An empirical estimation on the impact of freight traffic on the capacity drop

S.J. de Swart

TU Delft

21.04.2021

TNO innovation for life

An empirical estimation on the impact of freight traffic on the capacity drop

by

Sebastiaan Jacobus de Swart
4377443

May 11, 2020 - April 21, 2021

In partial fulfilment of the requirements for the degree of

Master of Science

in Transport & Planning

at the Delft University of Technology,
to be defended publicly on 21-04-2021

Assessment committee:

Dr. V.L. (Victor) Knoop (chair)	TU Delft
Dr. ir. S.C. (Simeon) Calvert	TU Delft
Dr. M. (Maaike) Snelder	TU Delft & TNO SUMS
Dr. ir. M.A. (Michiel) de Bok	TU Delft & Significance

Name: Sebas de Swart
T&P Thesis – CIE5060

An electronic version of this thesis is available at https://repository.tudelft.nl/

Key words:
Capacity drop; truck penetration; freight traffic; heavy vehicles

PREFACE

This thesis will finalize my Master programme Transport & Planning, specialization Transport Networks, at the faculty of Civil Engineering of the Delft University of Technology. The executed study was performed in association with the Dutch Organisation for Applied Scientific Research, TNO, at the Sustainable Urban Mobility and Safety department from 18 May 2020 till 21 April 2021.

While the world faced the greatest pandemic known in my existence, I was just about to start my graduation work. Luckily, I got the confidence from my supervisor at TNO, Maaike Snelder, to stay after my internship and execute my master thesis here. Although the study did not come close to my initial expected plan due to the change of subject and the lack of physical presence, I am very grateful for this opportunity and the support I received. New challenges arose, and my study topic seemed to be completely irrelevant due to the reduction in transportation movements. However, I am very happy with the final result and hope my study contributes to a better performance of roadway capacity and specifically gain insight into the relation between the capacity drop and heavy vehicles.

First, I would like to thank my daily supervisor Simeon for guiding me during the entire process. Not only did he provoke my thoughts with critical analysis, he also provided our meetings with energy and was concerned about my wellbeing. I would like to thank Maaike for providing me with this opportunity and bringing me in contact with the right people. When I asked Michiel to be part of my commission, I was intrigued by his enthusiasm. His perception of this research to put it in a broader perspective helped me a lot. I would like to thank Victor for his trust and help in setting up this research as well as for his input during our meetings. I would also like to thank Salil in helping me with discussing simulation approaches and setting up the simulation environment. I would also like to thank Wouter Schakel for his time and expertise in helping me with fixing the bug in OpenTrafficSim. I would also to thank the people at TNO for their open attitude and guidance.

Second, I would like to thank the people around me in supporting me during the process along. My colleagues at Albert Heijn, Stan & Romke for being one of my few physical contacts during this time and being flexible in taking over my shifts so I could focus on my research. My parents especially for their unconditional support in any form to the best of their abilities. As well as my girlfriend for listening to me talking about my research ceaselessly and supporting me unconditionally.

Thank you all!

“There have to be reasons that you get up in the morning and you want to live.
Why do you want to live? What's the point? What inspires you? What do you love about the future?
If the future does not include being out there among the stars and being a multi-planet species, I find that incredibly depressing.” – Elon Musk

Sebas de Swart
Delft, April 2021

EXECUTIVE SUMMARY

Introduction and Study Design

Introduction

Good transportation infrastructure is the backbone of economic welfare and is inseparably connected with the local economy of a country. Not only does it facilitate the possibility for the transportation of goods and commodities, but it also enables commuting and travelling of individuals. The number of transport movements is conversely often used as a gauge of the economy. The economic crisis in 2007 becomes clear from a traffic perspective when observing the decline in kilometres travelled by freight traffic. However, the recovery of the economy led to a fast-increasing number of kilometres travelled afterwards, and further increase is forecasted including the adverse consequences. Despite solutions to maximize roadway capacity, the drop of capacity after congestions sets in remains an active field of study and will be central to this research.

When a breakdown of traffic occurs, i.e. a drop in speed, the flow of the traffic decreases considerably. Once traffic does recover from the speed drop, it appears that the flow during queue discharge is lower than pre-queue flow. This restricts the recovery of the congestion. This phenomenon has been researched for years, but the main cause has not been identified yet. Many traffic flow related research appoint observed phenomena regarding traffic flow to the presence of freight traffic. However, the influence of heavy vehicles on traffic flow has received little attention, especially the relationship between the capacity drop and the presence of heavy vehicles.

New emerging technologies are becoming available rather soon with the production of electric semi-autonomous trucks. In addition, measures by local governments to reach zero-emission zones in cities could further stimulate this trend. How this will impact the change in physical and operational characteristics of heavy vehicles, and subsequently traffic flow, is yet to be discovered.

Study Design

This research aims to fill the gap by investigating the impact of freight traffic on the capacity drop. Although the number of heavy vehicles is relatively small compared to regular vehicles, their impact is prominent, particularly during busy traffic conditions. The following research questions have been established based on the problem(s) introduced:

Research Question 1: **Which variables influence the capacity drop rate on both strategic and operational levels?**

Research Question 2: **How and to what extend does freight traffic impact the capacity drop on highways currently?**

Research Question 3: **How do different vehicle heterogeneity compositions and, physical- and operational characteristics of freight traffic impact the capacity drop?**

This study is divided into three main research stages, such that each of the research questions is supported by a research stage. A research framework has been set-up and can be observed in Figure 1. Work packages have been added to clarify the research stages. Subsequently, activities were created to steer the research and

answer the sub-research questions. The arrows indicate the dependency between activities, while the labelled rectangles indicate in which chapter the information can be found.

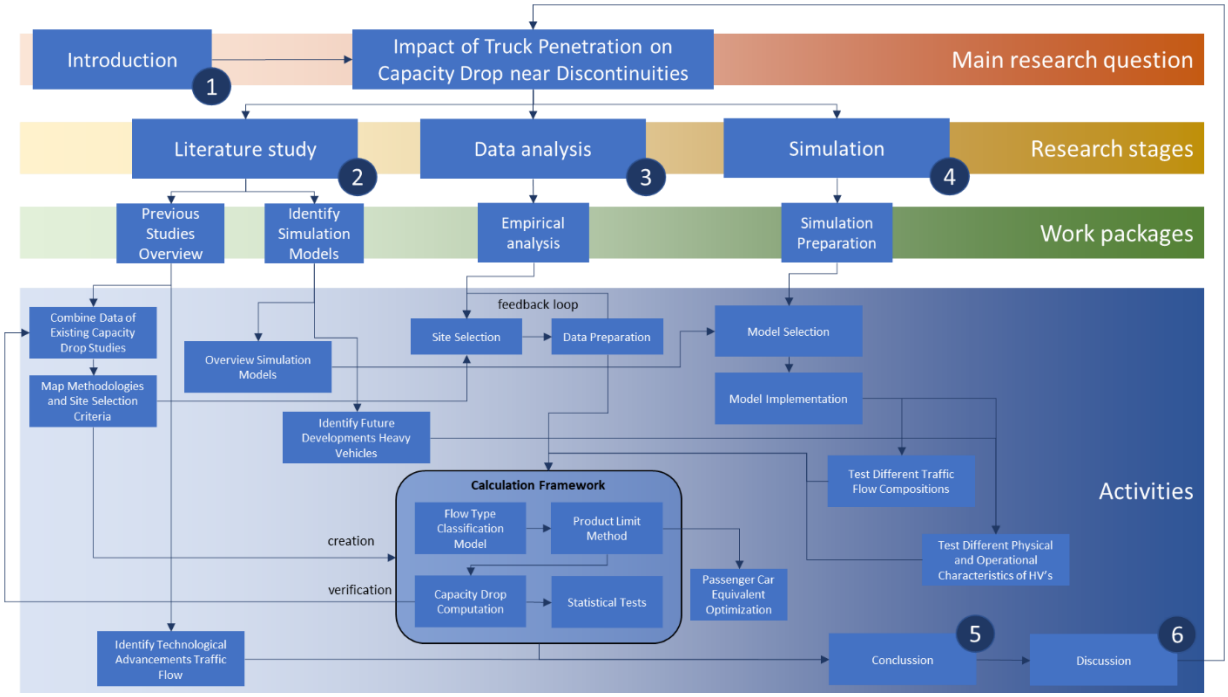


Figure 1 - Research Framework

Empirical Analysis

Methodology

To study the impact of freight traffic on the capacity drop empirically, collected data by loop-detectors along the Dutch highways was used. It was decided to use macroscopic data, as it turned out that microscopic data is scarce, especially for the requirements of this study. Several computational methods have been investigated for the estimation of the capacity drop. The Kaplan-Meier approach builds upon the stochastic nature of capacity and estimates the survival function from lifetime data and, therefore, fits the aim of the study best. Also, previous studies shown this methodology delivers appropriate results for German highways.

Eleven isolated bottlenecks were selected based on geometric, traffic and control properties on Dutch highways. The main criterion was that the sites had a wide range of heavy vehicle shares, so that the correlation between heavy vehicle share and capacity drop could be investigated. A bottleneck detector and a downstream detector were used to measure speed and flow, and to check for spillback of downstream bottlenecks. The effect of external impacts such as accidents and driver population were minimized in the data preparation so that these effects do not superimpose the effects of heavy vehicles. Furthermore, the raw measurement data was smoothed with a 3-minute moving average technique.

Next, the Flow Type Classification Model has been developed. This model identifies all breakdown and recovery events for the selected sites of the chosen study period, ranging between 4 and 8 years. The maximum flow values pre-queue were used as input data for the estimation of capacity distribution functions. Similarly, flow values at the moment of recovery were used as input data to estimate the recovery distribution functions. A Weibull distributions was fitted onto the non-parametric distributions as this turned out to be the best fit. Distribution curves were estimated for the entire study period as well as on a yearly basis to investigate how the capacity drop and heavy vehicle share changes from year to year. The capacity drop was estimated as the difference between the median of the breakdown distribution and median of the recovery distribution. An investigation into Passenger Car Equivalency values was also executed.

Results

The estimated capacity distribution functions show an increasing mean and standard deviation as the number of lanes increases for the selected sites. The estimated distribution functions of both the breakdown and recovery of only the 2-lane sites can be observed in Figure 2 and Figure 3 respectively. The observed breakdown means range between 4531 veh/h and 5272 veh/h with standard deviations ranging between 256 veh/h to 510 veh/h. The recovery distribution means range between 3428 veh/h and 4091 veh/h with standard deviations ranging between 237 veh/h to 326 veh/h. The estimated capacity drop for all sites ranged between 14% at the A1 up to 29.9% at both the A4 and A27. The share of heavy vehicles at the sites during breakdown ranged between an average 2% at the A6 up to 9% at the A15.

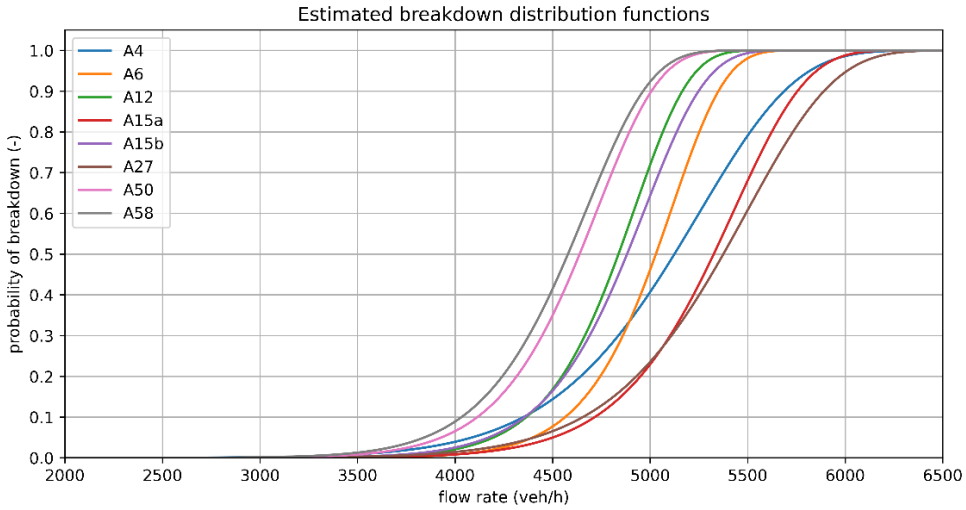


Figure 2 - Estimated breakdown distribution functions (only 2-lane sites)

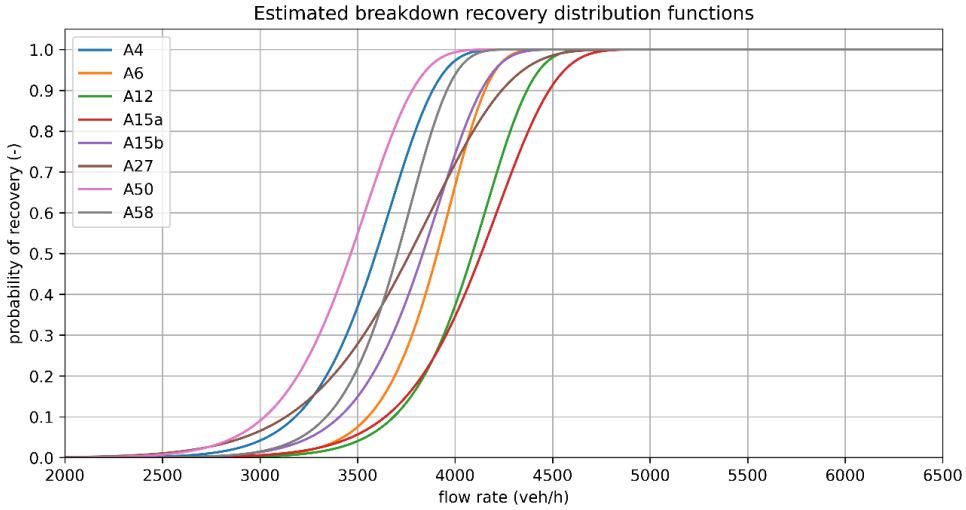


Figure 3 - Estimated recovery distribution functions (only 2-lane sites)

The yearly estimated distributions revealed an increase in heavy vehicle share each year for the A12, A15 and A50. The increase of heavy vehicles was about 10 to 14 veh/h per year during breakdown. This went hand in hand with a decline of capacity between 28 to 90 veh/h per year. These trends showed strong significance with a minimum R-squared of 0.73.

The heavy vehicle share – capacity drop correlation was tested both intra-location with aggregated distribution functions and intra-day by only considering corresponding breakdown-recovery events. The Pearson correlation test was used to examine the linear relationship between both variables, while the

Spearman correlation test assesses whether the two variables can be described using a monotonic relationship function. The intra-location results showed a wide variety of aggregated heavy vehicle share – capacity drop measurements. No significant relationship could be yielded. The intra-day results showed only significance for the A1, A15a, A15c and A50 with correlation factors between -0.16 and -0.48 for the Pearson test. However, a clear relationship was yielded between the pre-queue flow breakdown intensity and the capacity drop variables. Despite the wide variance in observed breakdown flows, the recovery flow was rather constant and predictable.

