between the backlanes and the network.
When this is done, the actual coordinates of each endpoint of the connectors is calculated. And although the actual connection is automatically made at the right point within the link, the coordinates still have to be determined, as they are used to determine the initial and final orientation of each connector. As such, each connector has at least four so-called spline points, two located at the start and two at the end. Misplacing these points results in strange turns by vehicles, increased link/connector lengths and most important a poor visualization of the network. The final visualization of the resulting intersection in VISSIM is shown in Figure A.28. The roads and intersections in the top corners belong to the subnetworks.

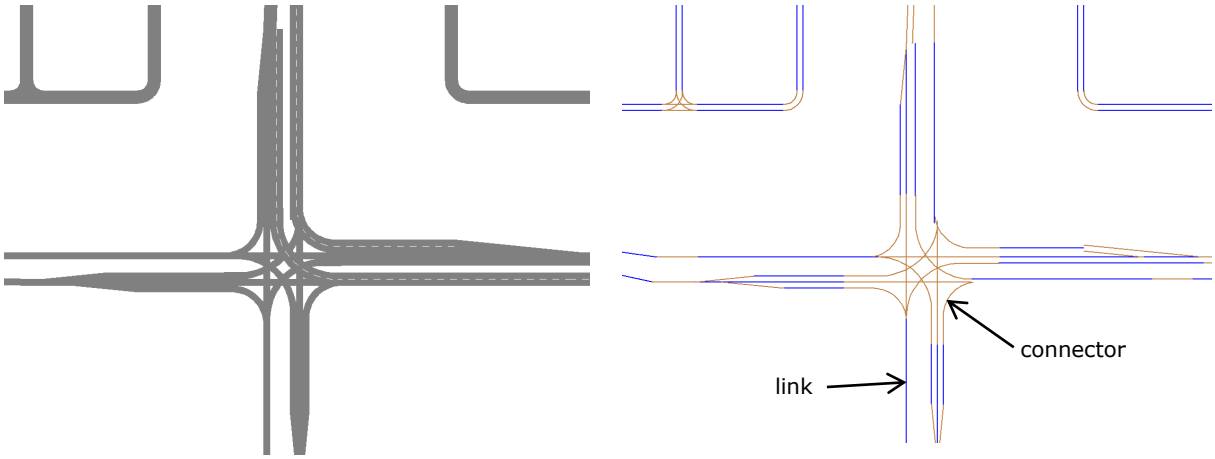


Figure A.28: *Intersection 2 in VISSIM*
(left = normal, right = centre line; blue = link, brown = connector)

Apart from the normal intersections, off-/onramps and freeway junctions are also reconstructed at this stage. This process is somewhat more elaborate, as also a merging lane and the ramps itself have to be created. For the sake of convenience, the number of lanes on the ramps has been set to 2 in all cases. Apart from that, this also circumvents some issues within VISSIM with some strange traffic behaviour at the off-ramps. This can however cause congestion in cases where the number of lanes should be higher. Due to time and scope-restrictions this will not be reprogrammed. Instead some modifications to the OD-matrix will be made to solve these issues and some congestion on the freeways is accepted. A visualization of an off-ramp and junction is given in Figure A.29 and Figure A.30 respectively. The radius of the turns in the junction can be adapted to obtain a more realistic representation of the freeway, but has been omitted, as this causes some problems with overlapping zones of the junction and subnetwork intersections.

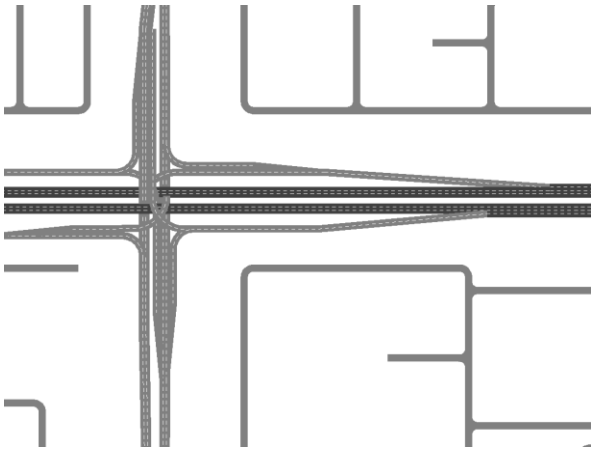


Figure A.29: *Off-/onramp in VISSIM*

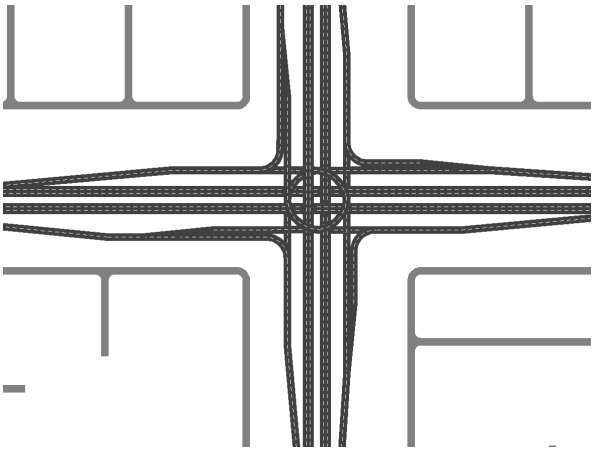


Figure A.30: *Freeway junction in VISSIM*

Further adaptations to the links will not have to be made, as these are simply defined by the x- and y-coordinate of their begin and endpoint and no additional intermediate points are needed.

Step 5.2: Connect subnetworks to arterial (snc_f05_connect_main)

After the layout of the intersections has been determined, it becomes possible to connect the arterial to the subnetworks. Using the links reaching from the subnetwork into the arterials, a connection to the newly formed intersections can be made, with which the network finally forms a completely interconnected structure, which is shown in Figure A.31.

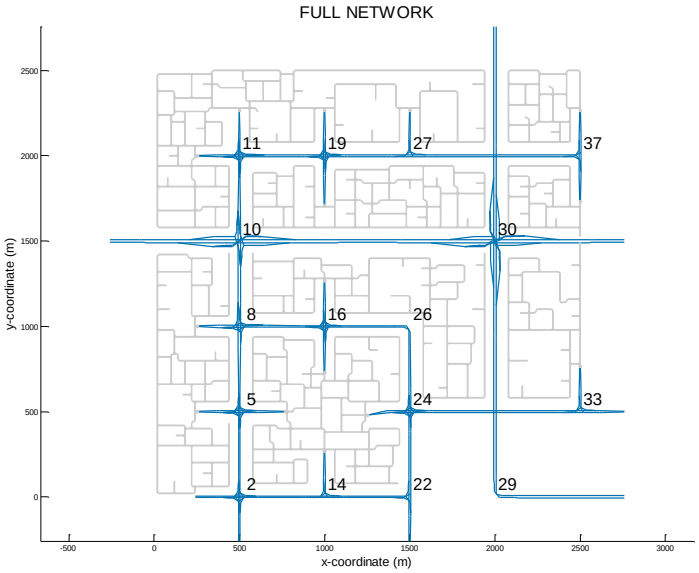


Figure A.31: *Complete network*

Step 5.3: Write data to VISSIM network file (vex_f03_00_write network)

Based on the format of the VISSIM network file, which is written in plain text, all the components created in the previous step are written to the VISSIM network file. Apart from the layout of the network (links, connectors feeders, signal schemes, speed decisions, et cetera), additional parameters for the driving behaviour and the link evaluation are set. Also conflict areas, priority rules and travel time sections are added.

The conflict areas and priority rules are added to improve traffic flow at intersections. A conflict area for example is the area, at which two links or connectors overlap with one another. By adding the conflict area, VISSIM knows that no two vehicles can be at that point at the same time. Also priority is given two one of the traffic streams. An example is given in Figure A.32.

The second addition to improve traffic flow at intersections, are priority rules. A priority rule consists of a stop line and one or more conflict markers. In Figure A.33 stop lines are given in red and the conflict marker in green. This does actually seem to be opposed to the above statement, but in this case multiple conflicts, with one conflict marker and one stop line each have been added. If a vehicle is present between the stop line and the conflict marker, no other vehicle can enter that region, e.g. the conflict marker functions as a red light. However, to prevent traffic halting every time a vehicle is within that region a threshold speed limit is added to the priority rule. The priority rule is only activated when the speed of the arriving vehicle is below the threshold

value, thus not hampering free flow. The threshold speed for subnetwork intersections is set to $u_{i,sn,th} = 15\,km/h$. For arterial intersections $u_{i,art,th} = 30\,km/h$.

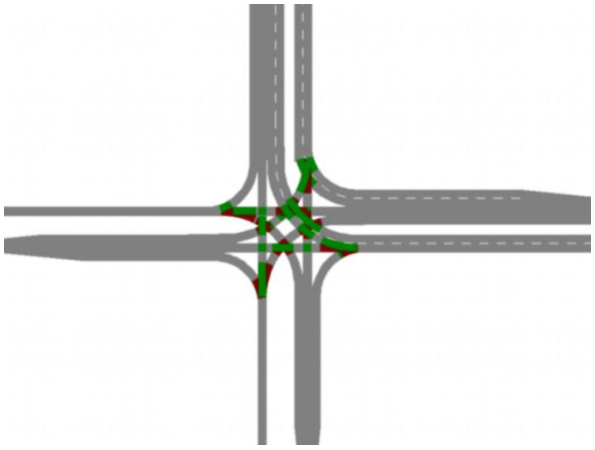


Figure A.32: *Conflict areas*

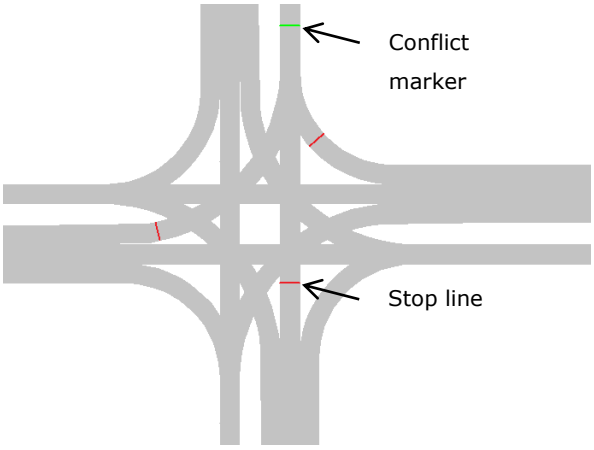


Figure A.33: *Priority rules*

The travel time sections that are added in the network are used at a later stage, in the perimeter control algorithm. Sections on arterial link start at the end of the link and end at the end of the link. At the end of the arterial link, a new section is started, which ends at the start of the next arterial link, thus measuring the average travel time over the intersection.

Furthermore, separate files are written for the signal schemes, the OD-matrix. Also additional files have to be made, containing information on the type of information that has to be measured, such as speed, density, volume for links and speed and distance travelled for vehicles. The resulting network after exporting it to VISSIM is shown in Figure A.34.

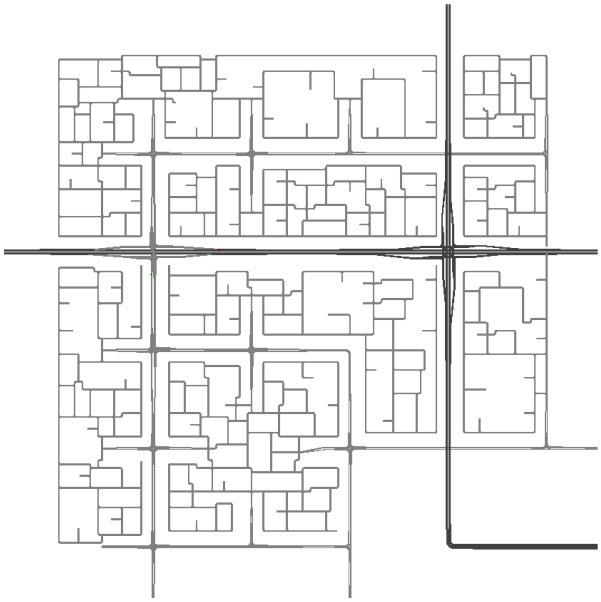


Figure A.34: *Complete network in VISSIM*

Step 6: Update network (dynamic)

After the network is created and has been VISSIMified, the network is updated using simulation results from VISSIM. To this end, an initial number of simulations is run to obtain a substantial path set. After the initial simulations, a simulation is run in which traffic volumes are read from the various links and connectors. Based on these volumes, the number of lanes of the arterial and freeway links are updated and the intersections are fully recreated, following the same process as step 3. Then step 5 is repeated, resulting in a fully updated network. As with the previous network, a number of simulation runs is made to obtain a path set for the complete network.
At this stage the network is finalised and the network is exported to a new folder, containing all data of the network.