Simulation

Methodology

The OpenTrafficSim lane change simulation model was used to test the effect on the capacity drop of different heavy vehicle penetrations on a 2-lane site with identical geometric, traffic and control conditions. An example scenario was adjusted so that the calibration and validation of the model parameters was not necessary. Six different heavy vehicle penetration scenarios were simulated, ranging from 0 – 40%, each containing 20 runs. A loop detector was placed in the bottleneck to collect data comparable as the empirical loop detector data. The output data of the simulation could therefore be used as input in the Calculation Framework, including the Flow Type Classification Model.

In addition to the six heavy vehicle scenarios simulated, a mix of adjusted physical and operational characteristics of heavy vehicles were simulated to investigate the impact of future emerging developments. Both a smaller size of heavy vehicles, in line with the expectation of local governments, and a higher acceleration capability due to electrification, were tested.

Results

The heavy vehicle share scenarios showed an increasing capacity drop, from 16.4% to 34%, until the heavy vehicle share was at 15% (S4). Hereafter, lower capacity drop values were measured, as can be observed in Figure 4. The 2.5%, 7.5%, 12.5%, 17.5%, 25%, 30% and 35% HV scenarios were added to increase the number of data points. The breakdown capacity decreased form 4272 veh/h to 2972 veh/h for an increased heavy vehicle share. The standard deviation increased from 198 veh/h to 715 veh/h up to the 15% HV scenario, whereafter it declined to 560 veh/h (S6). The recovery mean steadily decreases from 3575 veh/h to 2338 veh/h, while the standard deviation increased from 150 veh/h to 248 veh/h. The breakdown distributions show a lot of variance in correspondence to the empirical data, while the recovery distributions show a very constant, almost predictable pattern for increased heavy vehicle share scenarios.

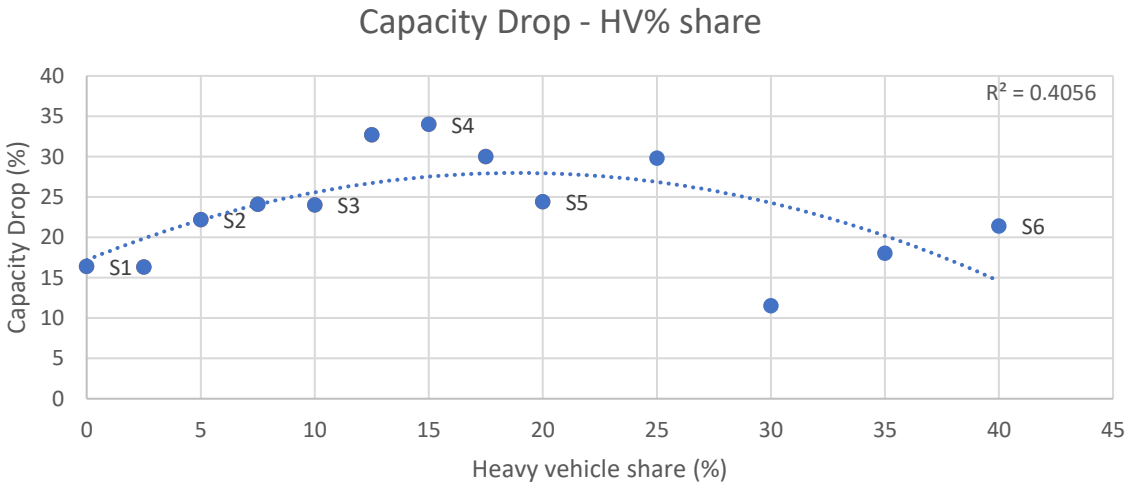


Figure 4 - Capacity drop (%) - Heavy vehicle share (%) correlation

The decreased truck size scenario resulted in a slightly higher mean of the breakdown distribution. The increased acceleration scenario resulted in a considerably higher mean of the recovery distribution.

A combination of both changed characteristics was simulated for the same heavy vehicle share scenarios as for the six base scenarios. The mean of the breakdown distributions remained the same, while the standard deviation declined compared to the base scenarios. The mean of the recovery distributions became slightly higher. The standard deviation became increasingly smaller, while, oppositely, the standard deviation of the recovery distributions of the six base scenarios became decreasingly larger.

Conclusion

Conclusion

The empirically observed capacity drop ranged between 14 – 29.9%, while the share of heavy vehicles at breakdown ranged between 2 – 9%. The empirical results seem to show a connection between the heavy vehicle share and the capacity drop, but this could not be statistically proven by both the Pearson and Spearman correlation test both intra-location and intra-day. Therefore, it can be concluded that the chaotic properties of the breakdown flow seem to superimpose the effects the share of heavy vehicles has on the capacity drop from an empirical perspective.

The simulation study aimed to fill this gap by equalizing external factors. The results showed that the estimated capacity drop tended to increase with an increasing amount of heavy vehicle share up to 15%. This reinforces the empirical evidence that a connection between heavy vehicle share and capacity drop exists. As the heavy vehicle share during breakdown only reaches up to 9% on roads currently, it is possible that the effect becomes more influential as the share of heavy vehicles grows further on the Dutch highways. Current empirical data does not show significant correlations, insights from the simulations show that this could be the case in the future as the percentage of heavy vehicle share increases continuously.

Future scenarios, with increased acceleration capabilities of vehicles and decreased truck sizes showed a decrease in the capacity drop. However, it also showed an increased standard deviation of the recovery distributions for higher heavy vehicles shares compared to the base scenarios. This shows that the estimation of the recovery flow shall become more difficult in the future.

Empirical data revealed that the number of heavy vehicles increased on the A12, A15 and A50 by about 10 to 14 veh/h per year during breakdown. Moreover, a decreasing capacity was observed at those sites of about 28 to 90 veh/h per year. The capacity decreased more than the increase in heavy vehicles, even after correction with the estimated PCE values. This raises concern about the underlying cause of the depreciation of roadway capacity. It is conjectured that new innovations such as ACC are responsible. However, this has not been verified. As the cause of this phenomenon has not been discovered yet, researchers should investigate this phenomenon because it could have a large impact on the Dutch roadway capacity.

Discussion

Reflection of the Research Methodology

The Flow Type Classification model currently has a one-size-fits-all approach. It appeared from flow-speed diagrams that the sites had different critical points regarding fundamental relationships. Detection of breakdown and recoveries by use of several breakdown and recovery speed threshold values could improve the classification. Furthermore, the requirements of a type 3 detector for the detection of spillback limits the number of available sites drastically. The occurrence of stop-and-go-waves also leads to some falsely classified breakdowns, mainly due to smoothing of the data with a lower temporal resolution (3 minutes) compared to other studies (5 minutes). However, the lower temporal resolution combined with the moving-average technique led to more accurate classified values, closer to the raw values.

Due to some simplifications, the simulation study must be put into perspective. The simulation model did only consist of two vehicle classes: passenger cars and heavy vehicles, while category 2 vehicles do lead to different traffic flow observations as illustrated in the empirical results of the A4 site. The lane utilization during congestion of OTS is known to be a weakness. It appeared that heavy vehicle had a very low incentive to change lanes, which was corrected by adjusting the VGAIN parameter. The other model parameters were not recalibrated after the change, because it was considered out of the scope of this research. Furthermore, the changed physical and operational characteristics scenarios are merely an investigation how these changes would affect traffic flow and specifically the capacity drop. The characteristics changes applied are not representative for actual future scenarios. Despite, a penetration rate of 100% was applied while studies have shown that mixed traffic does have negative impact on traffic flow.

Recommendations for Future Research

For this study, improvement of the Flow Type Classification Model is suggested. Any type of detector for the downstream measurement, increasing the site selection, replacing of the breakdown and recovery threshold constant with site-specific thresholds, as well as minimizing the impact of external factors would improve the reliability of the results.

The Passenger Car Equivalency value should be computed annually for the A12, A15 and A50. This would confirm whether the increase of heavy vehicles results in a more spatial impact on traffic flow due to for example lane flow underutilization. Besides, the declining capacity phenomenon should be investigated at other sites, on a large scale, to get insight into the observed phenomenon.

The lack of available microscopic data is hindering the investigation of the possible changed driving behaviour as the share of heavy vehicles increases. Therefore, more microscopic data should be collected in a similar way as the macroscopic loop detector data. Big data analysis would be helpful in investigation the traffic flow implications from a driving behaviour perspective and could provide monitoring of implemented infrastructural measurements.

Investigation into how physical and operational characteristics of traffic, and in particular freight traffic, will change the coming years, will help to improve future scenario estimation of simulation models. For the time being, it is clear that there will be a disruptive change to electrified vehicles, while the exact change of characteristics is hardly researched.

CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

PQF	Pre Queue Flow
QDF	Queue Discharge Flow
VPH	Vehicles Per Hour [veh/h]
NDW	(Dutch) National Database for Road Traffic Data
PCU	Passenger Car Unit
PCE	Passenger Car Equivalent
MLC	Mandatory Lane Changes
DLC	Discretionary Lane Changes
DBT	Driving Ban for Trucks
LFD	Lane Flow Distribution
VSL	Variable Speed Limit
PHL	Peak Hour Lane

1 INTRODUCTION

First, the research problem will be introduced in 1.1. Subsequently, a research objective and main research questions are formulated in 1.2. To answer these questions, a research methodology is proposed in 1.3, summarized into a research framework. Lastly, the scientific and societal relevance is highlighted in 1.4.

1.1 Problem Statement

Good transportation infrastructure is the backbone of economic welfare and inseparably connected with the local economy of a country. Not only does it facilitate the possibility for the transportation of goods and commodities, but it also enables commuting and travelling of individuals. The number of transport movements is conversely often used as a gauge of the economy. This may very well be understood when looking at the impact of the economic crisis starting in 2007, which stagnated the kilometres travelled on the Dutch road network as can be observed in Figure 5. The number of heavy vehicle kilometres travelled fell back to the level of the beginning of this century.

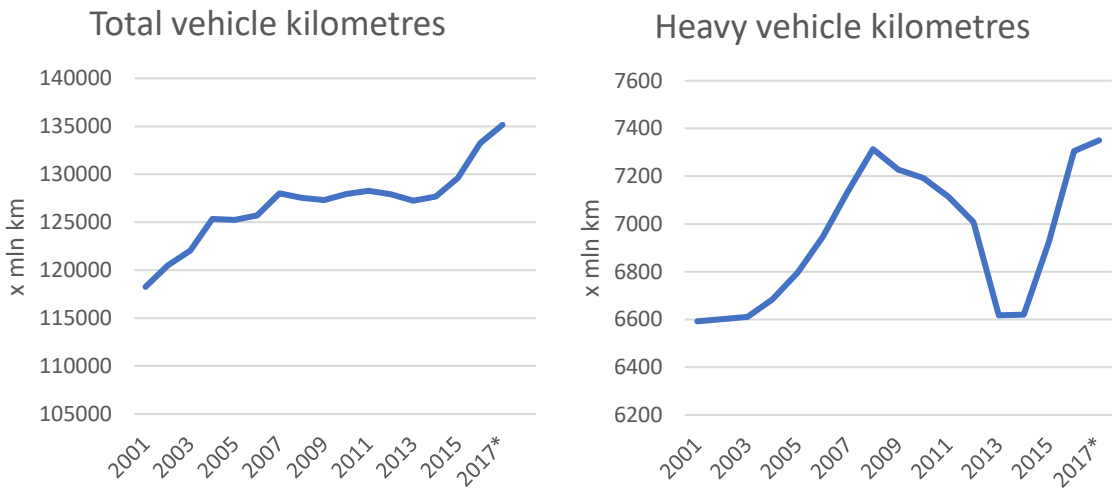


Figure 5 - Vehicle kilometres and truck kilometres in the Netherlands ("Verkeersprestaties motorvoertuigen; kilometers, voertuigsoort, grondgebied," 2019; "Verkeersprestaties vrachtvoertuigen naar gewicht en grondgebied 2001-2017," 2018)

The economic recovery becomes noticeable as the number of kilometres travelled raised considerably, particularly for heavy vehicles. This lead to an increase of 20% more congestion in 2017 compared to 2016 (ANWB, 2018), despite increasing expenditure on infrastructure each year by the government during the economic crisis to prevent future congestion (Wegenwiki, 2019). More asphalt might lead to less congestion in the short-term, however, it is questionable in the long-term. Therefore, innovative solutions were introduced such as peak hour lanes, matrix signs, and ramp metering systems to support enlargement or retention of capacity and prevent congestion. Despite these solutions to maximize roadway capacity, the drop of capacity after congestion sets in, remains an active field of study and will be central in this research.

The capacity drop phenomena can be observed in Figure 6. A sudden drop in flow rate after traffic breaks down can be observed rather than a descending slope opposite to the situation pre-breakdown. This phenomenon results in a discharge flow that is 3 – 18% lower compared to the pre queue flow(Yuan, Knoop, & Hoogendoorn, 2015). To better comprehend the capacity drop phenomenon, researchers investigated the impact on the flow rate of several circumstances, e.g. different merging sections (Chen & Ahn, 2018; Leclercq, Knoop, Marczak, & Hoogendoorn, 2016), number of lanes (Oh & Yeo, 2012) and speed in congestion (Yuan et al., 2015). The main cause of the capacity drop has not been identified yet, but relationships between drop rate and certain variables have been discovered. Researchers have pointed at heterogeneity of traffic to declare some of the observations regarding the capacity drop (Yuan et al., 2015). However, a lack of scientific evidence is present, and research is needed to verify this conjecture.

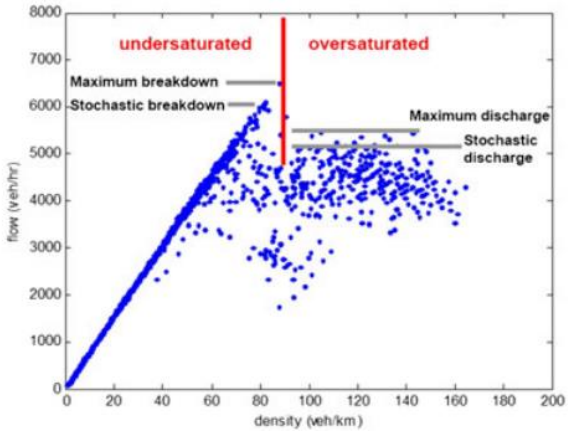


Figure 6 - Capacity drop phenomena explained (Calvert, Taale, & Hoogendoorn, 2016)

Although the number of heavy vehicles is relatively small compared to regular vehicles, their impact is prominent, particularly during heavy traffic conditions. Due to divergent physical characteristics (length and size) and operational characteristics (acceleration and deceleration) compared to regular vehicles, heavy vehicles impose physical and psychological effects on surrounding traffic (Moridpour, Mazloumi, & Mesbah, 2015). Despite the growing amount of freight transport on the road, as forecasted in a growing economy (de Bok, Wesseling, Kiel, Miete, & Jan, 2018), the influence of heavy vehicles on traffic flow has received little attention.

Furthermore, new emerging technologies in heavy vehicles are becoming available rather soon with the production of semi-autonomous trucks, such as the Tesla Semi in the second half of 2020 (Lamberts, 2019). Self-driving abilities such as Lane-keeping Assistance and Adaptive Cruise Control will likely be introduced, perhaps including truck platooning. Moreover, these heavy vehicles will be hybrid or electric resulting in different operational characteristics, such as a higher acceleration capability. Measures taken by local governments to introduce zero-emission zones in cities for logistics by 2025 (Rotterdam, 2019) might further push this trend. However, the effects of these changed operational characteristics and self-driving abilities of trucks on traffic flow are yet to be discovered. Some simulation models predict beneficial results regarding the smoothing of traffic flow (Tsugawa, Jeschke, & Shladover, 2016) while low-level automation penetration is expected to initially have a negative impact in mixed traffic on traffic flow (Calvert, Schakel, & van Lint, 2017). As the impact of freight traffic on the capacity drop is yet to be empirically discovered, changed operational characteristics of trucks are around the corner. Researchers reckon that vehicle heterogeneity of the traffic flow composition is not a necessary condition for the appearance of the capacity drop but could be a determining factor in reducing the queue discharge rate (K. Yuan, 2016). Current knowledge cannot explicitly explain how vehicle heterogeneity reduces the queue discharge rate. Filling this gap can improve road performance, e.g., by introducing vehicle-class specific control strategies or lane-specific management strategies. Roadway authorities could this way purposefully adapt their management strategy based on the actual truck penetration rate.

1.2 Research Objective & Main Research Questions

While kilometres travelled rise in a growing economy, so does the time travelled in congestion. As mentioned in paragraph 1.1, the amount of freight traffic kilometres is rapidly increasing and new emerging technologies in trucks will be available soon leading to changed operational characteristics. However, the influence of heavy vehicles on traffic flow and particularly the capacity drop phenomenon has received little attention.

One of the main concerns of traffic authorities is travel time losses. Breakdown of traffic will not only result in congestion but also in a drop of capacity during recovery hence the pre-queue inflow is higher than the queue discharge flow (Banks, 1991; Hall & Agyemang-Duah, 1991). Many measures have been taken to prevent breakdowns of traffic, such as ramp-metering and matrix signs, both aiming to stabilize traffic flow. Therefore, if the relation of trucks on the capacity drop could be identified, appropriate measures could be taken to purposefully prevent or minimize this drop, dependent on the truck penetration rate. Assumptions about observed phenomena, however, have been appointed to the heterogeneity of traffic as state-of-the-art traffic flow theory could no longer elucidate.

Collaborative research between TNO, Port of Rotterdam and Rijkswaterstaat, analyses logistic processes and traffic management systems to detect inefficiencies and to understand the potential of a combination of logistical and traffic management measures. Understanding the impact of freight traffic on the capacity drop could support traffic management operators to maintain travel times by preventing congestion. Highways, such as the A15, connecting the Dutch road network to the port of Rotterdam, are known as truck-dominant road sections and the performance from a traffic flow perspective might therefore be improved with an individual traffic management strategy.

Based on the problem statement, three main research questions are formulated:

Research Question 1: **Which variables influence the capacity drop rate on both strategic and operational levels?**

The mechanism of the capacity drop should be understood and captured into a theoretical framework. In addition, the impact of other factors should be understood. This led to the following question:

Research Question 2: **How and to what extend does freight traffic impact the capacity drop on highways currently?**

A calculation framework to estimate the capacity drop should be designed based on previous studies. This supports verifying and comparing the results. Based on the assumed relationship between freight traffic and capacity drop found in literature, the next follow-up question is formulated to further investigate the problem:

Research Question 3: **How do different vehicle heterogeneity compositions and, physical- and operational characteristics of freight traffic impact the capacity drop?**

The supporting research questions will be formulated in section 1.3, after the methodology and supporting research framework have been elaborated.

1.3 Research Methodology

The research framework in Figure 7 translates the main research questions into research stages with evident work packages. These work packages induce activities that eventually lead to a detailed time planning, which can be found in the research proposal. This research will be separated into three stages: the Literature Study, the Data Analysis, and the Simulation. The arrows indicate the order of the activities and the dependency of the different activities. The numbers provided next to some of the blocks indicate the chapter in which that part of the research can be found. The research will consist of six chapters. Below, each research stage will be elaborated based on the activities, which can be found in sub-chapters.

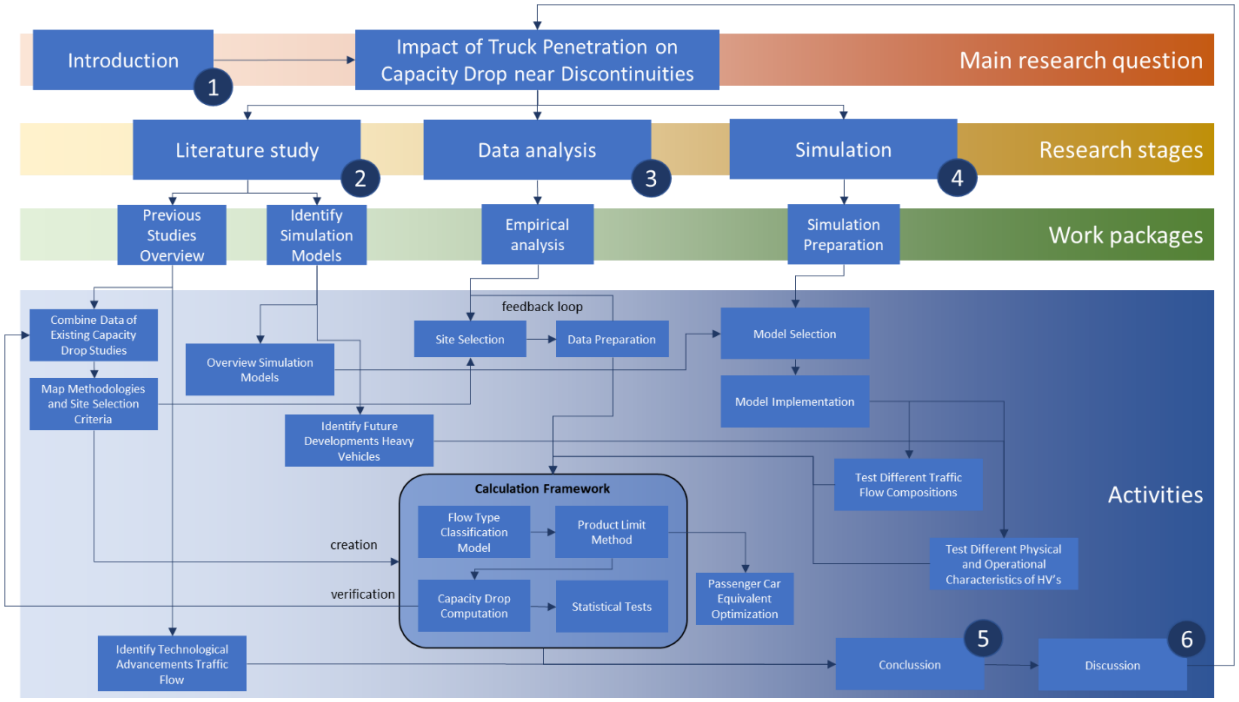


Figure 7 - Research framework

Stage 1 – Literature study

First, existing literature will be used to identify the relationship between truck penetration and capacity drop. Several studies measured the truck penetration rate without further use. Therefore, combining data of multiple existing capacity drop studies could identify fledgling relations. In addition, an overview of what factors influence the capacity drop, the breakdown, or recovery intensity will be created. Besides, methodologies of the different studies will be mapped to define site selection criteria. Furthermore, future developments will be researched. This will support the question of how future traffic will impact the capacity drop. Also, different microscopic simulation models will be investigated as one of these models will be used for the simulation in the last research stage. Lastly, technological advancements will be identified, to get an idea of how currently traffic flows are influenced and how possibly the capacity drop could be reduced.

Stage 2 – Data analysis

A set of sites will be selected with different configurations and vehicle compositions based on the site selection criteria denoted in the first stage. The Dutch loop detector data available by TNO will be prepared to use for the empirical analysis. The calculation framework will be designed on mapped methodologies, consisting of a Flow Type Classification Model, Product Limit Method, Capacity Drop Computation, and Statistical Analysis. Besides, the Passenger Car Equivalent (PCE) of heavy vehicles will be computed for each road to show the space heavy vehicles take in terms of passenger cars in traffic flow pre-queue and queue discharge.

Stage 3 – Simulation

A simulation model is chosen to fit the purpose and criteria of this research. One specific simulation set-up will be prepared and used for all tested scenarios. Subsequently, the Calculation Framework will be reused to compute the capacity drop. Therefore, the simulation data should be equivalent to the empirical loop-detector data. It should be reckoned that the focus of this study is not on designing or improving an existing simulation model. The focus will be solely on using the simulation software to the best of its abilities to investigate the impact of different truck penetration scenarios on the capacity drop.

To support the research stages, research questions have been formulated in Table 1. Each question can be identified within a block of the research framework. Some main supporting research questions, indicated with a capital, contain multiple sub-questions, indicated with a lowercase letter. The main supporting questions will be used to answer the corresponding chapters.

Table 1 - Supporting research questions

Research stage	Supporting questions
Stage 1	A. What relationship(s) can be yielded from previous research about the capacity drop for different truck penetration rates? a. What can be yielded from empirical research? b. What can be yielded from modelling research? B. How does freight traffic impact traffic flow now and in the future? a. What are the developments in freight traffic? b. What are the operational implications on traffic flow? C. What methodology is most suited to investigate the impact of freight traffic on the capacity drop? a. What is needed to compute the capacity drop? b. What methodologies are suited? c. How to select suitable locations? D. What measures have been previously taken (or researched) to reduce or prevent the capacity drop? E. What simulation models are available to investigate the capacity drop phenomena? a. What are the pros and cons of existing models? b. Which model is most suited for this research?
Stage 2	A. What data is measured by the Dutch loop detectors? a. What variables are measured? b. How can the data be used for this research? c. Which vehicle types are distinguished? B. Which roads are suited as a case study? a. How to select suitable roads? b. How do the theoretical case studies match the practical case study? c. What are the characteristics of the selected roads? C. How to use the chosen methodology to answer the main research question? a. How to classify different flow types? b. How to compute the capacity drop? c. How to calculate the PCE? d. How to test the results for statistical significance?

	D. What can be concluded from the results? a. How does the number of heavy vehicles impact the capacity drop? b. How do different site characteristics influence the capacity drop? c. What do the PCE values conclude about the impact of heavy vehicles? d. What other phenomena have been observed?
Stage 3	A. How to implement a case study for the simulation? a. Which variables should be adjusted? b. How does the simulation model distinguish different vehicle classes? c. How to output the data similarly to the empirical data? d. What scenarios will be tested? e. How to calibrate the model for the case study? B. What is the effect of different traffic flow compositions? a. What measurements are impacted? b. How does it affect the capacity drop? C. What is the effect of different truck characteristics? a. For different truck penetration rates? b. For changed operational characteristics of trucks?

1.4 Scientific & Societal Relevance

The scientific contribution of this research would be to investigate the impact of freight traffic on the capacity drop and place this in the general knowledge of other variables that influence the capacity drop. This study provides the ToGRIP project with new insights into the impact of freight traffic on the capacity drop. Currently, researchers are working on designing a lane change model for trucks, which would benefit from knowledge on the relation between heavy vehicles and the capacity drop. For example, the breakdown distributions for different heavy vehicle shares could help in estimating the probability of a breakdown and help when the lane change advisory should be activated to maximise performance.

From a societal perspective, this research could motivate roadway authorities to design road-specific management systems depending on the traffic composition of the road. An individual lane management system might be a solution, leading to a more improved traffic flow for all vehicles. It would also be in the interest of transport companies since such a management system would decrease travel time, which in the end improves business processes. Smart cooperative (heavy) vehicle systems could use the results to their benefit, as they might be able to reduce or diminish the capacity drop. This would imply that congestion could be solved quicker, or even be prevented.

2 LITERATURE REVIEW

The capacity drop has been researched in both empirical and simulation studies. First, established empirical and simulation studies related to the capacity drop are elaborated in 2.1. Second, the characteristics and development of heavy vehicles are discussed in 2.2. Different methodologies to compute the capacity drop are discussed in 2.3. Finally, different simulation models are explored 2.4.

2.1 Overview Capacity Drop Studies

The capacity drop phenomenon was first empirically confirmed with an observed reduction of maximum flow rate after the onset of congestion in 1991. Since that moment, a lot of research has been executed to grasp the cause and impact of the drop rate. A clear distinction can be made between empirical studies (section 2.1.2) which are solely based on measured data and modelling studies (section 2.1.3) which use calibrated simulation models based on driver models. Especially impact of emerging technologies in vehicles on the general traffic flow is difficult to measure empirically and therefore always simulated. Section 2.1.4 gives an overview of the impact on the capacity drop, related to the pre-queue flow and queue discharge flow. The results will be summarized and discussed in 2.1.5.

2.1.1 Magnitude of the Capacity Drop

Hall and Agyemang-Duah (1991) observed duality of capacity, a clear gap between the observed data in the congested branch and uncongested branch, by comparing mean flows before and after queue formation at bottlenecks. This drop of capacity would only restore after congestion disappeared, leading to a reduced throughput when traffic flow is high. This observation raised the question of how much road capacity drops during such breakdown in terms of Passenger Car Equivalents (PCE's). To do so, the researchers measured the number of heavy vehicles manually by counting the number of trucks passing by. An average of 6% heavy vehicles was observed. By doing so, they were able to compare their observations to the Highway Capacity Manual (HCM) and compute the difference between the theoretical and observed capacity after a breakdown. Subsequently, a drop in capacity of about 6% was computed after congestion sets in.

Simultaneously, Banks (1991) found the same traffic phenomenon in a different study. Moreover, the traffic flow on the left lane appeared to have a remarkable higher drop of 10%, compared to 3% on the right lane. This raises the question of what leads to the unequal distributed drop over the lanes. At some of the measured sites, the capacity drop for all lanes turned out to be negative, meaning that the average flow increased during recovery, compared to the breakdown moment. As the recovery intensity was measured 12 minutes after the breakdown, queues were sometimes cleared up spontaneously. The authors argued that the magnitude of the capacity drop mainly depends on site characteristics. An average heavy vehicle percentage between 1.9 – 4.5% for the different sites was observed. Both these studies are the first research to discover the capacity drop and initiated more research to estimate the capacity drop at different locations, unravel the main cause of the capacity drop and understand the magnitude of the drop.

Two decades later, follow-up studies were summarized by Oh and Yeo (2012), bundling different capacity drop values of studies before 2008. A distinction was made between different ramp types and the number of lanes to test for correlation with the capacity drop, as this was the aim of the study. Their overview can be observed in Table 2 and has been extended with more recent studies. More research showed that the observed drop rates of Hall and Agyemang-Duah (1991) and Banks (1991) were relatively low compared to the newer observations. Therefore, the capacity drop phenomena should be considered a serious issue within the traffic flow theory field as it deteriorates the capacity decently when the high traffic flow requires maximal throughput.

It should be noted that most research is based on American traffic flow data. This is important since the American keep-your-lane driving style causes different driving behaviour compared to the European stick-right driving style. How this impacts traffic flow, in general, is difficult to capture, and to the authors' knowledge has not been researched yet. Since most research investigates (near) congested conditions, this difference in driving style is assumed to be less important.