Step 6.1: Run initial simulations to create paths (sim_f01_init_VISSIM)

After the network has been exported to the VISSIM-format, a number of initial simulations is run, to create multiple paths in the network. Using the total travel costs in the network, VISSIM adds new paths between each origin and destination at every iteration and at given time intervals ($R_{sim,upd,\Delta t} = 600\ s$). To avoid that the number of paths becomes too large and unrealistic paths are used, the number of paths between every OD-pair has been restricted to $n_{pt,r,s} = 50$, in which the total travel cost of each path cannot be more than $M_{pt,r,s} \leq \alpha_{det} * min(B_{r,s})$ in which $\alpha_{det} = 1{,}5$ and $B_{r,s}$ is the set containing all $M_{pt,r,s}$.

To find the initial paths, a connection between Matlab and VISSIM is created through a COM-interface. Using this COM-interface, the simulations in VISSIM can be controlled completely by Matlab. In this way, data from the simulation can be collected by Matlab and real-time changes to the simulation can be made.
When the simulation is started, only the shortest path between every origin and destination is found and all vehicles are loaded onto those paths. As this can easily cause congestion, only a fraction of the total demand ($\gamma_{demand} = 0{,}2$) is loaded onto the network and a number of simulation runs is performed ($n_{runs,VIS} = 10$). Through a number of test runs, this number of runs was found to yield a sufficiently large path set. The paths found in VISSIM are stored in a separate file.

Step 6.2: Get link lengths from VISSIM and update network (sim_f01_init_VISSIM)

VISSIM calculates the intensities and densities (called volumes in VISSIM) of links and connectors, by splitting the link in multiple segments. The length of these segments can be defined for each link and connector individually. However, reading the data from every individual segment using the COM-interface is very time consuming. As such it is easier to set the segment length equal to the link/connector length. As the connectors in VISSIM are often curved, their actual lengths are larger than the originally calculated Euclidian distance, resulting in faulty data. To solve this, the actual length of each of these connectors is read from VISSIM and changed within the original text file.

Step 6.3: Run simulation to obtain link intensities

Step 6.3-a: Run simulation (sim_f02_getlinkdata)

After the segment lengths have been updated, the simulation is again called from Matlab and a new simulation is run. The simulation is paused at every $t_{data,i} = 600\ s$ and data is read from the links (density, volume, speed and cumulated number of vehicles). The simulation is run until (1) the end time of the simulation is reached $t_{tot} = 14400\ s$, (2) the network is empty $n_{veh} = 0$, or (3) the network is in full gridlock. Gridlock is assumed to be reached when no vehicle currently in the network has entered a different link/connector between the current and the previous time interval.

Step 6.3-b: Update link intensities (sim_f03_update flows)

From the simulation results, the total number of vehicles that have travelled over each link is taken and multiplied by a factor $1/\gamma_{demand}$, as to obtain the full demand. This value is then assigned to each arterial and freeway link. For intersections, the intensities are not updated for each link. Instead the number of vehicles that have travelled over the connector within the intersection are used. The value obtained is then assigned to that specific turn of the intersection.

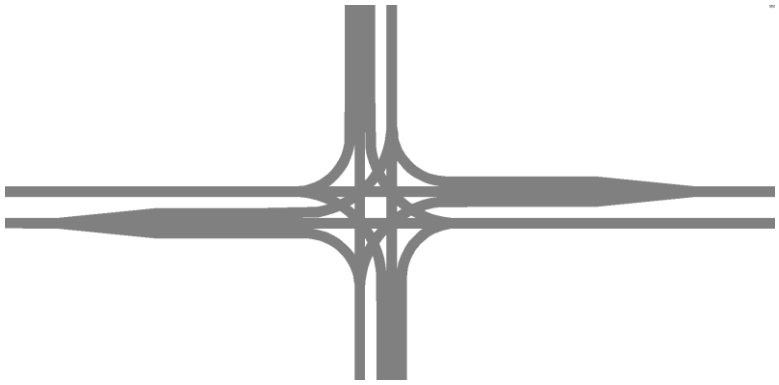


Figure A.35: *Dynamically updated version of intersection 2*

Step 6.4: Update network with VISSIM results (nwc_f07_update_network)

After the volumes have been updated, the network is updated using these results in the same manner as described in step 3. When the new number of lanes and the signal schemes for each intersection has been obtained, the network is again reconstructed, as described in step 5 (subnetworks are not changed). Reason that the full network is reconstructed instead of only updated is because unused links have been removed and this would result in empty and unused entries in the links and node matrices, resulting in inconsistencies. Downside of this method is that the link and connectors before and after the update have different values and cannot be related to each another anymore.

When the network has been updated, step 6.1 and step 6.2 are repeated, again obtaining paths in the new network and setting segment lengths. After these steps, all files used in VISSIM and the corresponding data structure created in Matlab are written to a separate folder, finalizing the network creation algorithm.