Table 2 - Capacity Drop in Previous Research

Researcher	Location	Sites	Collection period (days)	No. of lanes	Truck Penetration (%)	Capacity drop (%)
Hall and Agyemang-Duah (1991)	Toronto	1	36	3	6	5.8
Banks (1991)	USA	4	31	2 - 4	1.9 – 4.5	-1.2 – 3.2
Elefteriadou, Roess, and Williams (1995)	Chicago	3	2	3	-	10
Cassidy and Bertini (1999)	Toronto	2	3	3	-	8 – 9
Bertini and Malik (2004)	Minneapolis	1	4	2	-	4
L. Zhang and Levinson (2004)	Minneapolis – St. Paul	27	23 – 28	2 – 4	0.7 – 8.8	2 – 11
Regler (2004)	Germany	15	180 - 365	3	0 – 20*	8 – 40*
Cassidy and Rudjanakanoknad (2005)	San Diego	1	10	4	-	11.7
Banks (2006)	Minneapolis – St. Paul, San Diego and Seattle	21	60	2 – 5	2.9 – 8.8	2 – 17
Chung, Rudjanakanoknad, and Cassidy (2007)	San Diego, San Francisco and Toronto	3	3 – 12	2 – 4	-	5.8 – 18
Oh and Yeo (2012)	California	16	5 – 24	2 – 5	-	8.4 – 16.5
Yuan et al. (2015)	Netherlands	2	30	3	>15	25*

**No overview of individual results, only aggregate numbers*

As many studies did measure the number of heavy vehicles, although used for different purposes, it is possible to start exploring whether a relation between traffic composition and the capacity drop exists. Therefore, truck percentages have been added to Table 2. Measuring traffic composition is quite difficult, especially when advanced loop detectors were not available in early studies. Therefore, these observations should only be considered a rough estimation. Information about the sites, methodologies, and measurements including Queue Discharge Flow (QDF) and Pre-Queue Flow (PQF) was gathered from these studies to support verification in the data analysis section.

The relation between the capacity drop and the truck penetration based on data of previous studies has been visualized in Figure 8 from the studies from Table 2 who published their non-aggregated data. No clear relationship can be yielded between the percentage of heavy vehicles and the capacity drop. However, as the measurement methodologies differ substantially, the data analysis in this study will investigate more comparable data, leading to a more robust investigation.

The maximum measured truck penetration rate measured was about 9%. Nowadays, roadway authorities have guidelines for truck penetration rates up to 25%, which is considerably more than observed in these studies. Therefore, more data should be added of roads with higher penetration rates to see how this relationship holds for the more truck-dominant highways.

It should be noted that the studies did not use a similar methodology for collecting data, computing the capacity drop, nor have identical site characteristics. As mentioned, the main goal of these studies was not to relate truck penetration to the capacity drop. However, the necessity to investigate the impact of heavy vehicles on the capacity drop, especially for truck-dominant highway sections becomes clear as other studies do not find the main cause related to other attributes but do account for the vehicle composition to be an important factor.

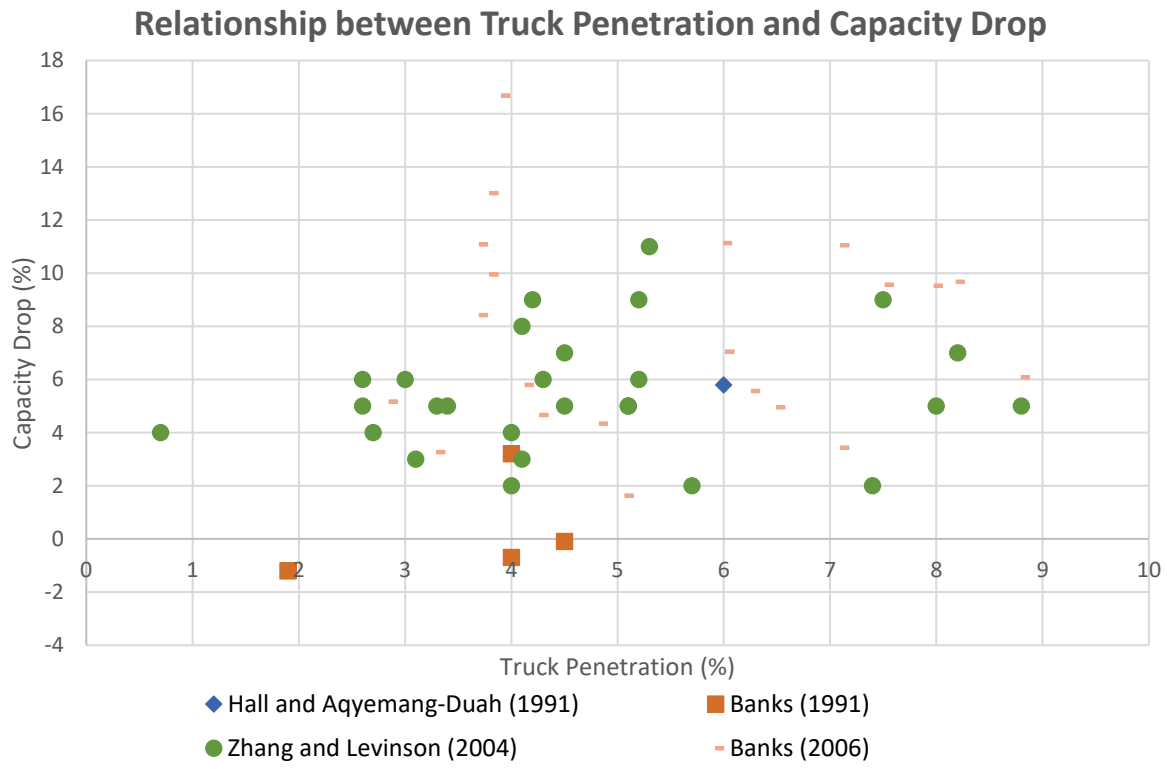


Figure 8 - Relationship between Truck Penetration and Capacity Drop

2.1.2 Empirical Factors

Oh and Yeo (2012) investigated the relationship between the number of lanes and the capacity drop in an empirical study. Sixteen on-ramps merging bottlenecks were selected in California ranging from 2-lane highways to 5-lane highways. A negative relationship was found between the number of lanes and capacity drop, decreasing from 16.33% for two-lane highways to 8.85% for five-lane highways. Microscopic analysis revealed that a reduction in disturbance (lane changes) causes an increase in the queue discharge rate. At four and five-lane highways, the capacity drop is found to be significantly lower in the outer lanes which the authors suggest might be caused by the number of trucks as well as merging and exiting vehicles in these lanes.

To investigate the impact of merging vehicles, Srivastava and Geroliminis (2013) compared mainline flow to ramp flow related to the capacity drop. The inclination in mainline volume compared to ramp volume during the congested phase suggests that volume stabilizes for a respective 2:1 ratio, implying that every merging vehicle is twice as detrimental to the congestion level compared to an additional mainline vehicle. Ramp flow does therefore have a huge impact on the capacity drop based on the volume of the merging flow. Besides, a ramp metering system was placed during the study period including different control strategies, to investigate how this would affect the drop. However, it appeared that the capacity drop phenomena remained at a similar

magnitude for all different ramp control strategies. Both these studies show that the design of a roadway does influence the impact on the capacity drop, but lack guidance on how to bring reduced capacity drop to practice.

A more traffic flow-oriented study executed by Yuan et al. (2015) reveals the relationship between vehicle speed in congestion and the queue discharge rate. Data analysis of Dutch loop detectors revealed that the queue discharge rate increases considerably with increased speed in congestion. A remarkable high queue discharge rate (6.840 veh/h) was found in this study compared to theoretical free flow capacity (6.300 veh/h). The studied Dutch freeways A4 and A12 are assumed to have considerably lower heavy vehicle share than 15%, which is used by the Dutch Highway Capacity Manual for capacity estimations, and thus the queue discharge rate could be considerably influenced by the share of heavy vehicles. Although a statistical relationship was found between queue discharge rate and speed in congestion it is suggested that local site characteristics such as traffic flow composition and weather could substantially influence the quantitative relationship found.

Another traffic flow relation with the capacity drop was researched by Oh and Yeo (2015). They investigated the impact of stop-and-go waves and lane changes on queue discharge rate on a microscopic scale. One of the influential factors found on decreased queue discharge flow was that drivers increased their following distance after passing stop-and-go waves. Moreover, it was observed that lane-changing can increase stop-and-go wave probability and those waves subsequently lead to a drop in capacity as traffic switches between an acceleration and deceleration state. Besides, a lower capacity drop in the outer lane was observed, which is argued by the authors can be caused by higher truck ratios in the inner lanes. Specific characteristics such as lower acceleration capability and larger safety gaps from surrounding vehicles support this reasoning. However, future research into this effect is proposed as quantitative research on this topic lacks.

Another possible relationship is time-dependency. Calvert et al. (2016) investigated the impact of the reduction in breakdown capacity on weekend days compared with workdays. They define two types of capacity: design capacity and operational capacity. The first is used for the design of a road, while the latter is the actual flow at which breakdown occurs, dependent on actual specific road conditions. In his study, the operational capacity is researched. The results show a significant decrease of 8% in queue discharge flow when comparing workdays with weekend days.

The influence of precipitation on capacity is researched by Calvert and Snelder (2013). It was found that the operational capacity drops by 3-4% during rain, followed by approximately 1.9% per increase in mm/h rain. It should be noted that both congested and uncongested observations were used for this analysis, resulting in a combination of pre-queue flow and queue discharge values for the estimation of the reduction. How precipitation influences each flow type individually might be another topic of research.

The long-term variation in traffic capacity has been investigated by Shiomi, Xing, Kai, and Katayama (2019). A case study in Japan was executed, consisting of 9 bottlenecks measured between 2008 and 2016. A decreasing tendency in traffic breakdown probability for the fifth percentile of traffic volume was shown long-term, while August and December both show a decrease compared to the yearly mean. The latter is explicable as both months are within the holiday season in Japan, leading to more less-experienced drivers. The decreasing tendency is attributed to characteristics of vehicles, drivers, and both. The researchers conjecture that the dramatically changing vehicle performance, e.g. more hybrid vehicles each year and more vehicles equipped with ACC, play a major role as hybrid vehicles, in general, have lower throttle response compared to gas-fuel vehicles. Following distances are manually set to a longer distance due to preference of the driver.

From these past studies, it can be concluded that many variables may impact the drop rate. The stochastic nature of the capacity drop becomes clear since many investigations to measure influential variables lead to diverse drop rate values. Each factor identified will be added to the overview in 2.1.4, and will, be used for the development of the Calculation Framework in 3.4.

2.1.3 Modelling Factors

A model that reproduces the drop in discharge rate when congestions sets in was developed by Laval and Daganzo (2006). It is based on the theory that voids are created due to lane changing. The reduced discharge rate phenomenon is very well captured in the four-parameter multilane hybrid model. Each lane is modelled individually and lets lane-changing vehicles blocks occupying lanes. A relationship between the capacity of a moving bottleneck and its speed seems to be apparent. A moving bottleneck with very low speed seems to cause a lot of voids due to the lane-changing vehicles which are unable to accelerate instantaneously afterward. Freight traffic is also a kind of moving bottleneck since its allowed speed is generally lower compared to regular traffic. In practice, two different truck-dominant roads with identical site characteristics except for free-flow speed should result in a decrease in queue discharge flow, and subsequently higher capacity drop.

One important aspect in the minimization of those voids is the reaction time of individual drivers. Calvert, Van Wageningen-Kessels, and Hoogendoorn (2018) modelled the effect of different reaction times in heterogeneous traffic on the capacity drop. The FOMSA (First Order Model with Stochastic Advection) is used. This is a Langrangian formulation of the kinematic wave model (LKWM) with vehicle specifics to capture heterogeneity. It was observed that the reaction time has a significant impact on the capacity drop, reducing it to 0% when a reaction time of 0 seconds was simulated. The heterogeneity of traffic seemed to chiefly impact the breakdown capacity, compared to the impact of reaction time. However, further improvements in the dynamics behind the application of heterogeneity are suggested since only a single variable was used to model vehicle specifics.

An analytical framework to estimate the effective capacity at freeway merges, with special attention for vehicle heterogeneity and physical interactions at merges was developed by Leclercq, Knoop, et al. (2016). Numerical results show that considering vehicle heterogeneity has no impact when not considering voids in estimating the effective capacity. However, heterogeneous vehicle characteristics are preliminarily found to be crucial for representing other traffic phenomena. Simulation with different flow variations for heterogeneous traffic flow shows to trigger stop-and-go wave appearance, which is not the case for homogeneous traffic flow. Subsequently, a set of analytical formulas was designed to transform discretionary lane changes into a lane-change flow, with capacity and on-ramp flow as variables (Leclercq, Marczak, Knoop, & Hoogendoorn, 2016). Sensitivity analysis showed that vehicle acceleration and truck ratio were the main factors influencing capacity. Different truck ratios per lane implementation are expected to further refine the analytical estimation.

Another study that examined the impact of heavy vehicle movements on traffic was performed by Moridpour et al. (2015). The research used NGSIM trajectory data in heavy traffic conditions. The researchers observed that passenger cars tried to avoid being in front or behind heavy vehicles by making a lane change or increase their space gap. Furthermore, lane-changing manoeuvres were researched using AIMSUN (Advanced Interactive Microscopic Simulator for Urban and Non-Urban Networks) and were found to be more frequent when more heavy vehicles were existent in the same lane. Especially in higher densities increased lane changing behaviour of surrounding passenger cars was observed. Also, worth mentioning is that when the percentage of heavy vehicles increases to 30% of traffic flow, an increase of 5% in the likelihood of accidents is estimated.

Besides the lane change behaviour caused by heavy vehicles, longitudinal driving behaviour in general also impacts roadway capacity. K. Yuan (2016) investigated strategies to increase the queue discharge rate from a longitudinal perspective. Two types of longitudinal behaviour mechanisms are defined which can reduce the queue discharge rate: (1) inter-driver spread and (2) intra-driver variation. The inter-driver spread explains the driving character composition into driving styles, e.g. timid or aggressive behaviours. While the intra-driver variation mechanisms catch the different behaviour patterns in different traffic conditions, from a perspective of reaction time. It was observed that the effect of (1) did barely influence the queue discharge rate spread and therefore is found to be not essential in modelling the capacity drop phenomena. However, the queue discharge rate effect of (2) was essential in capturing the capacity drop.

Continuing on previous research, Kai Yuan, Laval, Knoop, Jiang, and Hoogendoorn (2019) incorporated the empirically observed desired acceleration stochasticity into a car-following model to capture the capacity drop phenomenon. The results are verified against empirical traffic data and show different capacity drop rates under different traffic situations. The researchers found that the stochasticity of desired accelerations is a significant reason for the capacity drop. The queue discharge rate can be increased by either removing human error when accelerating out of the queue or by increasing acceleration at lower speeds. They, therefore, argue that new vehicle technologies might be able to reduce or eliminate the capacity drop and suggest follow-up research.

Furthermore, the spatial distribution of lane changes was also found to contribute to the magnitude of the capacity drop, depending on the type of bottleneck (Chen & Ahn, 2018). Several bottlenecks have been investigated, and the following phenomena were observed. First, lane changes closer to the downstream end of the bottleneck have a higher chance to create persisting voids. Second, weave bottlenecks have two counteracting effects of lane changes that affect the capacity drop: persisting voids and the utilizations of gaps created by diverging vehicles. Third, the more balanced the merging and diverging flows are, the lower the capacity drop. Last, the capacity drop is minimum if merging lane changes occur downstream of diverging lane changes, while a capacity drop is maximum for the opposite alignment. Furthermore, it is found that variation in acceleration affects the void size. The researchers conjecture that different spatial distributions of lane changes across vehicle types would affect the mean as well as the variation of the capacity drop.