Appendix B: Simulation results

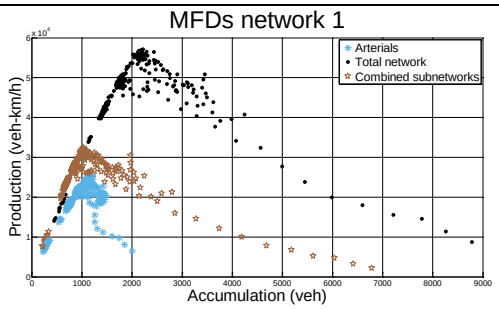


Figure B.1: *Production vs. Accumulation network 1*

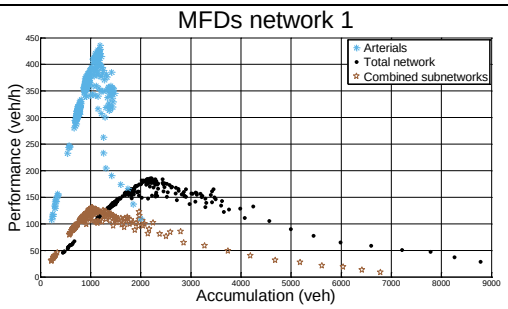


Figure B.2: *Performance vs. Accumulation network 1*

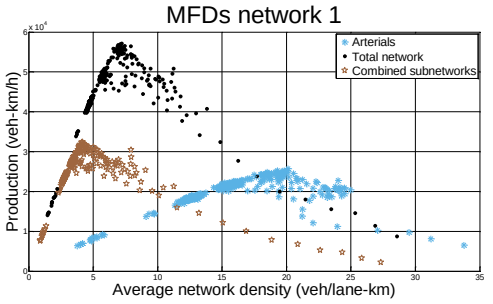


Figure B.3: *Production vs. Average network density network 1*

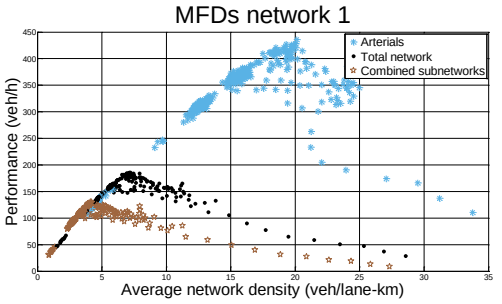


Figure B.4: *Performance vs. Average network density network 1*

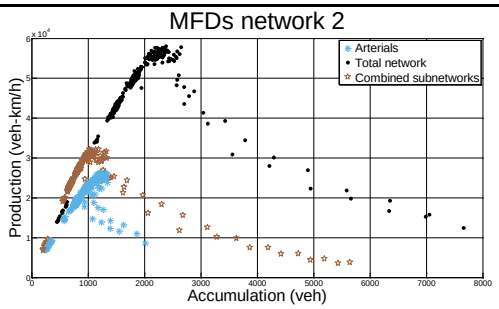


Figure B.5: *Production vs. Accumulation network 2*

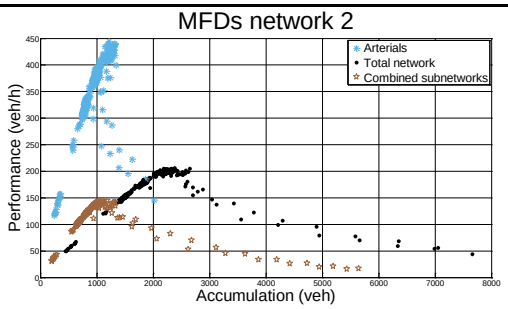


Figure B.6: *Performance vs. Accumulation network 2*

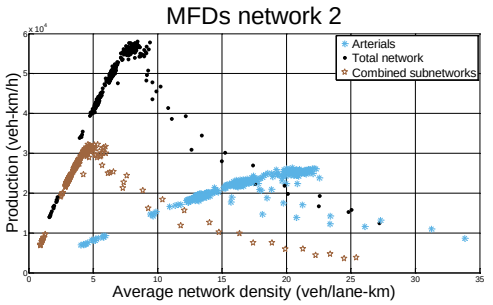


Figure B.7: *Production vs. Average network density network 2*

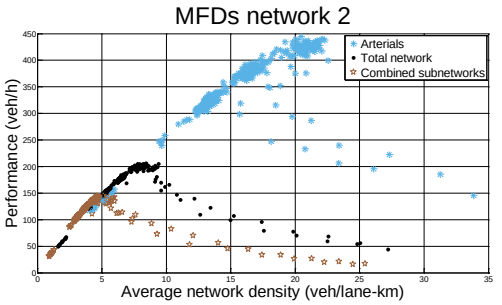


Figure B.8: *Performance vs. Average network density network 2*

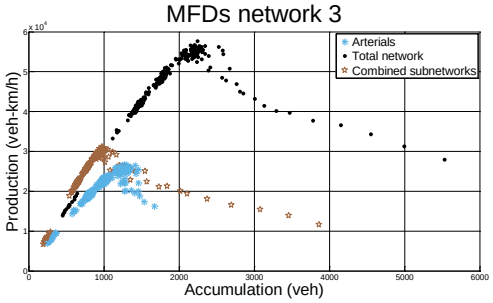


Figure B.9: *Production vs. Accumulation network 3*

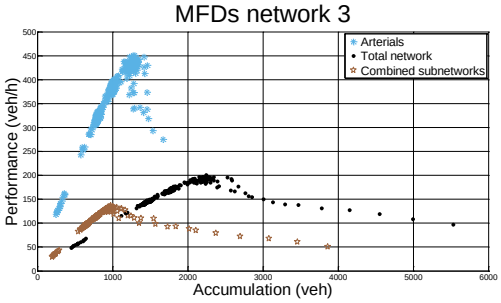


Figure B.10: *Performance vs. Accumulation network 3*

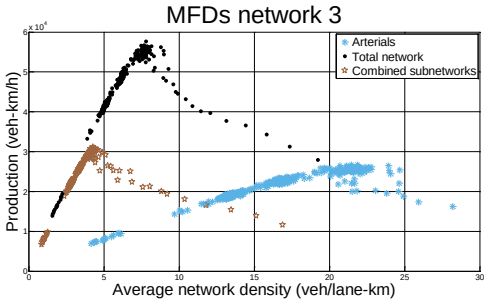


Figure B.11: *Production vs. Average network density network 3*

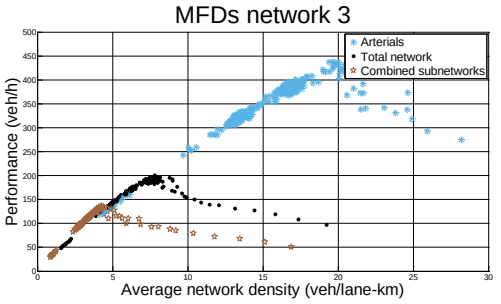


Figure B.12: *Performance vs. Average network density network 3*

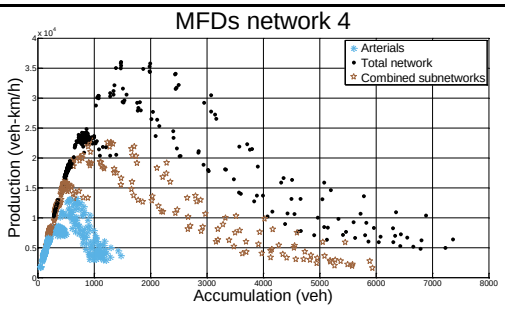


Figure B.13: *Production vs. Accumulation network 4*

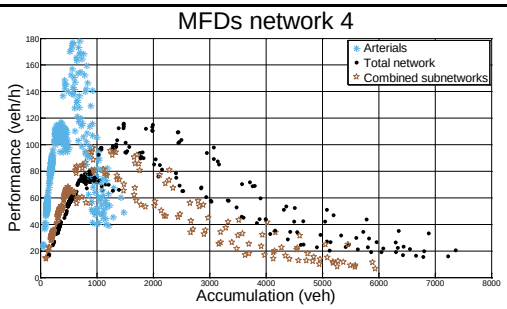


Figure B.14: *Performance vs. Accumulation network 4*

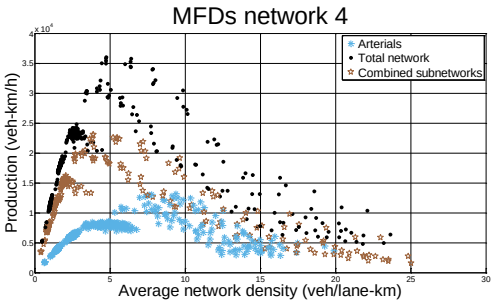


Figure B.15: *Production vs. Average network density network 4*

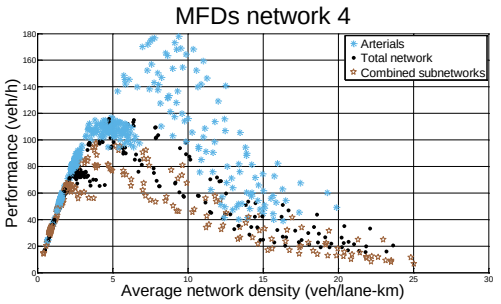


Figure B.16: *Performance vs. Average network density network 4*

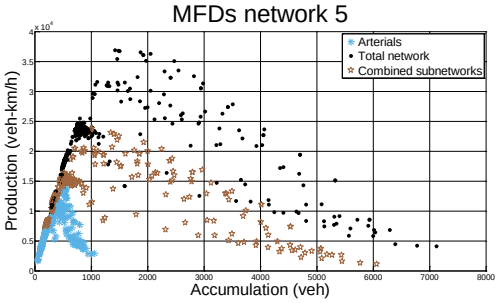


Figure B.17: *Production vs. Accumulation network 5*

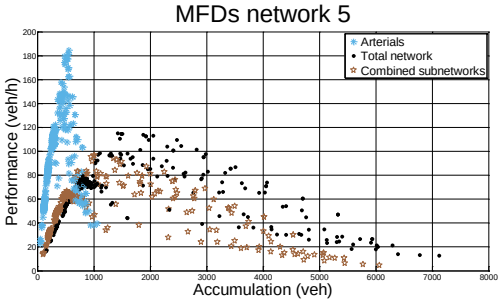


Figure B.18: *Performance vs. Accumulation network 5*

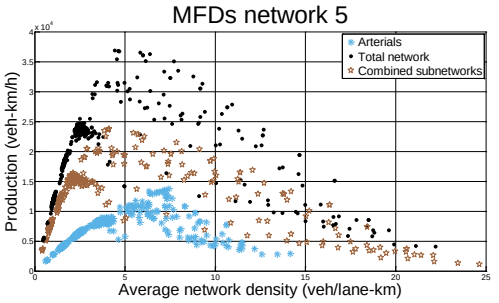


Figure B.19: *Production vs. Average network density network 5*

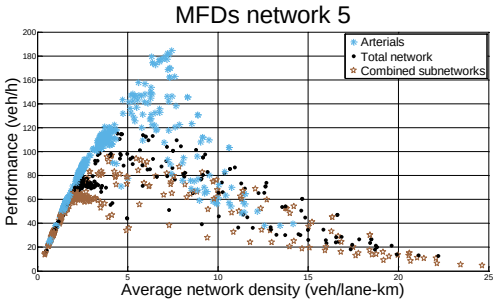


Figure B.20: *Performance vs. Average network density network 5*

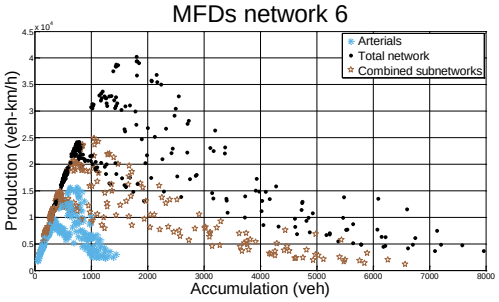


Figure B.21: *Production vs. Accumulation network 6*

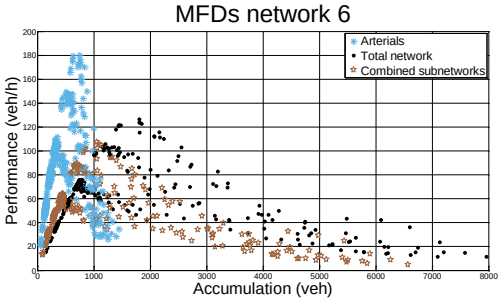


Figure B.22: *Performance vs. Accumulation network 6*

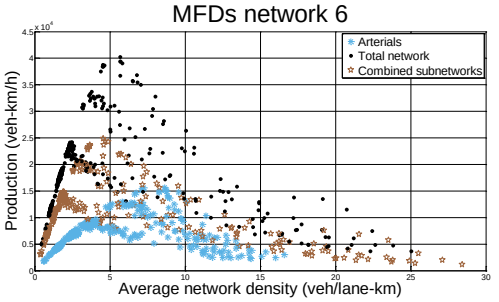


Figure B.23: *Production vs. Average network density network 6*

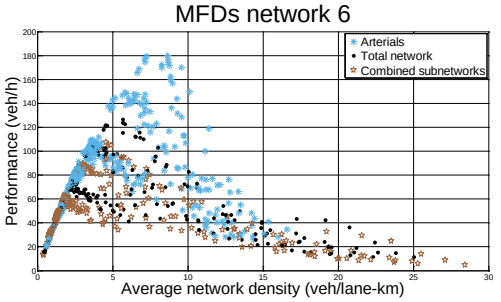


Figure B.24: *Performance vs. Average network density network 6*

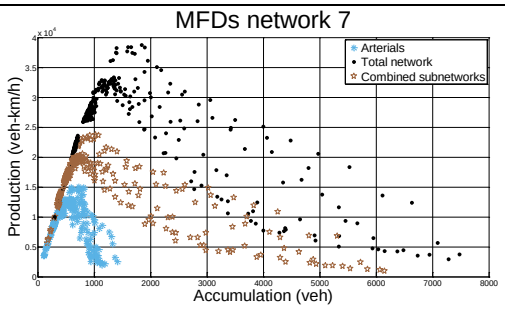


Figure B.25: *Production vs. Accumulation network 7*

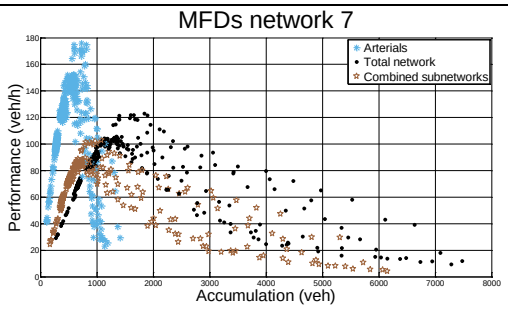


Figure B.26: *Performance vs. Accumulation network 7*

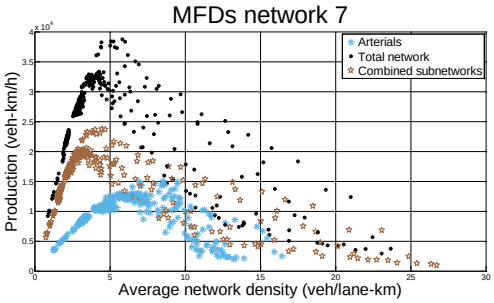


Figure B.27: *Production vs. Average network density network 7*

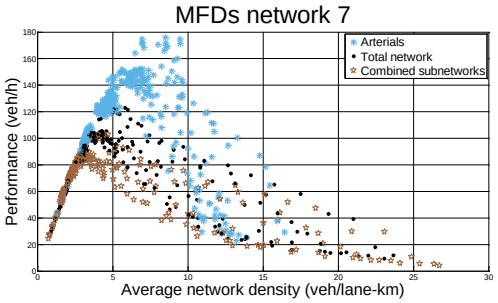


Figure B.28: *Performance vs. Average network density network 7*

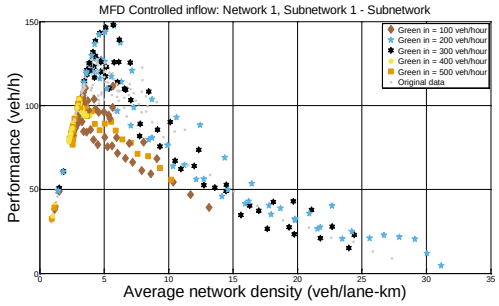


Figure B.29: *Controlled inflow, network 1 – subnetwork 1: subnetwork*

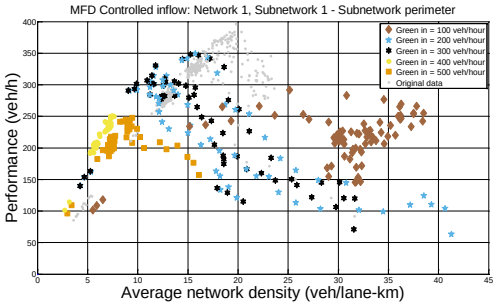


Figure B.30: *Controlled inflow, network 1 – subnetwork 1: perimeter*

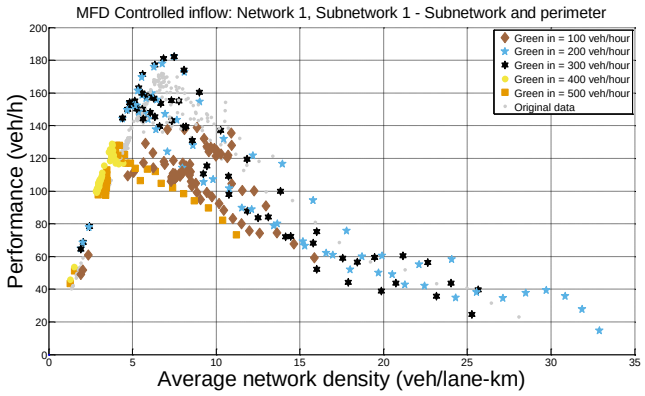


Figure B.31: *Controlled inflow, network 1 – subnetwork 1: subnetwork and perimeter*

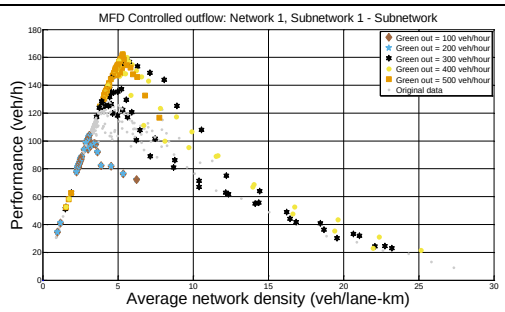


Figure B.32: *Controlled outflow, network 1 – subnetwork 1: subnetwork*

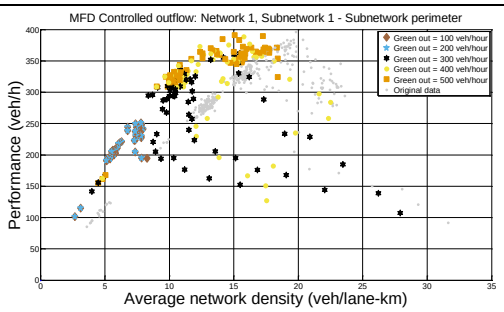


Figure B.33: *Controlled outflow, network 1 – subnetwork 1: perimeter*

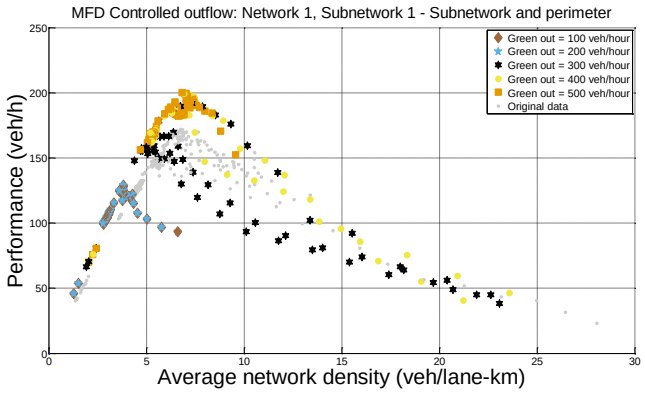


Figure B.34: *Controlled outflow, network 1 – subnetwork 1: subnetwork and perimeter*

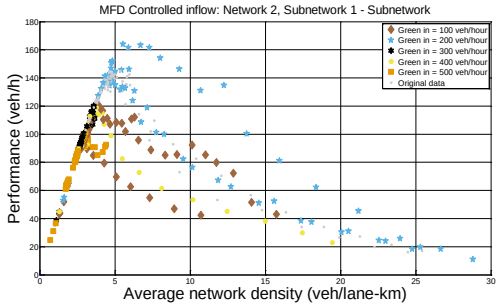


Figure B.35: *Controlled inflow, network 2 – subnetwork 1: subnetwork*

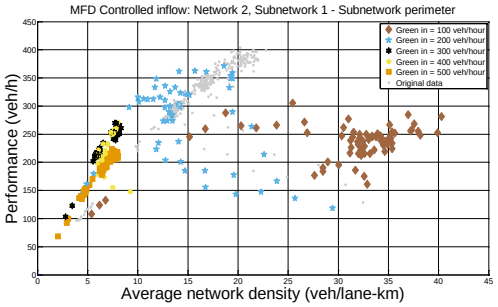


Figure B.36: *Controlled inflow, network 2 – subnetwork 1: perimeter*

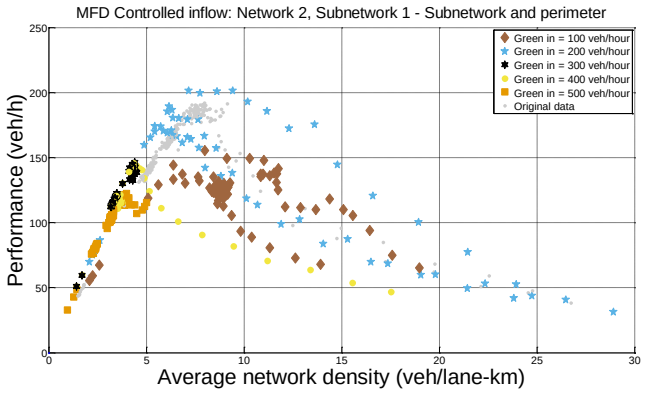


Figure B.37: *Controlled inflow, network 2 – subnetwork 1: subnetwork and perimeter*

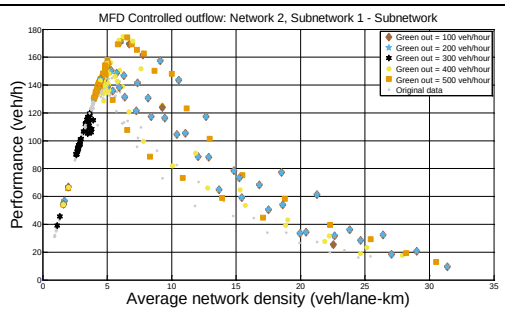


Figure B.38: *Controlled outflow, network 2 – subnetwork 1: subnetwork*

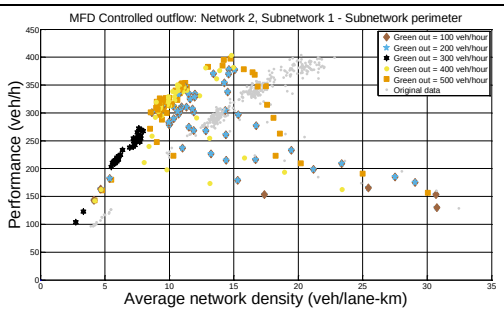


Figure B.39: *Controlled outflow, network 2 – subnetwork 1: perimeter*

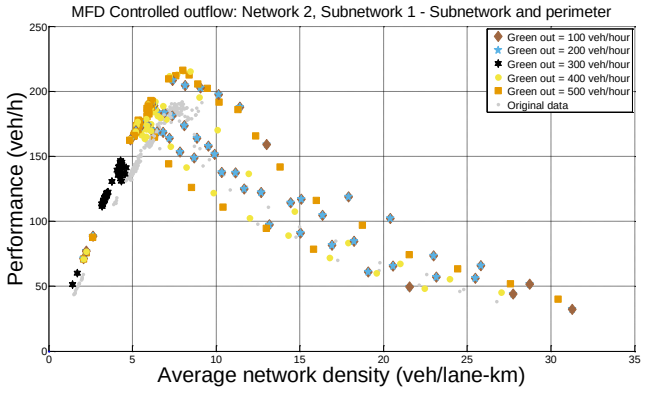


Figure B.40: *Controlled outflow, network 2 – subnetwork 1: subnetwork and perimeter*

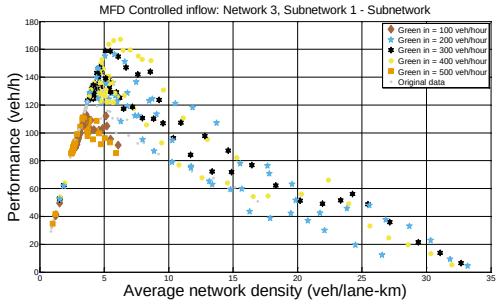


Figure B.41: *Controlled inflow, network 3 – subnetwork 1: subnetwork*

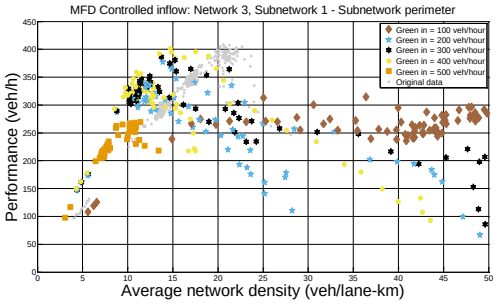


Figure B.42: *Controlled inflow, network 3 – subnetwork 1: perimeter*

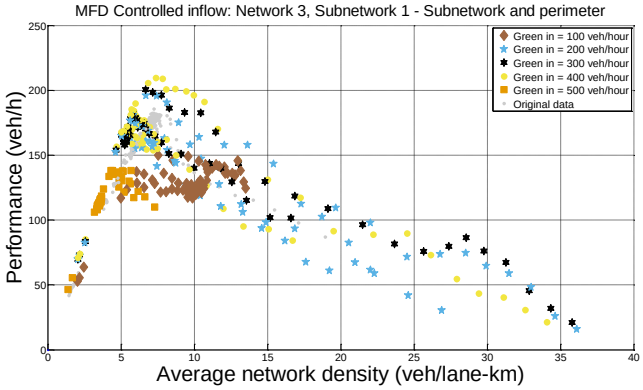


Figure B.43: *Controlled inflow, network 3 – subnetwork 1: subnetwork and perimeter*

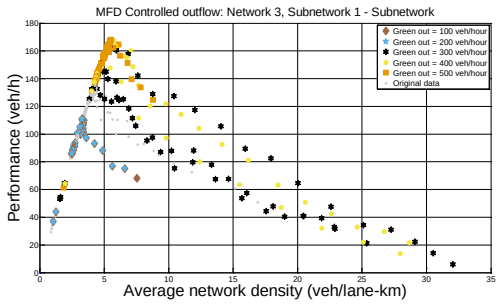


Figure B.44: *Controlled outflow, network 3 – subnetwork 1: subnetwork*

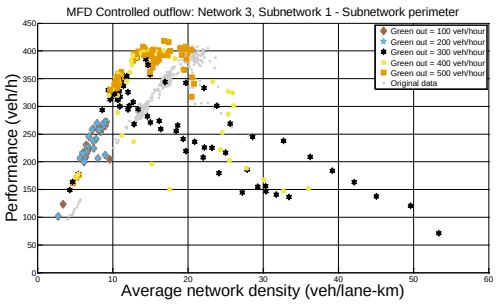


Figure B.45: *Controlled outflow, network 3 – subnetwork 1: perimeter*

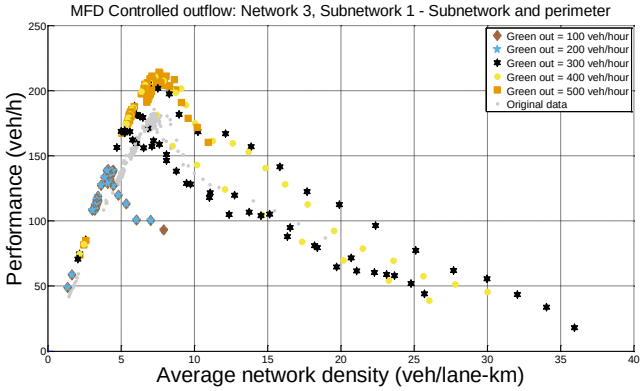


Figure B.46: *Controlled outflow, network 3 – subnetwork 1: subnetwork and perimeter*

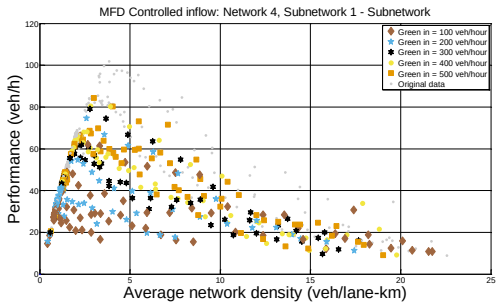


Figure B.47: *Controlled inflow, network 4 – subnetwork 1: subnetwork*

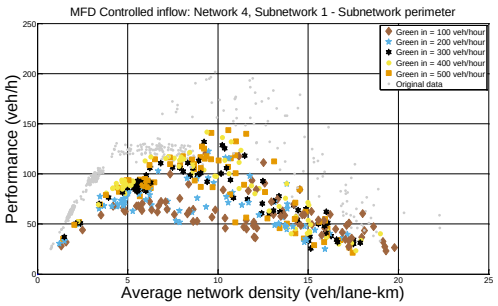


Figure B.48: *Controlled inflow, network 4 – subnetwork 1: perimeter*

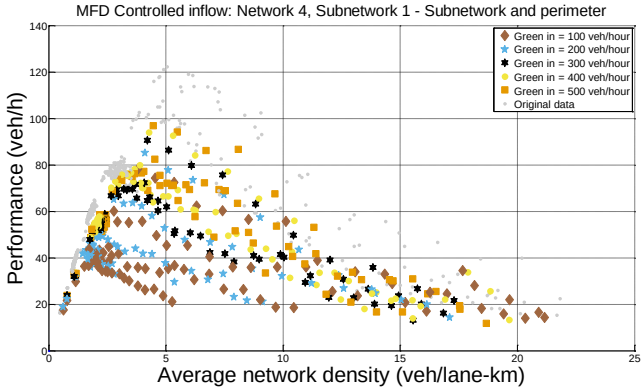


Figure B.49: *Controlled inflow, network 4 – subnetwork 1: subnetwork and perimeter*