2.1.4 Overview

Based on the studies elaborated in 2.1.2 and 2.1.3, an overview of factors has been made, which can be observed in Table 3. Factors that affect the queue discharge flow, pre-queue flow, or drop rate have been added. The impact overview is separated into empirical and modelling studies. The plus or minus sign indicates the relationship between the variable and observed traffic flow phenomena, as either positive correlated or negative correlated. For example, when the number of lanes increases on a highway, the drop rate decreases. The explanation is added in the last column to clarify the relationship.

Table 3 - Overview factors influencing the capacity drop

	Researcher	Variable	Traffic flow	+/-	Explanation
EMPIRICAL	Oh and Yeo (2012)	Number of lanes	Drop Rate	-	-2,49% per lane
	K. Yuan (2016)	Speed in Congestion	Queue Discharge Flow	+	+0,58% per km/h for 3 lanes
	Srivastava and Geroliminis (2013)	On-ramp flow	Drop Rate	+	2:1 saturation main line vs. ramp
	Oh and Yeo (2015)	Lane changes	Queue Discharge Flow	-	Transient increase until relaxation
		Stop-and-go waves	Queue Discharge Flow	-	Tendency to take significant headway
	Calvert et al. (2016)	Weekend days compared to work days	Pre Queue Flow	-	+8% for weekend days
	Calvert and Snelder (2013)	Precipitation	Pre Queue Flow	-	+3-4% + 0.8% per mm/h
	Shiomi et al. (2019)	Data collection period	Breakdown probability	-	Emerging technologies in vehicles hypothesis
		Holiday period	Breakdown probability	+	Less experienced drivers

MODELLING	Laval and Daganzo (2006)	Speed moving bottleneck	Drop Rate	-	Void utilization
	Calvert et al. (2018)	Reaction time	Drop Rate	+	Eliminating reaction-time eliminates capacity drop
	Leclercq, Knoop, et al. (2016)	Vehicle heterogeneity	Stop-and-go waves	+	Homogeneous traffic flow does not trigger stop-and-go waves
	Moridpour et al. (2015)	Number of trucks in the same lane	Lane changes	+	If density changes, so does lane changing behaviour
	K. Yuan (2016)	Inter-driver/vehicle spread	Queue Discharge Flow	+	Driver character composition. Only 3% at most
		Intra-driver variation	Queue Discharge Flow	+	Behaviour patterns in different traffic conditions
	Chen and Ahn (2018)	Lane change distance to bottleneck	Drop Rate	-	Persisting void utilization
	Kai Yuan et al. (2019)	Acceleration stochasticity	Queue Discharge Flow	-	New vehicle technologies seem to improve this

2.1.5 Summary and Discussion

The capacity drop conjecture was simultaneously confirmed and showed a difference between Pre Queue Flow (PQF) and Queue Discharge Flow (QDF) of 2 – 6% (Banks, 1991; Hall & Agyemang-Duah, 1991). Studies that are more recent showed that this drop rate could increase up to 18% and thus have a significant impact on the capacity after the onset of congestion. Combining data from earlier studies showed that with current data, no conclusion could be drawn from the correlation between heavy vehicles and the capacity drop. It should be noted that the highest observed truck penetration was only about 9%, where it is assumed that nowadays the number of trucks on the roads is considerably higher in the Netherlands.

When looking at the impact of certain variables on the capacity drop, a clear distinction was made between empirical studies and modelling studies. Empirical studies are based on measured data and therefore can qualitatively describe the observed phenomena. However, a lack of data or adaptability leads to less quantitative results. Contrary, modelling studies can capture different circumstances and generate a lot of data. State-of-the-art literature showed that on an empirical basis; the number of lanes, speed in congestion, on-ramp flow, number of lane changes, type of congestion, type of day, and precipitation are of influence on the drop rate. Besides, it was shown that the drop rate deteriorated over time, possibly due to the implementation of emerging technologies in vehicles. Modelling studies showed that void creation, reaction time, number of trucks, inter-driver/vehicle spread, intra-driver variation, spatial lane change location, and acceleration stochasticity are also decisive for the observed drop rate.

Some variables showed to have a more decisive effect than others. In the data analysis part, the focus will be on equalizing site characteristics. One of the factors that should be varied is the number of lanes, as this might be of high importance in truck-dominant highway sections (Moridpour et al., 2015) based on the theory of void utilization (Laval & Daganzo, 2006). The remarkable significant relationship between individual lane drop rate and the number of lanes (Oh & Yeo, 2012) should be verified for the spatial distribution of heavy vehicles, as this seems to be one of the decisive factors in determining the stochasticity of the drop rate.

2.2 Developments and Characteristics of Freight Traffic

Over time, the percentage of freight traffic has increased on the highways. To quantify that impact, three levels of impact on traffic flow can be divided: strategical, tactical, and operational level. The strategical level covers a wide time frame from policymaking of forecasted emerging technologies and sustainability changes down to the day-to-day variability of freight traffic on the road (2.2.1). At the operational level, the traffic flow impact on the smallest time scale caused by heavy vehicles will be identified for both longitudinal and lateral movements (2.2.2). The tactical level is in between both other levels and compromises measurement systems to maximise flow throughput. Technological advancements to improve traffic flow will be discussed in 0.

2.2.1 Strategic Level Developments

2.2.1.1 Long term developments

The amount of freight traffic on the road is highly dependent on the economic development of a country. Therefore, strategic freight transport models are developed to quantitively analyse long-term forecasts. Governments are interested in freight transport forecasts to satisfy the needs of policy making as infrastructure construction needs to be executed far ahead. The strategic freight transport model Basgoed is used by de Bok et al. (2018) to examine the impact of economic development (domestic growth by industry sector and international trade), infrastructure development, fuel prices, and logistic efficiency for different future scenarios. The freight traffic on the road has the potential to increase most, compared to other logistic sectors, as can be seen by the wide bandwidth in Figure 9. The steep rise in 2014 is, besides the economic welfare, is also due to the changed base registration for road transport, as 'Light Goods Vehicles' have since been added into the road transport registration.

This increasing trend is consistent with the growth of heavy traffic forecasted by Knoope and Francke (2019) commissioned by the Dutch government, which expects an increase of heavy vehicles between 1.2 - 8.4% for 2024. This forecast is based on the gross domestic product, the expected fuel price, the number of working inhabitants, and the capacity of the roads. Therefore, roadway authorities should expect an increase in heavy vehicles on the road in the coming years. How this will affect the overall heterogeneity of traffic is difficult to forecast as the number of commuters might be changed long-term due to the Coronavirus.

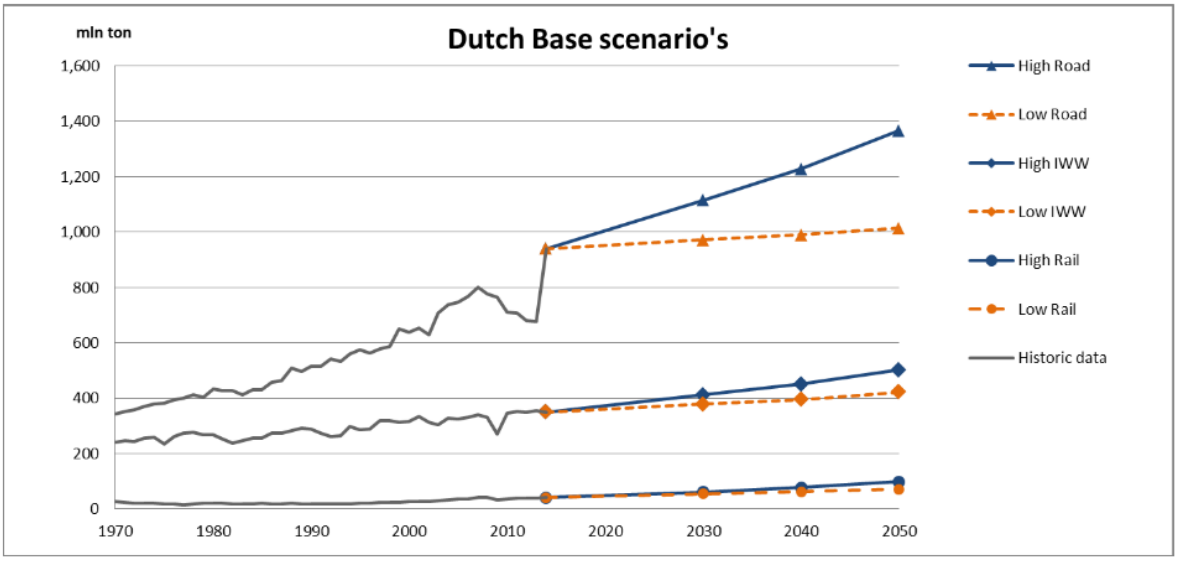


Figure 9 – Freight forecasts in High and Low scenarios, in mln. Tons and by mode: 2014-2050, compared to observed volumes 1970-2014 (de Bok et al., 2018).

On the other hand, the automotive industry is evolving. This leads to hybrid and electric trucks whereas they are nowadays based on diesel. Large consultancy agencies show different estimations for the expected share of electric or hybrid heavy vehicles. IDTechEX (2019) estimates the share of electric trucks to be about 9.4%

by 2030. BloombergNEF (2019) splits the commercial industry into light, medium, and heavy commercial vehicles and estimates that especially light commercial vehicles will most rapidly become electric as can be observed in Figure 10. By 2030, between 2 – 8% of the commercial traffic will be electric.

This adoption of electric or hybridized commercial vehicles will mainly depend on regulations made by local governments. For example, the zero-emission policy of the municipality of Rotterdam (2019) will no longer permit non-electric vehicles from the logistics sector within the city area from 2025. It is estimated that the number of heavy vehicles will decrease as smaller commercial vehicles, e.g. vans and lorries will be more common as the commercial industry will shift to smaller vehicles due to the new regulations. However, with the increase of tons transported as estimated by de Bok et al. (2018) the result will be that number of logistical movements will increase even more. So, it might eventually increase the number of heavy vehicles within traffic flow and push the truck penetration even higher on certain routes.

Besides, the impact of the changed operational characteristics on traffic flow is yet unclear. Moreover, it is difficult to estimate how these operational characteristics will change. Electric passenger cars can accelerate faster compared to gasoline cars, which will likely be the case for electric trucks as well. Unfortunately, to the authors' knowledge, no research has been executed in mapping the impact of electric vehicles on traffic flow nor the change in operational characteristics of trucks. Therefore, a study investigating the changed operational characteristics of regular vehicles compared to electric vehicles would help future studies to estimate operational parameters. This would lead to more accurate simulation studies of future scenarios.

Moreover, more high-level automation will become available. Technologies such as Adaptive Cruise Control (ACC) and Lane Change Assistant (LCA) are introduced and already available in passenger cars. These technologies will have a severe effect on car-following, lane-changing, and reaction time (Calvert et al., 2017). It was even shown that some of these automated features lead to a small decrease in capacity due to higher gap times and a marginal increase of capacity drop in mixed traffic. The idea of Cooperative Adaptive Cruise Control (CAAC) in trucks or Truck Platooning showed mixed results. If communication was only restricted to longitudinal control, Close-CACC platoons prevent other vehicles from merging eventually leading to a negative effect on traffic safety. Reduced time gaps and improved string stability in high-traffic volume conditions could reduce the number of shockwaves and standard deviation of speed (Van Arem, Van Driel, & Visser, 2006).

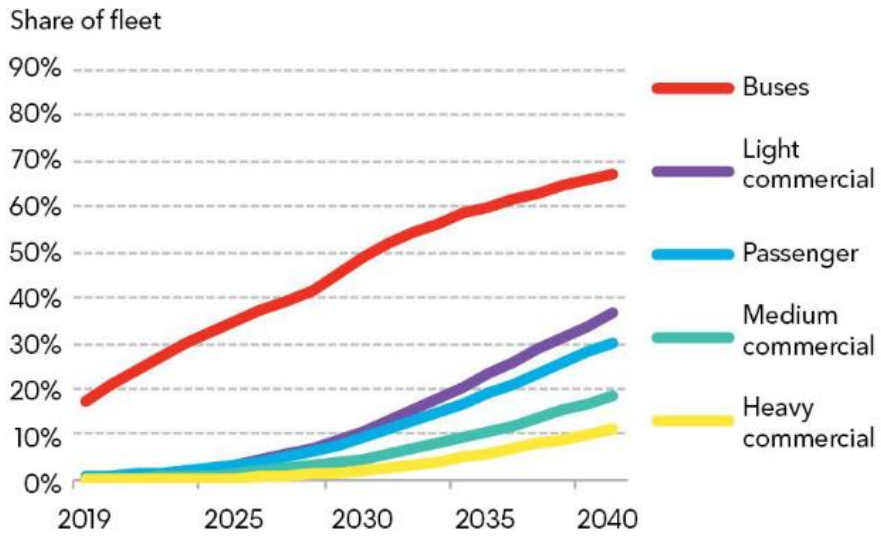


Figure 10 - EV share of the vehicle fleet by segment (BloombergNEF, 2019)

2.2.1.2 Passenger Car Equivalency (PCE)

The Dutch highway facilitates many vehicles per day with different travel purposes. This is paired with a wide variety of vehicles that travel on the road simultaneously, resulting in a heterogeneous traffic flow composed of vehicles with different physical and operational characteristics. Roadway authorities have the challenge to facilitate this traffic under different circumstances and minimize travel time losses. Therefore, sufficient infrastructure should be available according to the roadway capacity required.

Hence, it is necessary to account for heavy vehicles when computing roadway capacity as they occupy more space compared to passenger cars. A heterogeneous traffic flow is often described in a uniform measure of vehicles called passenger car unit (PCU) or passenger car equivalent (PCE), converting all vehicle types into an equivalent traffic flow of exclusively passenger cars. The main idea is to not only consider the static occupancy of the vehicles but somehow account for interactions between vehicles in a dynamic environment.

Many studies have investigated how to account for this PCE-value, and several methods can be found in Raj, Sivagnanasundaram, Asaithambi, and Ravi Shankar (2019). It becomes clear that Highway Capacity Manuals use different methods of how to account for PCE in the traffic steam and subsequently use different PCE-values. The National Research Council (2010) defines PCE as *"the number of passenger cars which will result in the same operational conditions as a single heavy vehicle of a particular type under specified roadway, traffic and control conditions"*, while TRRL (1965) London defines this as: *"on any particular section of road under particular traffic condition, if the addition of one vehicle of a particular type per hour will reduce the average speed of the remaining vehicles by the same amount as the addition of, say, x cars of average size per hour, then one vehicle of this type is equivalent to x PCU"*.

It seems straight-forward to convert physical characteristics from heavy vehicles into passenger car equivalents. Doing this while accounting for the operational characteristics is more complicated as this depends on many variables, e.g. the number of lanes or amount of heavy vehicles in the same lane. The PCE-value is useful in roughly estimating the capacity needed.

2.2.1.3 Heavy vehicle share

The (Dutch) National Database Road Traffic Data (NDW) distinguishes and measures different vehicle categories at measurement locations in the Netherlands based on physical characteristics of vehicles to make forecasts of the number of vehicles on the road. This initiative was implemented in 2009 to give insights into traffic jams to make better use of current infrastructure by providing better and more information. Vehicles are either classified into three or five categories based on the capability of the loop detector, see Table 4. This data collection allows for computing the truck penetration for any road for any time.