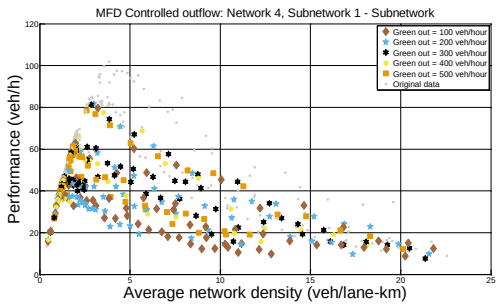


Figure B.50: *Controlled outflow, network 4 – subnetwork 1: subnetwork*

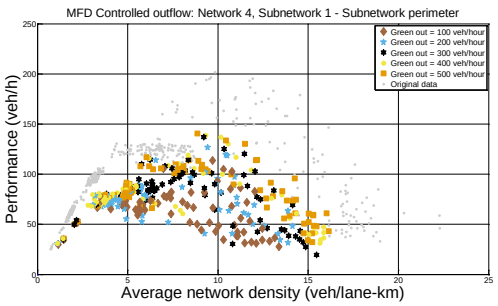


Figure B.51: *Controlled outflow, network 4 – subnetwork 1: perimeter*

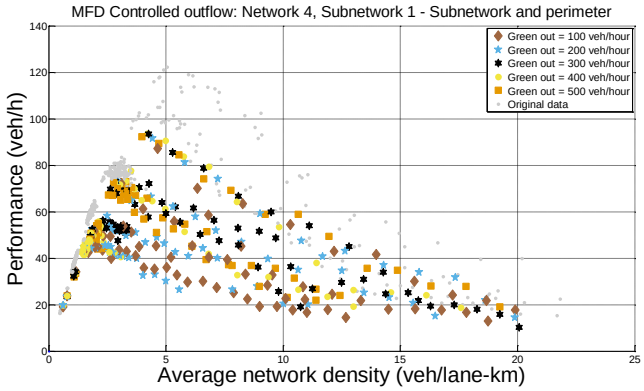


Figure B.52: *Controlled outflow, network 4 – subnetwork 1: subnetwork and perimeter*

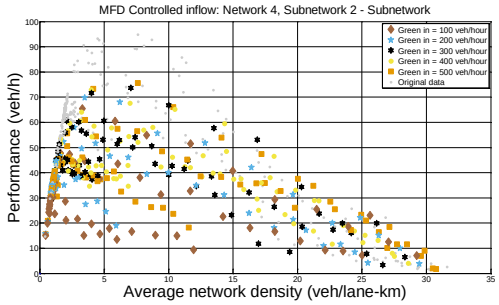


Figure B.53: *Controlled inflow, network 4 – subnetwork 2: subnetwork*

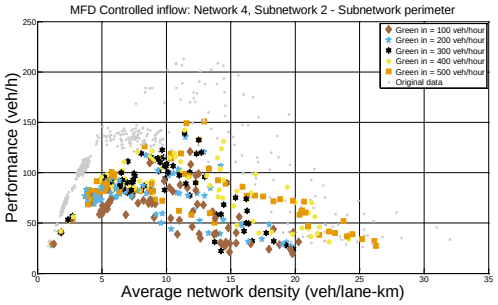


Figure B.54: *Controlled inflow, network 4 – subnetwork 2: perimeter*

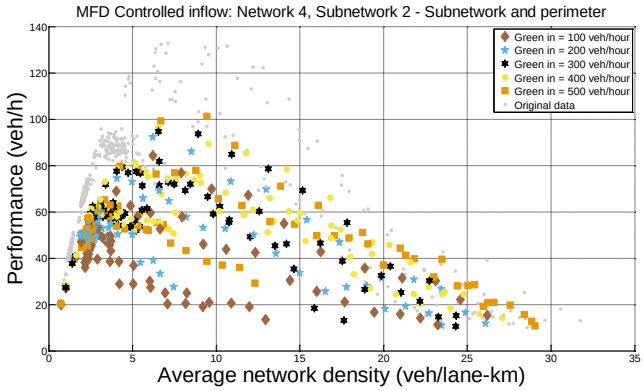


Figure B.55: *Controlled inflow, network 4 – subnetwork 2: subnetwork and perimeter*

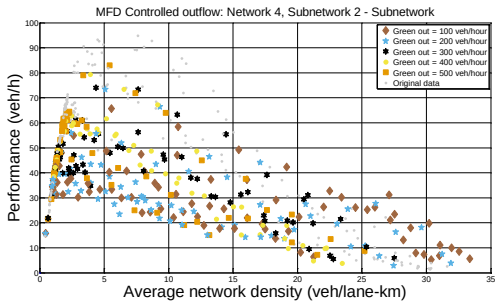


Figure B.56: *Controlled outflow, network 4 – subnetwork 2: subnetwork*

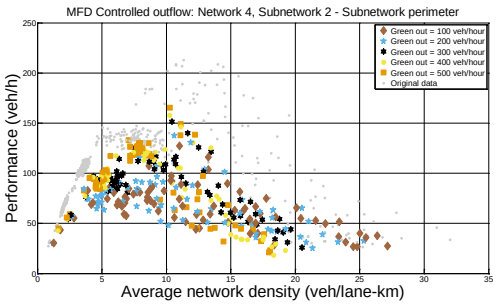


Figure B.57: *Controlled outflow, network 4 – subnetwork 2: perimeter*

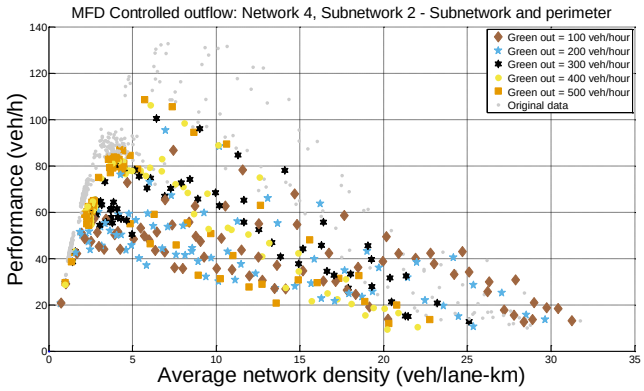


Figure B.58: *Controlled outflow, network 4 – subnetwork 2: subnetwork and perimeter*

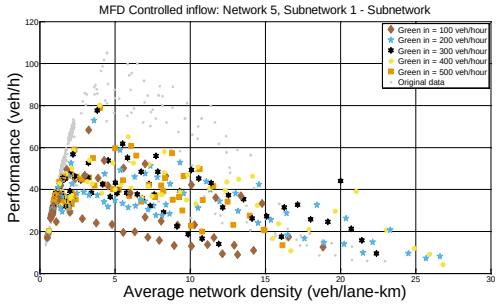


Figure B.59: *Controlled inflow, network 5 – subnetwork 1: subnetwork*

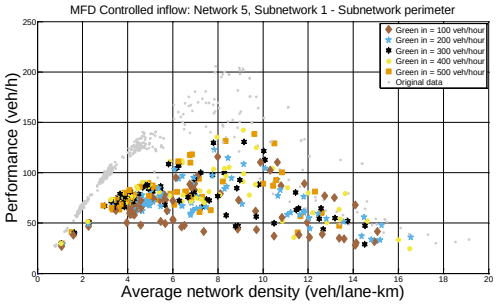


Figure B.60: *Controlled inflow, network 5 – subnetwork 1: perimeter*

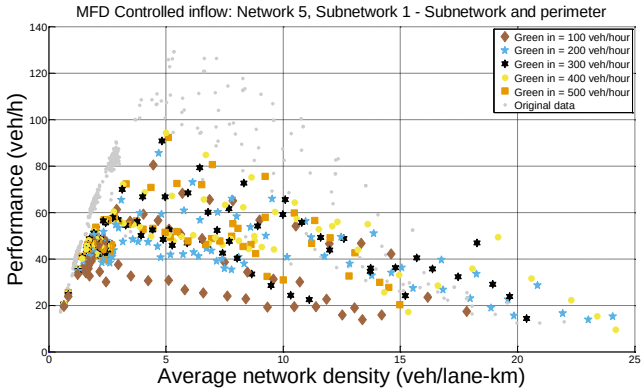


Figure B.61: *Controlled inflow, network 5 – subnetwork 1: subnetwork and perimeter*

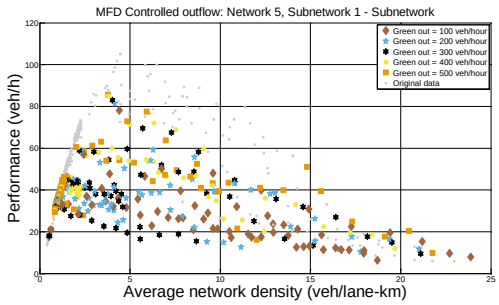


Figure B.62: *Controlled outflow, network 5 – subnetwork 1: subnetwork*

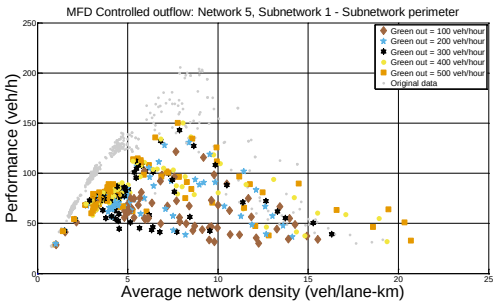


Figure B.63: *Controlled outflow, network 5 – subnetwork 1: perimeter*

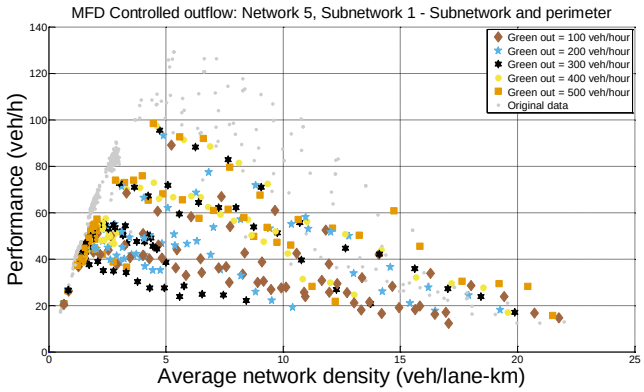


Figure B.64: *Controlled outflow, network 5 – subnetwork 1: subnetwork and perimeter*

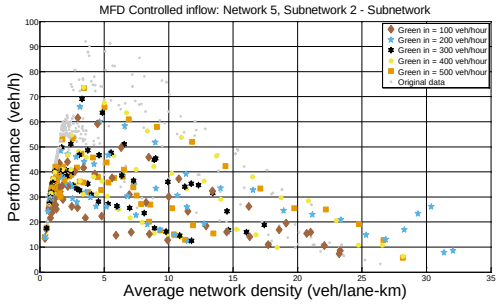


Figure B.65: *Controlled inflow, network 5 – subnetwork 2: subnetwork*

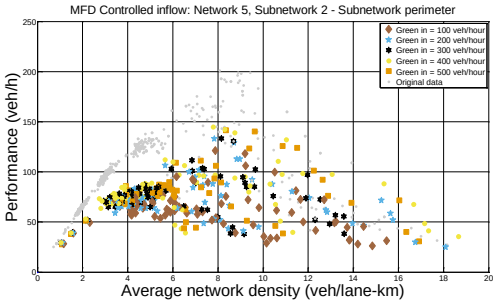


Figure B.66: *Controlled inflow, network 5 – subnetwork 2: perimeter*

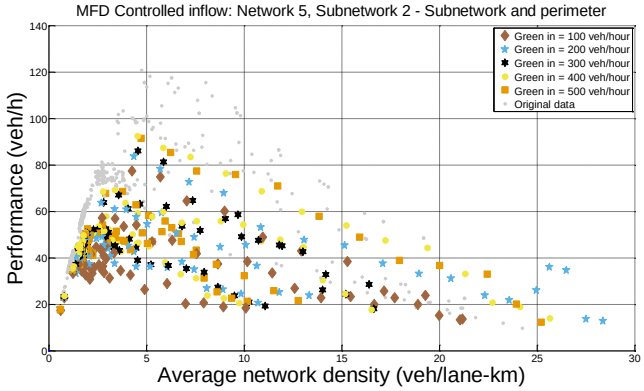


Figure B.67: *Controlled inflow, network 5 – subnetwork 2: subnetwork and perimeter*

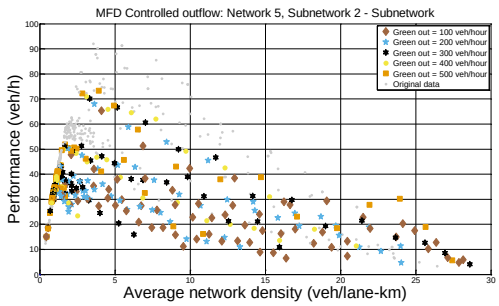


Figure B.68: *Controlled outflow, network 5 – subnetwork 2: subnetwork*

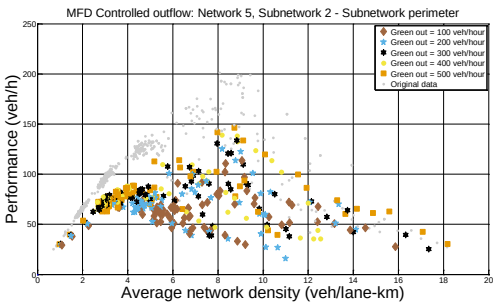


Figure B.69: *Controlled outflow, network 5 – subnetwork 2: perimeter*

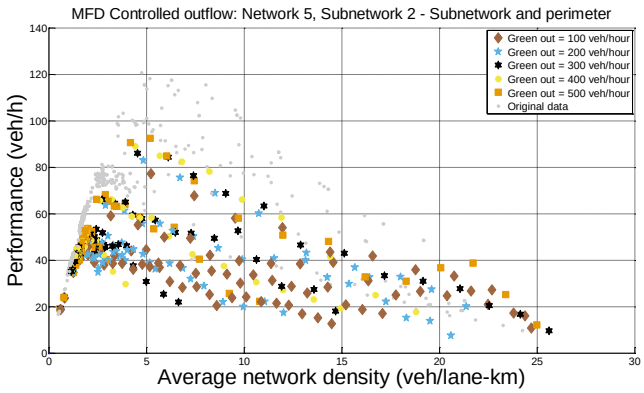


Figure B.70: *Controlled outflow, network 5 – subnetwork 2: subnetwork and perimeter*

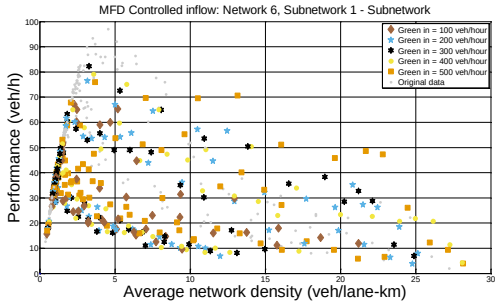


Figure B.71: *Controlled inflow, network 6 – subnetwork 1: subnetwork*

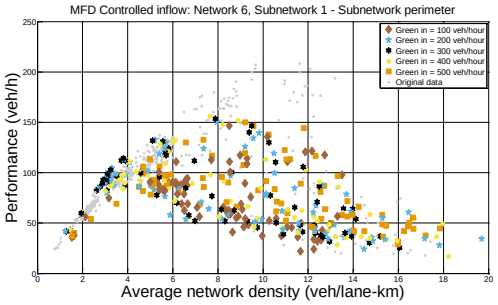


Figure B.72: *Controlled inflow, network 6 – subnetwork 1: perimeter*

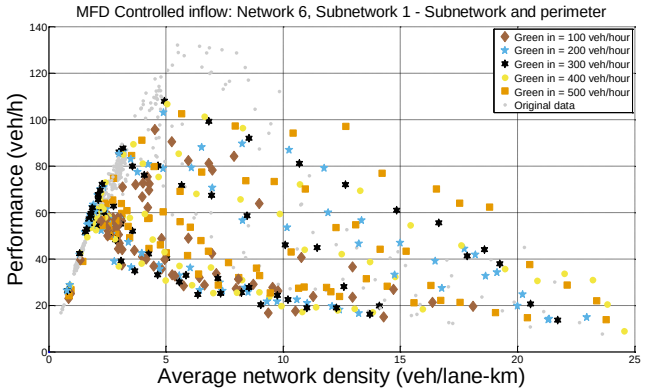


Figure B.73: *Controlled inflow, network 6 – subnetwork 1: subnetwork and perimeter*

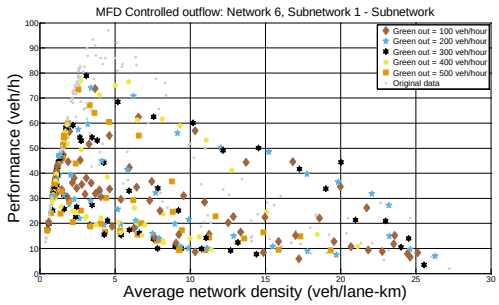


Figure B.74: *Controlled outflow, network 6 – subnetwork 1: subnetwork*

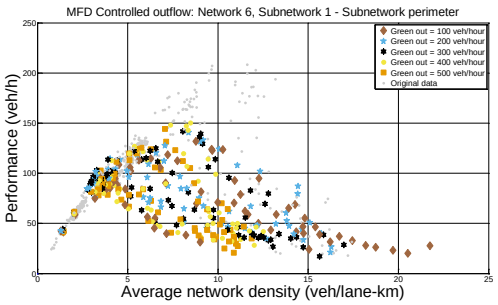


Figure B.75: *Controlled outflow, network 6 – subnetwork 1: perimeter*

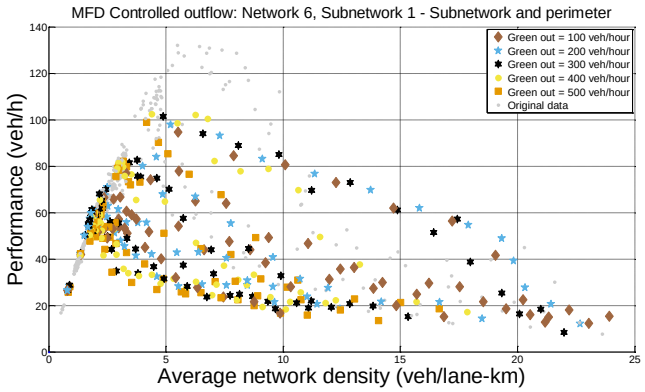


Figure B.76: *Controlled outflow, network 6 – subnetwork 1: subnetwork and perimeter*

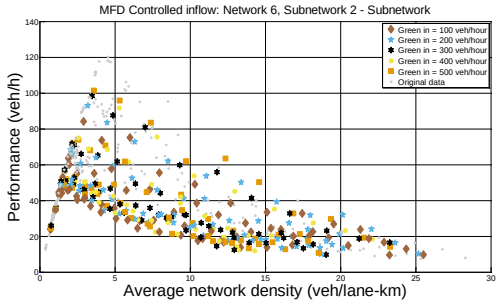


Figure B.77: *Controlled inflow, network 6 – subnetwork 2: subnetwork*

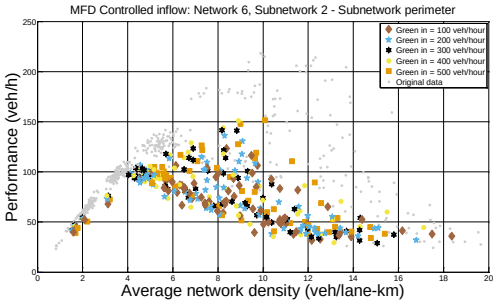


Figure B.78: *Controlled inflow, network 6 – subnetwork 2: perimeter*

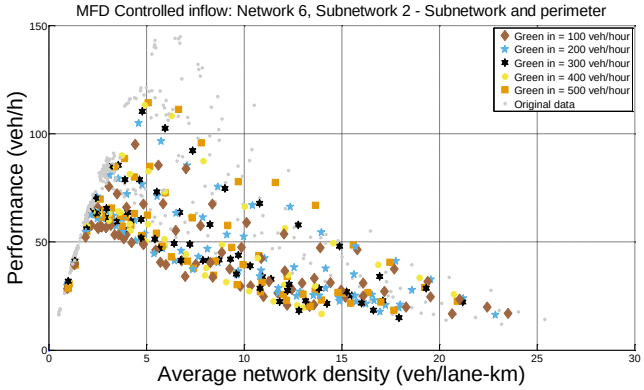


Figure B.79: *Controlled inflow, network 6 – subnetwork 2: subnetwork and perimeter*

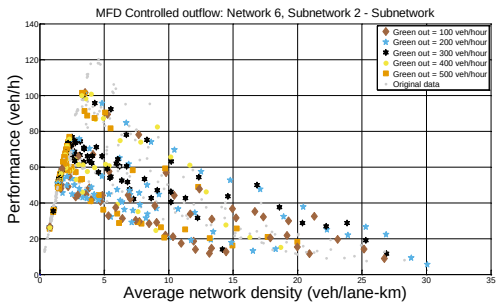


Figure B.80: *Controlled outflow, network 6 – subnetwork 2: subnetwork*

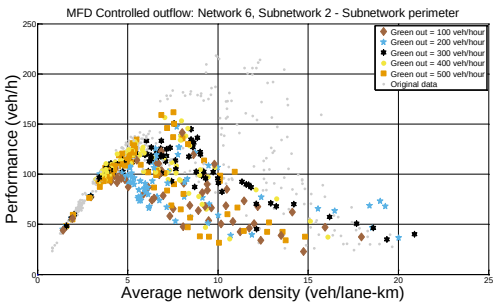


Figure B.81: *Controlled outflow, network 6 – subnetwork 2: perimeter*

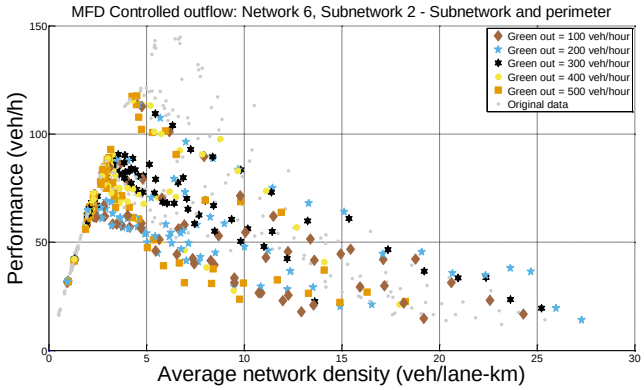


Figure B.82: *Controlled outflow, network 6 – subnetwork 2: subnetwork and perimeter*

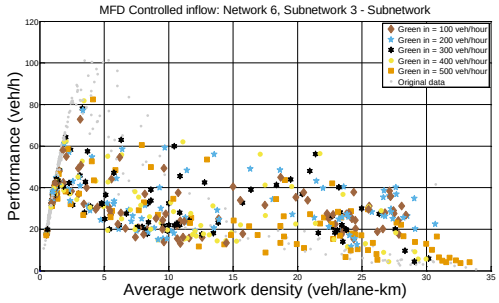


Figure B.83: *Controlled inflow, network 6 – subnetwork 3: subnetwork*

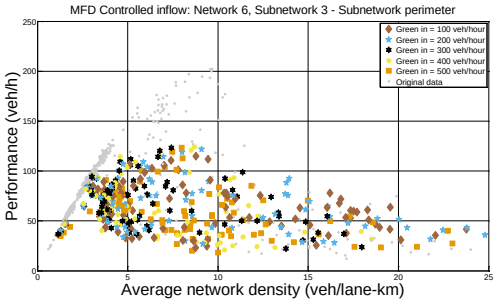


Figure B.84: *Controlled inflow, network 6 – subnetwork 3: perimeter*

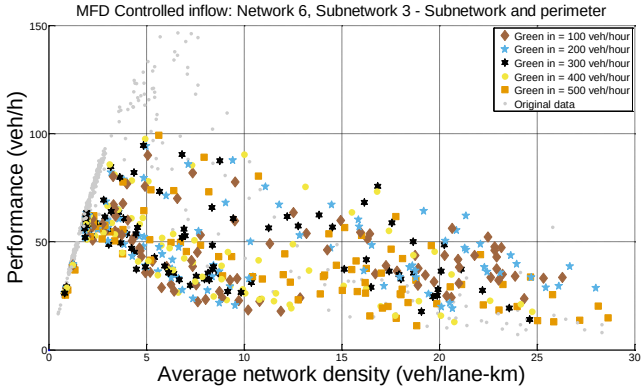


Figure B.85: *Controlled inflow, network 6 – subnetwork 3: subnetwork and perimeter*