Table 4 - Vehicle classification NDW

Type	**C-3**	**C-5**	**Length (m)**	**Max. Speed (km/h)**	**Deceleration (m/s²)**
Motorcycle, scooter	Cat 1	Cat 1	1.85 – 2.40	100	-
Passenger car, delivery van		Cat 2	2.40 – 5.60	100	5.2
Lorry	Cat 2	Cat 3	5.60 – 11.50	80	4.5
Bus		Cat 4	11.50 – 12.20	80	
Articulated lorry	Cat 3	Cat 5	>12.20	80	

The individual measurements will be used in the Empirical Analysis (3), but an aggregation of the share of heavy vehicles is provided by INWEVA per road section. Now, the data is used to get insight into the share of heavy vehicles on the Dutch highways in general and will be useful in the Site Selection (3.2) for selecting sites with different shares of heavy vehicles. In Figure 11, the flow of passenger vehicles and commercial (category 2 and category 3 vehicles) over the day is displayed. Roads with considerable traffic (>50000 veh/day) were selected, resulting in a total of 173 road sections of about 300-7000 meters. The number of passenger vehicles

varies strongly along the day and peaks between 8:00-10:00 and 16:00-19:00 were observed, while the peak for commercial vehicles is slightly earlier, between 7:00-8:00 and arguably between 16:00-17:00. Besides, the mean vehicle accumulation for these locations, also the 25% and 75% thresholds were computed to show the variability of the data. In Figure 12, the heavy vehicle share was computed as part of the traffic flow. As expected, the heavy vehicle share is minimum during peak hours, as the share of passenger cars increases considerably.

The heavy vehicle share measured in Figure 8, investigating the heavy vehicle share – capacity drop relation, has a maximum measured truck penetration below 9%. The average truck penetration in 2019 in the Netherlands during morning peak hour vary between 7 – 14% for 75% of the roads. This indicates that multiple roads have a truck penetration above 9% and confirms that currently more trucks are on the highway compared to the earlier studies. Furthermore, the evening peak seems to have more passenger cars and fewer commercial vehicles compared to the morning peak. It should be noted that individual roads have different patterns from the average, e.g. the A67 connecting Belgium, the Netherlands, and Germany which appeared to have a truck penetration of 25% during the morning peak and 20% during the evening peak. Besides, individual days could have a strong variance in truck penetration.

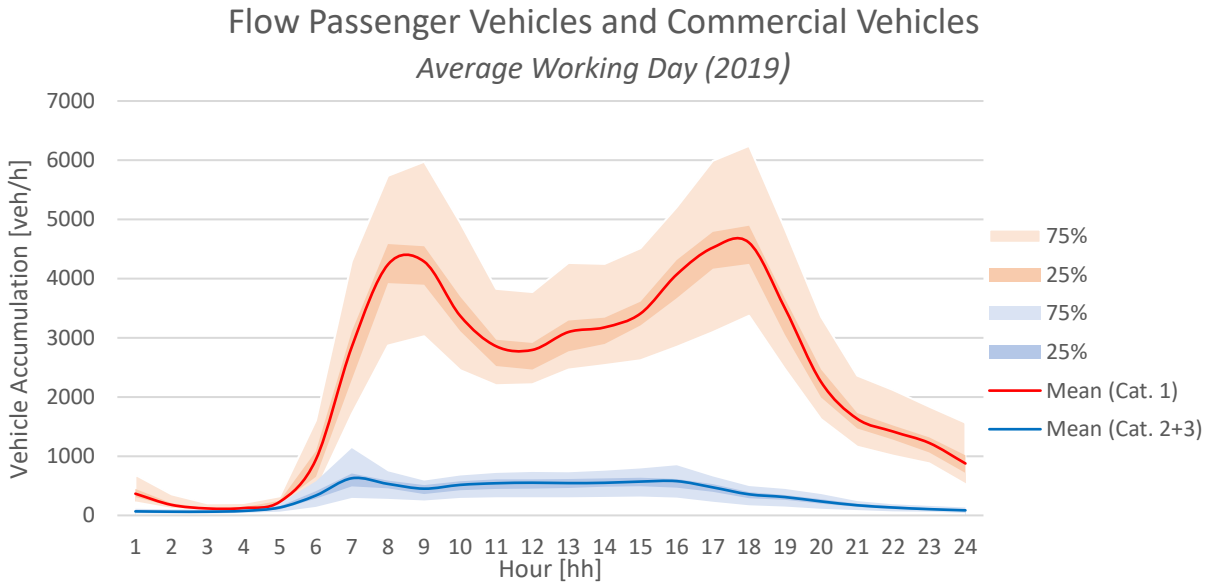


Figure 11 - Ratio Passenger Vehicles - Commercial Vehicles

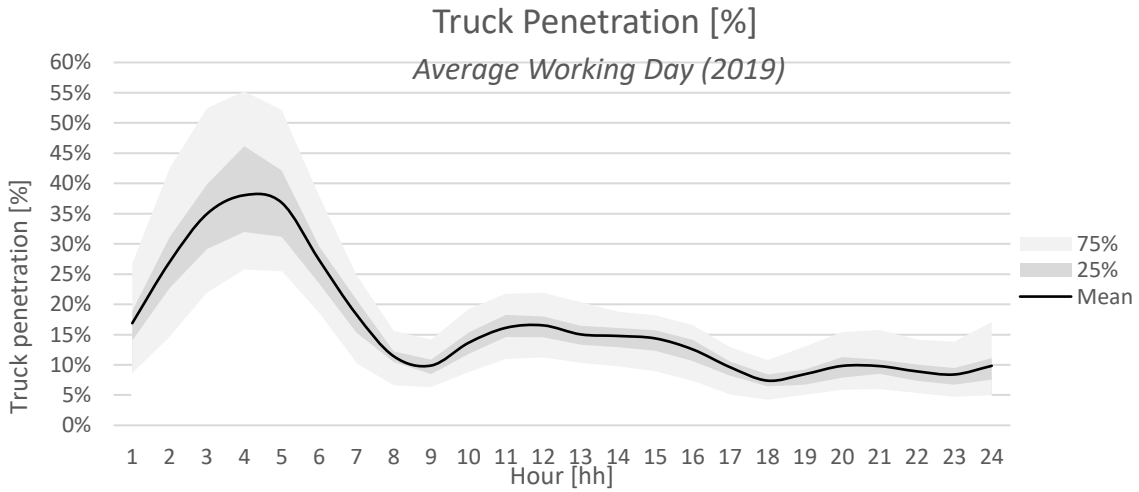


Figure 12 - Truck Penetration on an Average Working Day (2019)

2.2.2 Operational Implications on Traffic Flow

Heavy vehicles have different physical and operational characteristics in comparison to the average traffic flow. Besides, they have the potential to cause psychological disadvantages on surrounding passenger car drivers resulting in e.g. increased following distances or increased lane changing by other road users (Moridpour et al., 2015). Only a few studies empirically investigated the impact of heavy vehicles on traffic flow because loop detectors are often not able to classify vehicle type. Microscopic data is scarce and requires very extensive analysis because video images must be converted into trajectory data. Unfortunately, the largest microscopic dataset available, Next Generation SIMulation (NGSIM) data, where trajectories of vehicles are automatically created by an algorithm based on video images, is almost exclusively the empirical foundation for the majority of microscopic research today, but has shown to be more often wrong than right (Coifman & Li, 2017). Therefore, a knowledge gap seems to be apparent on the operational level while on the strategical level, methods to compute car equivalents for heavy vehicles are developed to account for heterogenous traffic flow.

Scientific research executed often aimed at investigating either longitudinal or lateral movements of heavy vehicles. Longitudinal implications concern the acceleration, relative speed, and following distance which will be elaborated in 2.2.2.1. The lateral movements concern lane changes of either the heavy vehicles themselves or caused by the heavy vehicles and will be elaborated in 2.2.2.2. Each of the implications will be discussed below. It should be kept in mind that most studies are only based on very few data collections. Therefore, more microscopic data collection is needed to verify and extend current research. Besides, this data could be used in calibrating microscopic simulation software, to test current assumptions regarding heavy vehicles on traffic flow.

2.2.2.1 Longitudinal

Four different car-truck combinations were formulated on a longitudinal level by Aghabayk, Sarvi, and Young (2012) based on the FHWA trajectory dataset. These combinations can be summarized as follows:

- Heavy vehicle following a passenger car (H-C)
- Passenger car-following a heavy vehicle (C-H)
- Passenger car-following a passenger car (C-C), and
- Heavy vehicle following a heavy vehicle (H-H)

The researchers observed that the H-H case has the longest space headway while the C-C case has the shortest space headway (Figure 13). The C-H and H-C case are both located in-between the other two cases. Below 30 km/h heavy vehicles tend to have a shorter space headway, while above 30 km/h, passenger cars tend to have a shorter space headway. It is argued by the author that this different behaviour is caused by the importance of sight distance and the braking power to mass ratio. Hence, freight traffic has an adverse effect on the following distance while the braking power becomes more important for the heavy vehicles when driving at higher speeds since their deceleration is not as high as for passenger cars.

A similar relationship between space headway and time headway is observed when comparing the relative order in Figure 14 to Figure 13. However, different time headway relationships were observed by other researchers. Ossen and Hoogendoorn (2011) stated that C-H time headway is the smallest among the C-H and H-C pair, whereas (Sarvi, 2011) found that the C-H pair has a longer time headway compared to the H-C pair. Ossen and Hoogendoorn (2011) argue that the assumption of restoring large deviations from the desired distance is more appropriate for passenger cars than for trucks and leads to smaller headway gaps for passenger cars. The opposite found by Sarvi (2011) could be explained by the psychological disadvantages that heavy trucks impose on surrounding traffic (Moridpour et al., 2015), such as the willingness to keep larger time headways.

The main difference between the studies is the collection site and video capturing. Ossen and Hoogendoorn (2011) collected data on the Dutch A15 and A2 during peak hour with a helicopter, while Sarvi (2011) used video cameras mounted on the top of buildings on freeways in both Tokyo and Melbourne. Aghabayk et al.

(2012) used the FHWA trajectory data set for some freeways in California, which is part of the NGSIM data collection and makes use of video cameras mounted on the top of buildings. All researchers argue that the results found are tentative since the amount of data is too small to conclude from. All three researchers found a slightly different relationship. Further research on this topic should investigate why different headway relationships were observed by the researchers, and if or how this affects traffic flow.

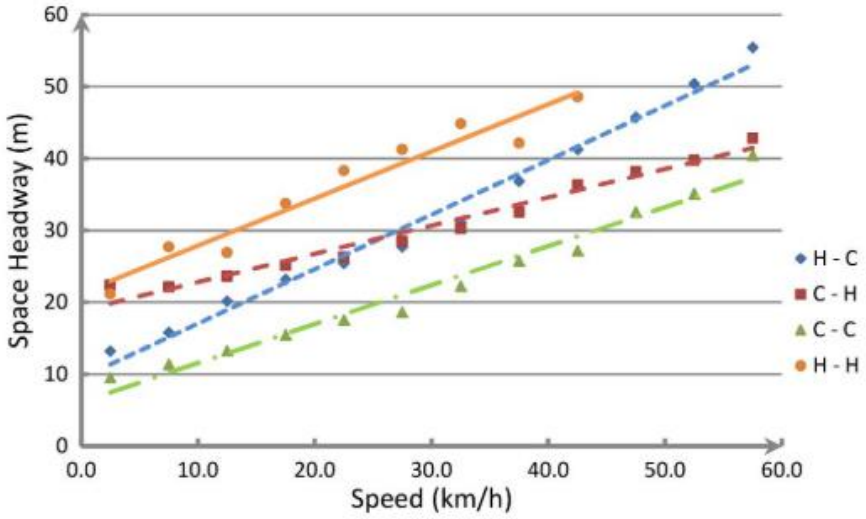


Figure 13 - Comparison of space headways (Aghabayk et al., 2012)

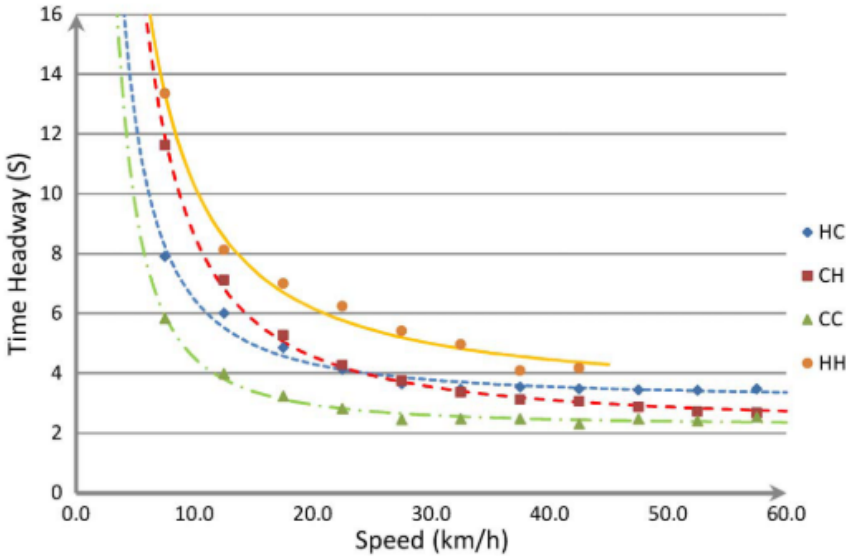


Figure 14 - Comparison of time headways (Aghabayk et al., 2012)

Another important factor in determining the truck-car following behaviour is the relative speed compared to the relative distance. Sarvi (2011) found a statistically significant difference between the following behaviour of heavy vehicles compared to passenger cars. It was observed, see results presented in Figure 15, that less than 5% of the heavy vehicles have a relative speed higher than 2 m/s, while less than 7% of the passenger vehicles have a relatively higher speed than 3 m/s. This indicates that cars behind heavy vehicles might have relative speeds up to 3 m/s, where this is only 2 m/s for a heavy vehicle following a car. It is argued that this is due to the lower acceleration-deceleration performance of heavy vehicles but also due to the experience of the heavy vehicle drivers since heavy vehicle drivers are professional drivers with high driving skills.

The acceleration threshold was similar for both heavy vehicles and cars. However, a different deceleration threshold was found for trucks compared to cars, which indicates that cars relatively decelerate later compared to trucks when the following distance decreases. Similar results were found in the research by Ossen and Hoogendoorn (2011) where truck drivers appear to have a more robust car-following behaviour, meaning that truck drivers were found to drive with a more constant speed compared to passenger cars.