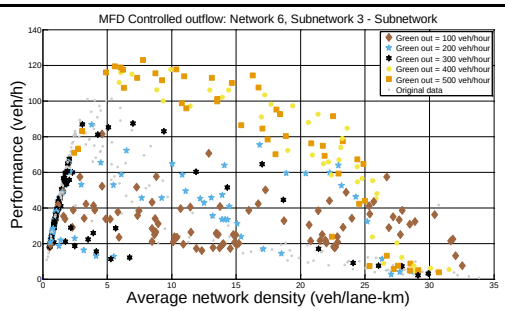


Figure B.86: *Controlled outflow, network 6 – subnetwork 3: subnetwork*

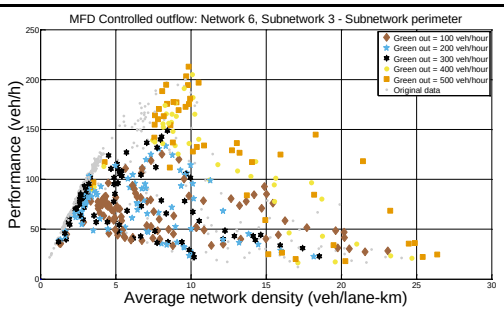


Figure B.87: *Controlled outflow, network 6 – subnetwork 3: perimeter*

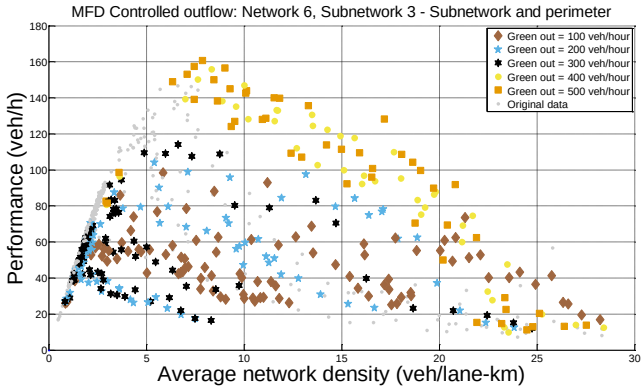


Figure B.88: *Controlled outflow, network 6 – subnetwork 3: subnetwork and perimeter*

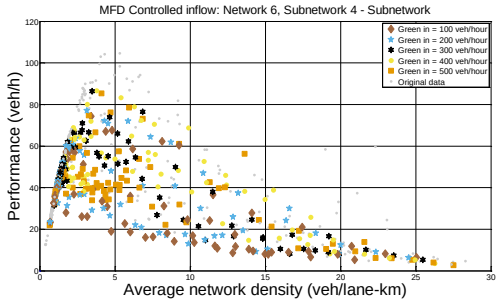


Figure B.89: *Controlled inflow, network 6 – subnetwork 4: subnetwork*

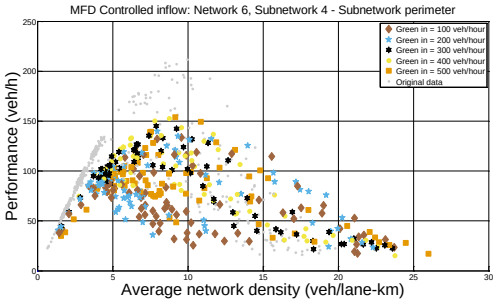


Figure B.90: *Controlled inflow, network 6 – subnetwork 4: perimeter*

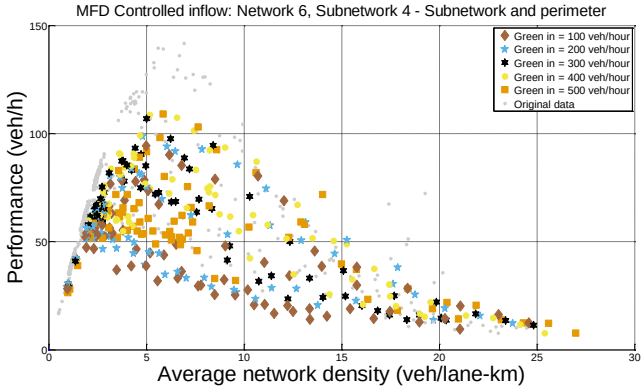


Figure B.91: *Controlled inflow, network 6 – subnetwork 4: subnetwork and perimeter*

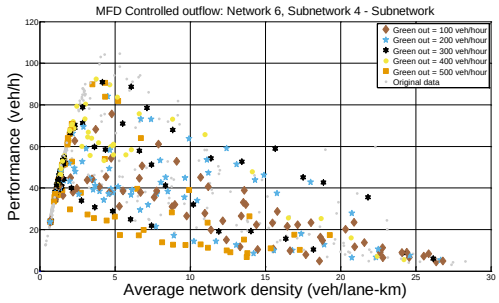


Figure B.92: *Controlled outflow, network 6 – subnetwork 4: subnetwork*

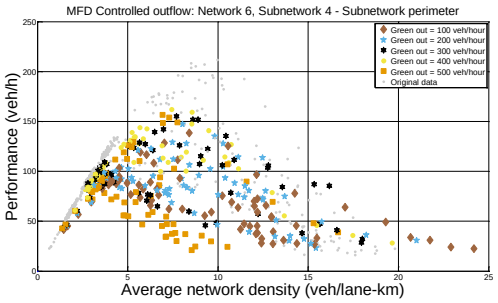


Figure B.93: *Controlled outflow, network 6 – subnetwork 4: perimeter*

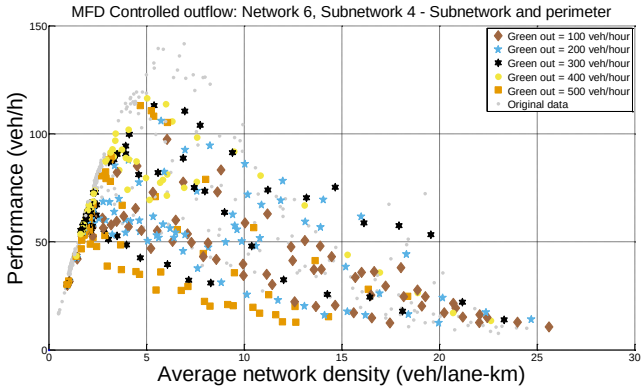


Figure B.94: *Controlled outflow, network 6 – subnetwork 4: subnetwork and perimeter*

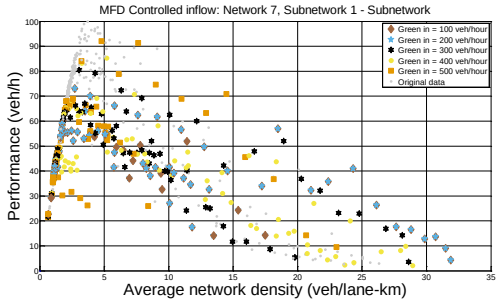


Figure B.95: *Controlled inflow, network 7 – subnetwork 1: subnetwork*

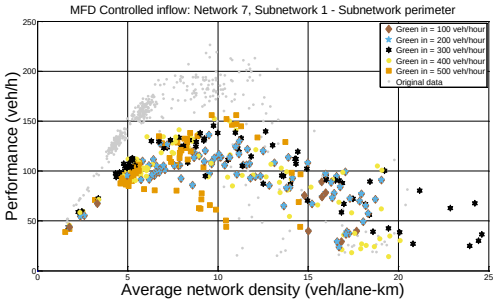


Figure B.96: *Controlled inflow, network 7 – subnetwork 1: perimeter*

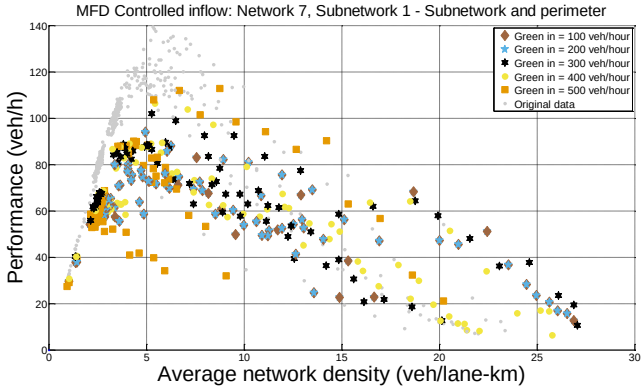


Figure B.97: *Controlled inflow, network 7 – subnetwork 1: subnetwork and perimeter*

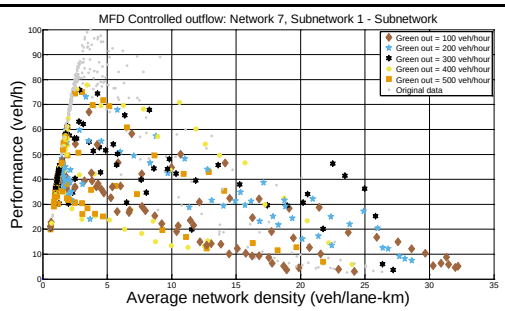


Figure B.98: *Controlled outflow, network 7 – subnetwork 1: subnetwork*

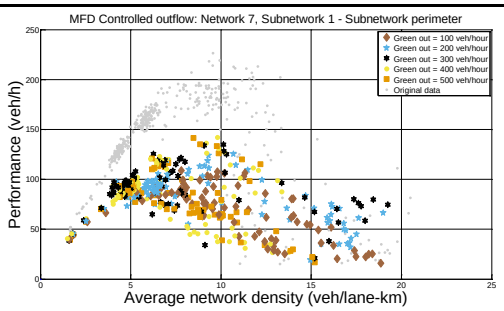


Figure B.99: *Controlled outflow, network 7 – subnetwork 1: perimeter*

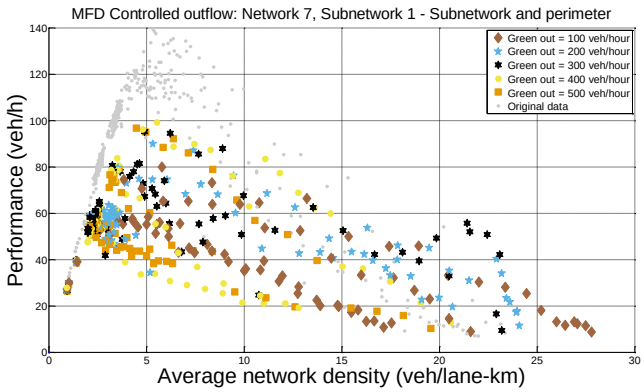


Figure B.100: *Controlled outflow, network 7 – subnetwork 1: subnetwork and perimeter*