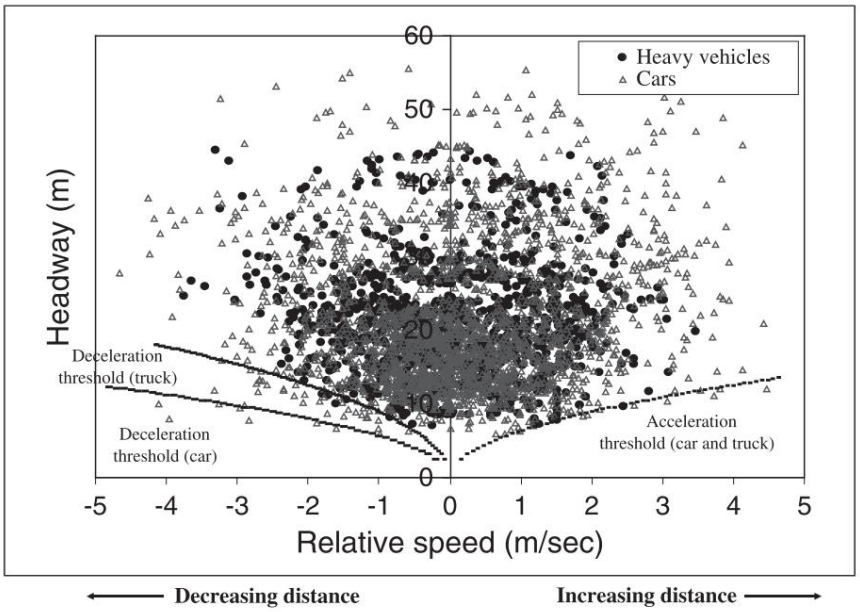


Figure 15 - Observed space headway as a function of relative speed (Sarvi, 2011)

Chen, Ahn, Bang, and Noyce (2016) go even further and argue that heavy vehicles tend to damper stop-and-go waves when following passenger cars by reducing speed variations. The authors did this based on NGSIM trajectory data. They defined five different reaction patterns, based on the time headway τ, namely:

- Concave triangle - in which τ increases but restores to the initial state
- Convex triangle - in which τ decreases but restores to the initial state
- Nondecreasing - in which τ increases (momentarily)
- Nonincreasing - in which τ decreases (momentarily)
- Constant - in which τ remains constant

For the H-C case, drivers mainly adopted a convex or nonincreasing reaction pattern. In contrast, for the C-H case, the concave reaction pattern dominated. The most common case, the C-C case is not dominated by either reaction pattern but is predominantly dependent on individual driver characteristics. The H-H case was not researched because of the lack of available pairs.

These findings explain the development of oscillations and are corresponding to the findings in Leclercq, Knoop, et al. (2016), who did not achieve stop-and-go waves in their simulation model when excluding vehicle heterogeneity. Hence, it is conjectured that heavy vehicles, when following passenger cars, have a dampening effect on traffic oscillations, while in contrast passenger cars following heavy vehicles have the opposite effect. However, the overall effect should be even since the number of H-C pairs should be quiet even compared to C-H pairs. Theoretically speaking, the composition of vehicle pair composition on the road would thus lead to different stop-and-go wave speeds.

The results are presented in Table 5 and indicate the dampening and growth effect of car-following pairs. Extending on these numbers, the average wave-speed estimated in this study is about -21 ft/s. The C-H case has the highest probability to grow the wave-speed, opposite to the H-C case where the dampening effect is the greatest. Since this study is based on very small sample size, the researchers impose further research with

more data to accept this conjecture. The different reaction patterns found for the acceleration-deceleration ability of heavy vehicles may be caused by the experience of the drivers, which is in line with the driving skill reasoning in Sarvi (2011).

Table 5 - Dampening and Growth effect of Car-Following Pairs (Chen et al., 2016)

	Heavy Vehicle–Passenger Car[a]		Passenger Car–Heavy Vehicle[a]		Passenger Car–Passenger Car[b]	
Effect	Magnitude (ft/s)	Percentage	Magnitude (ft/s)	Percentage	Magnitude (ft/s)	Percentage
Dampen	9.38	42	3.00	16	4.33	10
Grow	–5.00	5	–3.50	21	–4.40	17
No change	0	53	0	63	0.00	73

[a]Sample size = 19.
[b]Sample size = 30.

2.2.2.2 Lateral

Besides the imposed longitudinal changes in driving behaviour due to heavy vehicles, they also impose lateral changes on surrounding traffic. Empirical studies investigating lane-changing behaviour on a microscopic level caused by heavy vehicles are very scarce to the authors' knowledge. Therefore, general lateral traffic flow studies are elaborated below while relationships related to heavy vehicles are made.

First, it is necessary to comprehend the two types of lane changes: (i) mandatory lane changes (MLC) and (ii) discretionary lane changes (DLC) (Laval & Daganzo, 2006). The first happens when drivers have to make a lane change to get to the target destination lane, e.g. by taking the on- or off-ramp. The latter is done when drivers are not satisfied with the driving situation in their current lane and therefore make a lane change. Each type has its motivation and therefore result in different behaviour of the driver.

For mandatory lane changes, the number of lane changes towards the shoulder lane just upstream of an off-ramp was found to be higher than towards the median lane at 500 - 1500m upstream and downstream. The opposite was found downstream for an on-ramp (Knoop, Hoogendoorn, Shiomi, & Buisson, 2012).

Van Beinum, Farah, Wegman, and Hoogendoorn (2018) researched the merging behaviour of 14 bottlenecks in the Netherlands. The authors account different distributions in the mean accepted gap for merging found, with a long weaving segment and under high traffic flow conditions, to the difference in truck volume for different sites. Particularly when comparing the A13 on-ramp near Delft, with as low as 6% truck composition to the A15 weaving near Ridderkerk-south with a truck composition of about 20%. Despite both sites had a similar number of lanes and number of on-ramp/weaving vehicles, the KS-test proved that the mean-gap distributions are fundamentally different.

In regards to the discretionary lane changes, it is found that lane change rates increase when the density in the origin lane increases, resulting in drivers changing lanes on average approximately 0.5 times per kilometre driven (Knoop et al., 2012). Although mandatory lane changes cannot be entirely disclosed from these observations, it fits the theory of slugs and rabbits (Daganzo, 2002). This theory illustrates drivers by slugs and rabbits, whereby the rabbits want to drive to overtake slower drivers. This means that they have to switch lanes to overtake other drivers. Hence, from a game-theory perspective, the higher the density in the fastest lane, the more eager the rabbits are to drive in that lane, which corresponds to the findings by Knoop et al. (2012).

Chen et al. (2016) investigated the number of lane changes behind heavy vehicles compared to passenger cars. As expected, they found that in lanes with heavy vehicles the lane changing rate 'behind-out' is higher than the 'behind-in' rate, which corresponds to the typical driving behaviour to avoid driving behind heavy vehicles. Based on these observations, Chen et al. (2016) argue that the presence of heavy vehicles leads to underutilization of road capacity from a lateral view, since they induce lane changing but on the other hand improve traffic stability by the dampening effect from a longitudinal view. The number of lanes on the road is

conjectured to be an important factor of the lateral implications. One can imagine that on a 2-lane road with a lot of heavy vehicles, only a single lane is available to pass these heavy vehicles. However, on a 5-lane road, there is plenty of available lanes for passenger cars to choose the desired lane to fit the intended driving behaviour.

2.2.3 Summary and Discussion

The implications of trucks on traffic have been investigated on a strategical and operational level. Although commercial traffic is highly dependent on economic welfare, the increase in commodities transported via road is forecasted most (de Bok et al., 2018). Therefore, an increase between 1.2 – 8.4% of heavy vehicles is expected in 2024 (Knoope & Francke, 2019). Currently, the average truck penetration in the Netherlands is estimated at 10% during the morning peak and 7% during the evening peak, which is higher than previously investigated (<8%) and therefore raises concern about the impact on the capacity drop.

A change to hybrid or electric commercial traffic is expected and mainly dependents on the measures taken by local governments to stimulate emission-free transportation, e.g. only electric commercial traffic will be allowed in Rotterdam from 2025 onwards (Rotterdam, 2019). This change will be paired with the implementation of emerging technologies such as Adaptive Cruise Control and Lane Change Assistant. However, the impact of these technologies has yet to be discovered although shown to have a severe effect on car-following, lane-changing, and reaction time (Calvert et al., 2017).

The operational impact can be divided into longitudinal and lateral effects. Ossen and Hoogendoorn (2011) argue that the H-C time headway is larger compared to the C-H case due to the desire of drivers to restore large deviations. The opposite was found by Sarvi (2011) and can be explained by the imposed psychological disadvantages of trucks (Moridpour et al., 2015). Aghabayk et al. (2012) found that below 30 km/h, passenger vehicles keep larger distances to prevent lack of sight, while above 30 km/h trucks keep larger distances due to the larger mass-to-breaking-power ratio. The relative speed compared to the following distance also seemed to show different behaviour (Aghabayk et al., 2012). The reaction pattern of the time headway of truck drivers appeared to be convex and non-increasing, while for passenger car drivers a concave reaction pattern dominated (Chen et al., 2016). Therefore, it is argued that heavy vehicles impose a dampening effect on stop-and-go waves, while passenger vehicles have the opposite effect from a longitudinal perspective.

Lane changes performed or caused by heavy vehicles also cause an impact on traffic flow from a lateral perspective. Mandatory lane changes are made when needed to reach the destination, while discretionary lane changes are done when drivers are not satisfied with the driving situation in the current lane and result in different behaviour. The mean-gap distribution, which visualizes the accepted gap for merging, was found to be significantly different for ramps with a low truck penetration (estimated 6%) compared to high truck penetration (estimated 20%) (Van Beinum et al., 2018). On average, drivers change lanes approximately 0.5 times per kilometre driven (Knoop et al., 2012). More heterogeneity in vehicle composition will besides increase the number of lane changes. This fits the theory of slugs and rabbits (Daganzo, 2002). Subsequently, the lane-changing rate ‘behind-out’ was found to be higher compared to ‘behind-in’ for the C-H case (Chen et al., 2016).

The lack of research symbolizes the difficulty of collecting data. Observations in Ossen and Hoogendoorn (2011), Sarvi (2011), and Aghabayk et al. (2012) all showed different following behaviour regarding the C-H and H-C case. The underutilization of road capacity due to trucks is opposite to the improvement of traffic stability by dampening stop-and-go waves. How these two interact under different traffic compositions is yet unknown. Especially the number of lanes seems to be an important factor as more lanes result in more possibilities for discretionary lane changes and fits the theory of slugs and rabbits (Daganzo, 2002). Furthermore, previous research is based on truck percentages (<8%) below the current average measured truck penetration (10%) of the most crowded Dutch highway sections and urges the need for more research on this topic. Even more, as increasingly more commercial road-traffic is forecasted (de Bok et al., 2018).

2.3 Methodologies to Compute the Capacity Drop

To compute the capacity drop different established methodologies will be explored. First, different definitions of capacity will be given (2.3.1). Next, computation methodologies will be discussed based on case studies (2.3.2). Finally, site selection criteria will be illustrated (2.3.3).

2.3.1 Definition Capacity

A central concept in roadway design and traffic control is roadway capacity. Many computation methods have been developed depending on the available data and purpose to fulfil. Minderhoud, Botma, and Bovy (1997) distinguished three different types of roadway capacity, all serving a different purpose:

- **Design capacity:** Maximum traffic volume that may pass a cross-section of a road under predefined road and weather conditions. This value is most often prescribed by highway capacity manuals of governments and used for planning and designing purposes. It is derived from indirect-empirical capacity estimation methods.
- **Strategic capacity:** Maximum traffic volume that a certain road section can handle, which is useful for analysing conditions in the road network. Its value is derived from observed traffic flow data by static capacity models.
- **Operational capacity:** The actual maximum flow rate of a roadway, which may differ from time to time due to a specific road segment due to capacities stochastic nature. This value is useful in forecasting traffic flow and may be used for traffic control procedures. It is calculated with direct-empirical capacity methods in dynamic capacity models.

In this study, the focus will be on operational capacity and the direct-empirical capacity methods. Measuring the capacity of a certain road section can only be done with the corresponding reliability of traffic flow (Daamen, 2010). Therefore, capacity can be better described by a capacity distribution rather than a single capacity value. How factors influence the roadway capacity is not only influenced by the factor as such, but also by the dependency of that factor on other factors (Figure 16). A combination of ideal factors is used to compute the capacity of a highway (Highway Capacity Manual, 2010). Ideal factors are defined as (i) geometric construction of the road according to design criteria, (ii) excellent condition of the road, (iii) good weather conditions (iv) level terrain, and (v) homogeneous traffic flow. Since this study aims to investigate the impact of vehicle heterogeneity effects, it is important to minimize the impact of other factors.

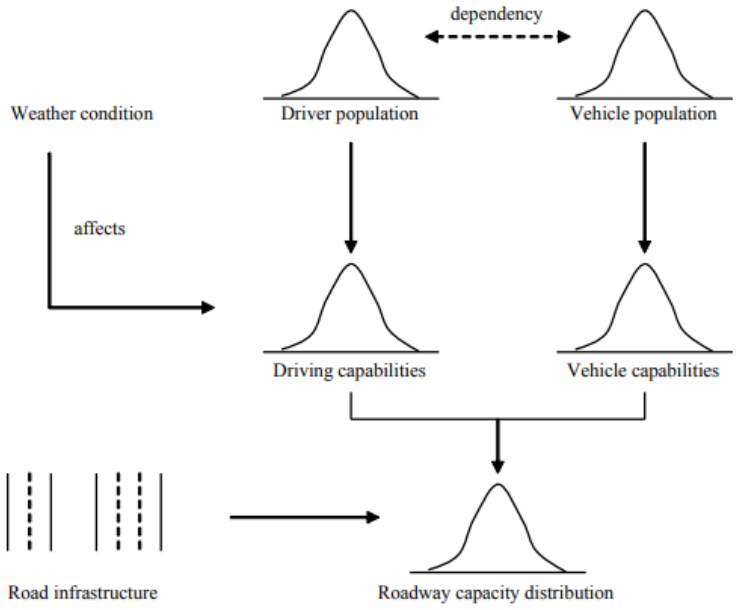


Figure 16 – Factors affecting the roadway capacity distribution (Daamen, 2010)

2.3.2 Computation Methods

Measuring capacity is complicated as the data is censored, more specifically right censored data (Lawless, 2003). This implies that estimating key characteristics has to be done with data points that are above a certain value, while it is unknown by how much (Daamen, 2010). Therefore, many capacity estimation methods have been developed over the years. Minderhoud et al. (1997) provided an early overview of capacity computation methodologies based on collected data (Figure 17). Although the research is over twenty years old, the capacity computation methodologies explored in this study still form the basis of current and future capacity estimation methods. The estimation methodologies were judged based on criteria such as computational demand, the viability of the theory, and expected uncertainties of the outcomes. Positively labelled methodologies will be explored below with the use of case studies to investigate which are useable for the data analysis, including advantages and disadvantages.