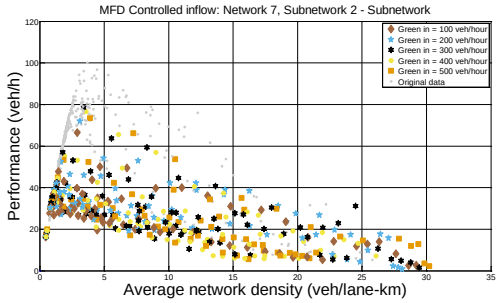


Figure B.101: *Controlled inflow, network 7 – subnetwork 2: subnetwork*

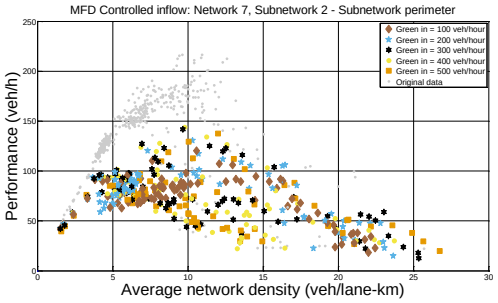


Figure B.102: *Controlled inflow, network 7 – subnetwork 2: perimeter*

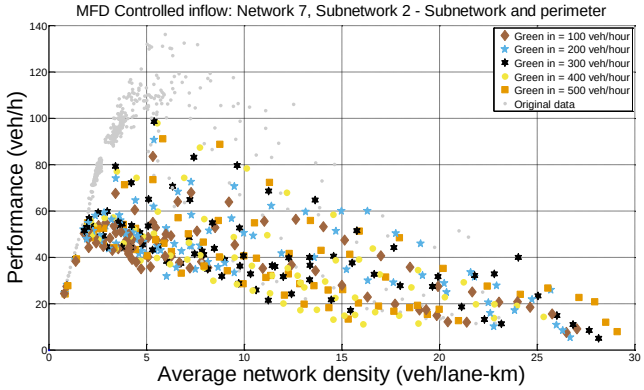


Figure B.103: *Controlled inflow, network 7 – subnetwork 2: subnetwork and perimeter*

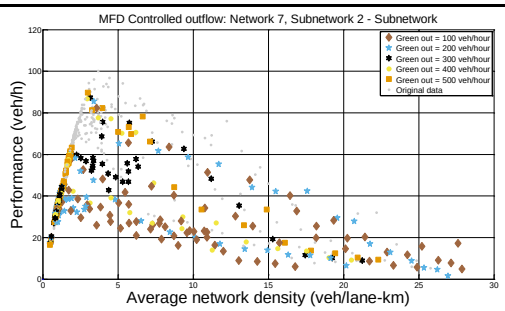


Figure B.104: *Controlled outflow, network 7 – subnetwork 2: subnetwork*

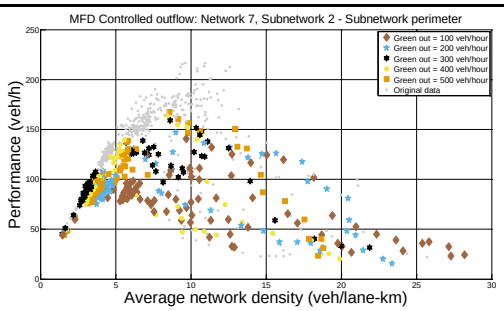


Figure B.105: *Controlled outflow, network 7 – subnetwork 2: perimeter*

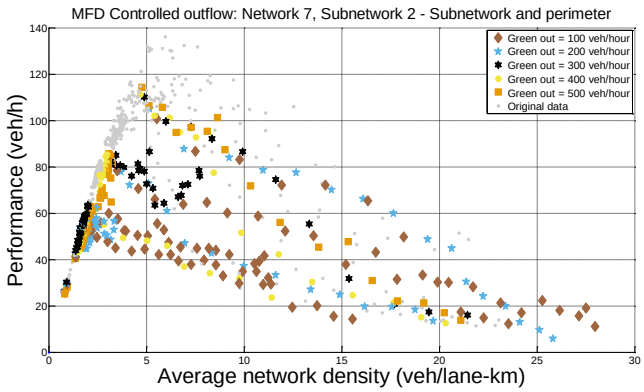


Figure B.106: *Controlled outflow, network 7 – subnetwork 2: subnetwork and perimeter*

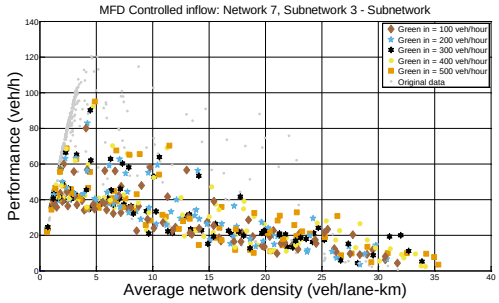


Figure B.107: *Controlled inflow, network 7 – subnetwork 3: subnetwork*

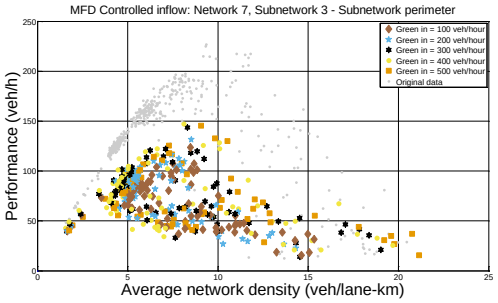


Figure B.108: *Controlled inflow, network 7 – subnetwork 3: perimeter*

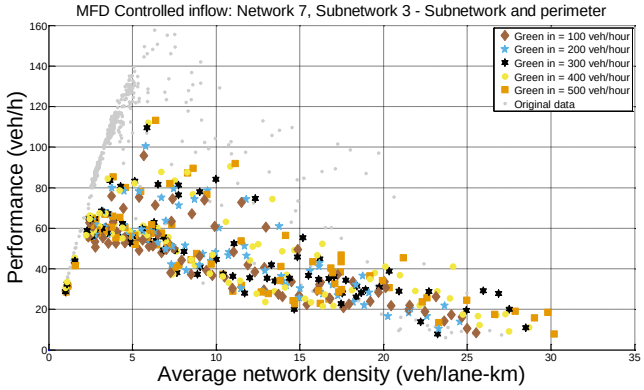


Figure B.109: *Controlled inflow, network 7 – subnetwork 3: subnetwork and perimeter*

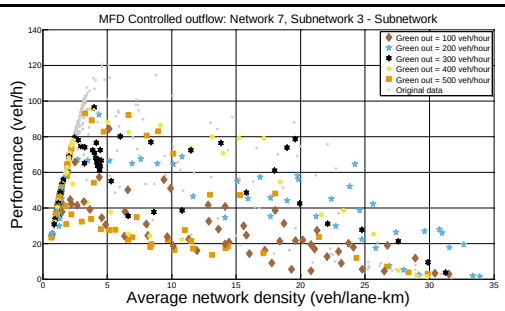


Figure B.110: *Controlled outflow, network 7 – subnetwork 3: subnetwork*

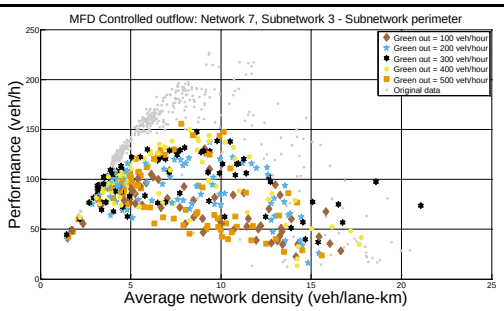


Figure B.111: *Controlled outflow, network 7 – subnetwork 3: perimeter*

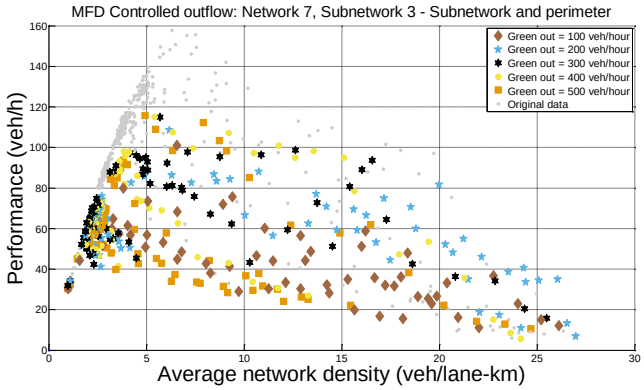


Figure B.112: *Controlled outflow, network 7 – subnetwork 3: subnetwork and perimeter*

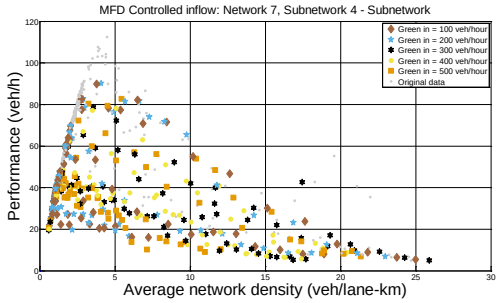


Figure B.113: *Controlled inflow, network 7 – subnetwork 4: subnetwork*

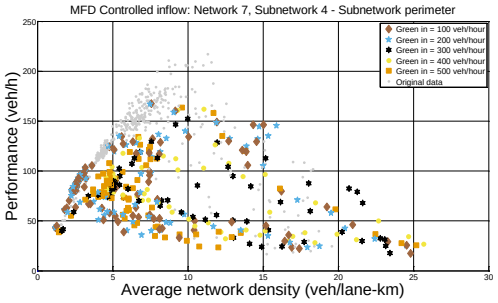


Figure B.114: *Controlled inflow, network 7 – subnetwork 4: perimeter*

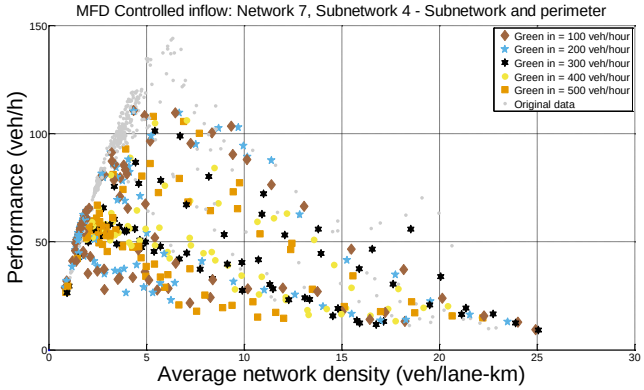


Figure B.115: *Controlled inflow, network 7 – subnetwork 4: subnetwork and perimeter*

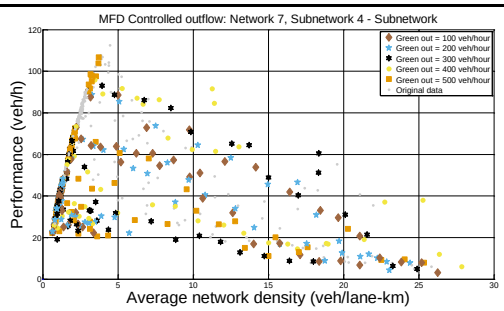


Figure B.116: *Controlled outflow, network 7 – subnetwork 4: subnetwork*

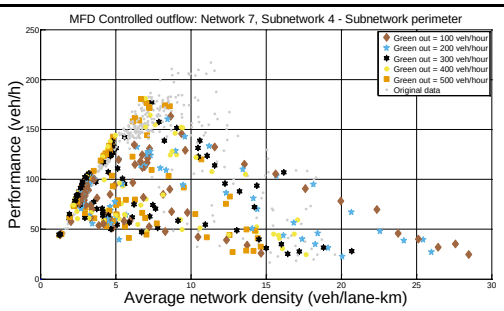


Figure B.117: *Controlled outflow, network 7 – subnetwork 4: perimeter*

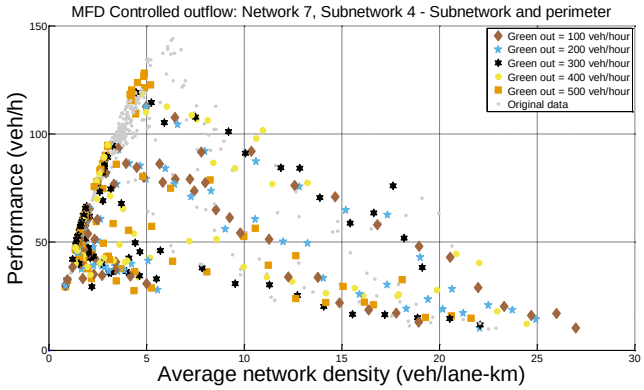


Figure B.118: *Controlled outflow, network 7 – subnetwork 4: subnetwork and perimeter*

Appendix C: Digital data

Appendix C.1:	Report
Appendix C.2:	Matlab model files
Appendix C.3:	VISSIM network files
Appendix C.4:	Simulation results: MFDs scenario 1
Appendix C.5:	Simulation results: MFDs scenario 2
Appendix C.6:	Simulation results: Final traffic states
Appendix C.7:	Simulation results: Data tables