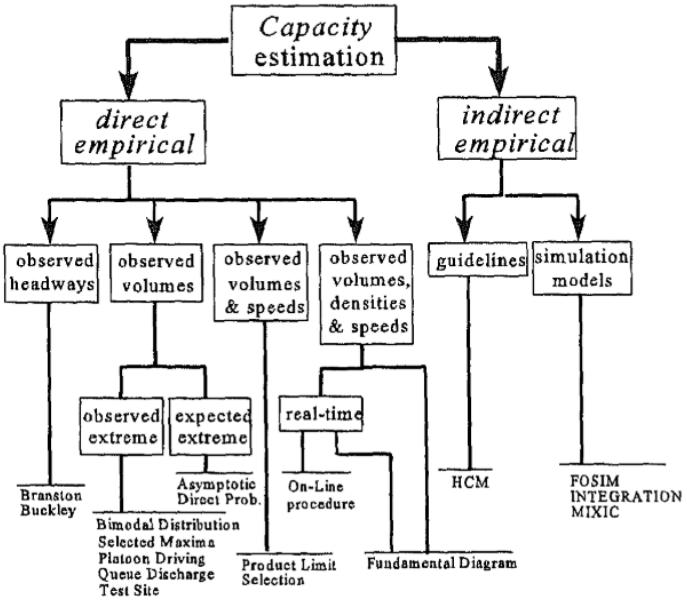


Figure 17 - Overview capacity computation methodologies (Minderhoud et al., 1997)

2.3.2.1 Estimation composite headway models

The composite headway model uses measured headways to determine the distribution of the minimum headway, leading to a deterministic capacity value. The Van Aerde model describes the speed-flow-density relationship based on a simple car-following equation (Shojaat, Geistefeldt, Parr, Escobar, & Wolshon, 2018). The distance headway between consecutive vehicles (h) only depends on free-flow speed (s_f), the current speed (s), and three parameters (c_1,c_2,c_3). The approach does not require capacity observations:

$$d = \frac{1}{h} = \frac{1}{c_1 + \frac{c_2}{s_f - s} + c_3 \cdot s} \qquad \text{Eq. 1}$$

where

- d = density (veh/km)
- h = distance headway between consecutive vehicles (km)
- s_f = free flow speed (km/h)
- c_1 = fixed distance headway parameter (km)
- c_2 = first variable headway parameter (km^2/h)
- c_3 = second variable distance headway parameter (h^{-1})
- s = speed (km/h)

Estimating the model parameters is done based on assuming key traffic flow variables (i.e. capacity, free-flow speed, the speed at capacity, and jam density). Non-linear regression om the speed-density-volume plot is used in an iterative way to minimize the sum of squared errors for the dependent variable. As it remains unclear which variable is dependent and which is independent, the orthogonal sum of squared errors can be minimized as an unbiased compromise using multivariate calibration. Subsequently, the capacity can be computed as the apex volume of the speed-flow diagram (Shojaat et al., 2018). The downside is that no distinction between the PQF and QDF can be made which makes this method unusable in estimating the capacity drop.

2.3.2.2 Fitting a fundamental diagram

The fitting of a fundamental diagram is used to estimate the capacity value by identifying the relationship between speed and density. Some of the most elementary shapes are the Greenshields, Triangular, Truncated triangular, Smulders, Drake, Inverse Lambda, and Wu fundamental diagram. Depending on the shape, the PQF and QDF can be estimated by shifting the diagram in such a way that it best fits the measurement. The downside is that empirical data often, e.g. Figure 6 - Capacity drop phenomena explained (Calvert, Taale, & Hoogendoorn, 2016), has a distinctive shape, slightly different for each particular road section. Therefore, a one-size-fits all approach would not work if multiple road sections are investigated.

2.3.2.3 Queue discharge distribution method

The queue discharge distribution method is based on the notion that available measurements (downstream of a bottleneck) of flow can be divided into measurements representing (Daamen, 2010):

- Traffic demand (free flow)
- Capacity state of the road (congestion)
- Capacity state measurements representing the capacity state of an active bottleneck upstream of the measurement site (free flow at the measurement location, but congestion upstream an upstream bottleneck)

In the case of the latter, the capacity of an activated bottleneck can only be measured only if the upstream state has been classified. The queue discharge distribution method is straightforward and unbiased. An example of syntactic data is provided in Figure 18. An important drawback is that the approach does not use non-capacity observations.

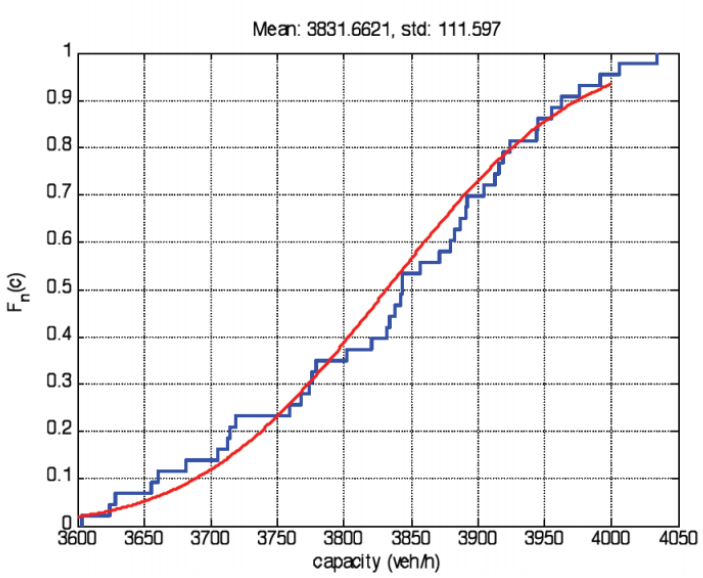


Figure 18 - Queue discharge capacity distribution (Daamen, 2010)

2.3.2.4 Use of cumulative curves

The use of cumulative curves is similar to the queue discharge distribution method. The slope of the curves determines the queue discharge rate which can be compared to the pre queue flow to estimate the capacity drop. However, assumptions should be made on key traffic variables to indicate differences in traffic state.

L. Zhang and Levinson (2004) used a density above 39 veh/lane/km as congested, and below 31 veh/lane/km as uncongested based on experiments in which congestion prediction by the two thresholds was confirmed after visual inspection. The bottleneck was considered active when the upstream detector measured densities above the congestion threshold for 5 minutes and the downstream detector measures densities below the uncongested threshold. For illustration, see Figure 19. A pre queue transition period was observed in stage 2 as the upstream occupancy increases, followed by a queue discharge period in stage 4 with a density above the congested threshold. Stages 3 and 5 were removed from the analysis as they fall in between both thresholds. The method with two occupancy thresholds was considered successful and constituted the algorithm detection of active bottlenecks. Furthermore, the researchers found that insufficient capacity of a diverging bottleneck was often the case for activation of a bottleneck. If occupancies higher than 25% were measured at these ramps, they were removed from the capacity drop computation.

The researchers also investigated whether the QDFs followed a normal distribution once averaged across a long period, leading to the reasonable assumption that daily QDFs are also normally distributed. These hypotheses were accepted by the Shapiro-Wilk test at the 0.05 level. The variability day-to-day should be explained by demand patterns, traffic composition, driver population, and weaving volumes. Hence, the researchers adjusted the streak of vehicles to PCEs to test whether the normality and variation could be explained by traffic composition. However, the results showed an even better fit for the normal distribution, which contradicts the conjecture of this research. It should be noted that the researchers measured an overage of about 2-8% heavy vehicles, which is way lower compared to some truck penetration seen now on the Dutch road network. Whether these findings hold with increased truck penetration seen now on the Dutch road network. Whether these findings hold with increased truck penetration should follow from the data analysis.

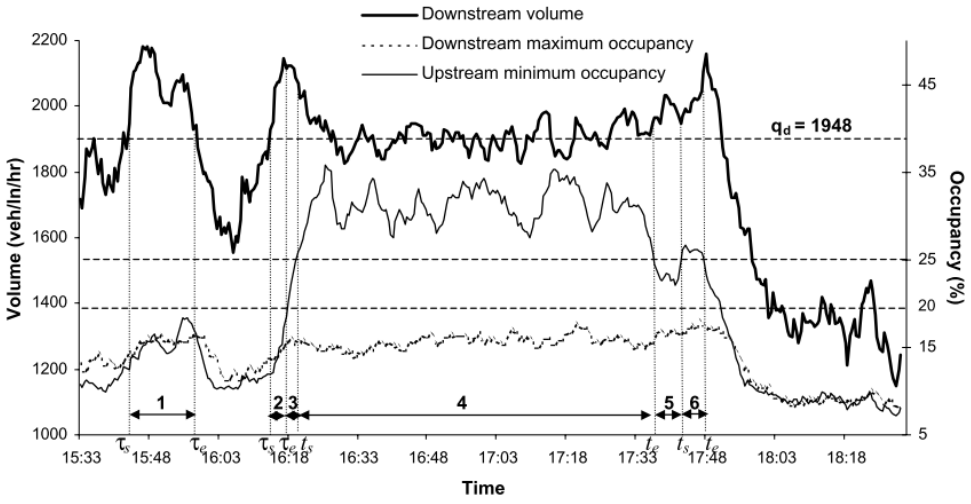


Figure 19 - Application of method with two occupancy thresholds to Bottleneck 5, Nov. 21, 2000 (L. Zhang & Levinson, 2004)

Cassidy and Rudjanakanoknad (2005) display the capacity drop by transformed curves of cumulative vehicle count, N, vs time, t, at locations [X_1, …, X_4]. Each N-curve was shifted horizontally to the right so that these V-curves show the "virtual" departures past the last measurement location, X_4. By doing this, the vertical displacements between the V-curves show the excess vehicle accumulations due to delays at measured times. To visualize this phenomenon, the authors plotted the V-curves on an oblique coordinate system to amplify changes in the slope (Figure 20) by reducing background noise:

$$O(t) = V(t) - q_0 * (t - t_0) \qquad \text{Eq. 2}$$

In which:

$V(t)$ = cumulative virtual vehicle count at time t
q_0 = specified rate close to estimated QDF capacity (=2300 vphpl in this example)
t_0 = start time t

Subsequently, key traffic features can be obtained from these O-curves. An uncongested high flow of 10.480 veh/h can be observed between t = 6:07 and 6:11, with virtually no vertical displacements. From 6:13 and onwards, the vertical displacement between X_1 and the other measurement locations becomes visible. In this case, the breakdown occurred due to a queue that formed on the shoulder lane. The queue discharge rate fell by 10% after the breakdown as indicated by the dashed lines.

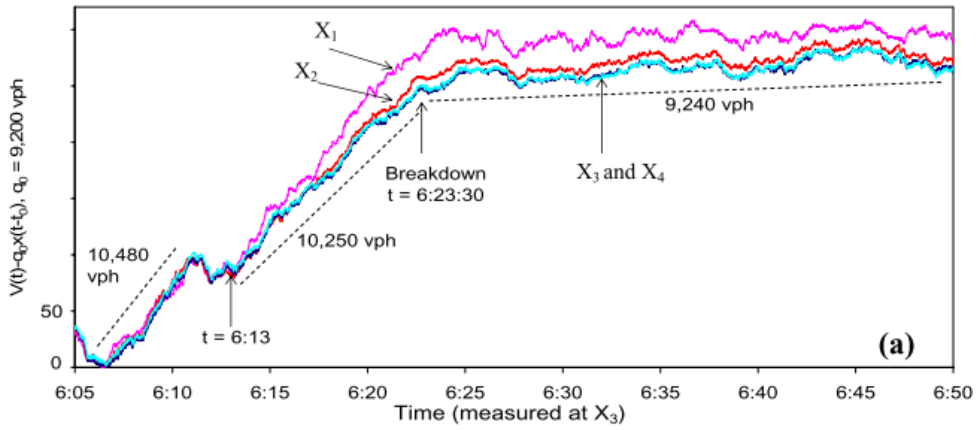


Figure 20 - O-curves at X_1 through X_4 (Cassidy & Rudjanakanoknad, 2005)

Oh and Yeo (2012) developed a systematic approach for finding capacity drops based on measured speeds. A 5-min window average for the 30-sec measurements was used to define the maximum capacity value. The researchers compared speed data upstream and downstream to identify bottleneck location, activation time, bottleneck duration, and traffic state. Four traffic states were identified:

1. Free-flow state:

$$V_{u,} \cap V_d \geq 50\ mph$$

2. Transition to bottleneck (or recovery from bottleneck):

$$40\ mph < V_u < 50\ mph \qquad V_d \geq 50\ mph$$

3. Bottleneck:

$$V_u \leq 40\ mph \qquad V_d \geq 50\ mph$$

and

$$(SD\ V_u + SD\ V_d)\ for\ 15\ \min(t-1, t, t+1) \leq 5mph$$

The authors used re-scaled cumulative curves, which are slightly different from the O-curves shown in Figure 20. Two measurement locations are used, one upstream including the on-ramp (VDS 801552) and one downstream (VDS 811408). The different states are marked by the vertical red dashed lines in Figure 21, corresponding to the speed thresholds mentioned. The maximum flow measured was about 7284 veh/h then followed by bottleneck activation, leading to a discharge flow of 6120 veh/h. The flow and speed plots were used to avoid error accumulation caused by mismatches due to detector errors as the curves do not meet

until the end. Thus, the re-scaled cumulative curves are very useful in determining the difference in slope between the upstream and downstream detector.

The individual capacity drop per lane was also computed, to investigate whether differences on a cross-section of a road exist. For 2-lane to 5-lane highways, they investigated for each lane the drop rate. It was found that the drop rate increases per lane from the shoulder lane up to the median lane. Therefore, the researchers concluded that the capacity drop is not significant on the outer lanes. Unfortunately, the researchers did not mention the actual truck percentage measured despite the use of microscopic data which could be used for measurement purposes. However, it was noted that the outer lanes contained many trucks and lane-changing vehicles.

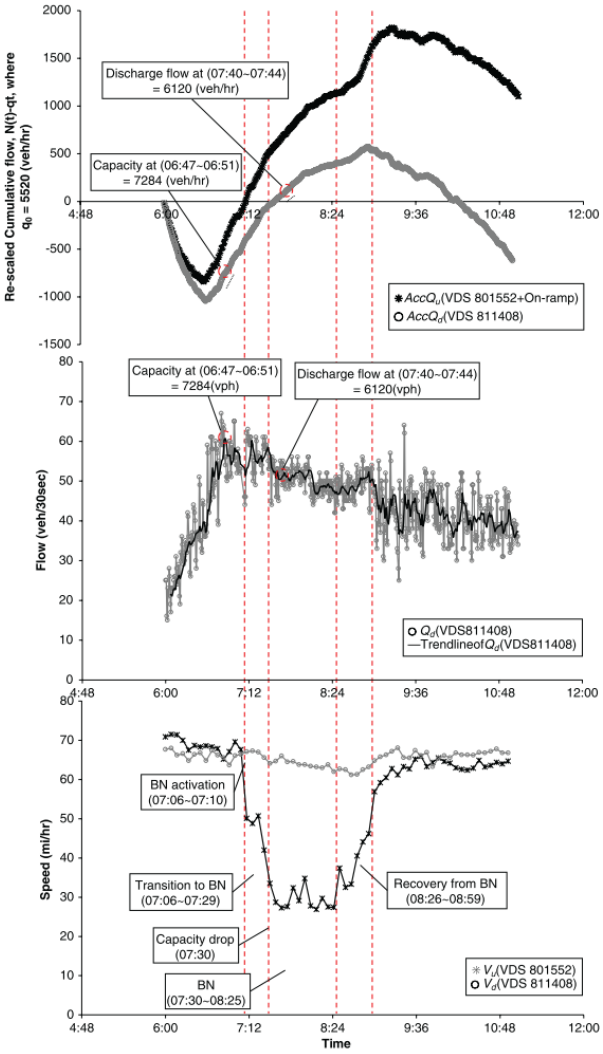


Figure 21 - Analysis at SR-91 East, rescaled cumulative flow [N(t) – q0t]; flow (Qd); and speed (Vu, Vd) (Oh & Yeo, 2012)

2.3.2.5 Selected maxima method

The selected maxima method aims to combine different measurements in congested states over a longer period to estimate a capacity distribution. This method is useful in smoothing out individual effects which may lead to different PQFs or QDFs but requires a large amount of data. The capacity q_c is assumed to be equal to the averaged traffic flow maxima observed during the total observation period (Minderhoud et al., 1997):

$$q_c = \sum_i \frac{q_i}{n} \qquad \text{Eq. 3}$$

where

q_c = capacity value (veh/h)
q_i = maximum flow rate observed over period i
n = number of cycles
i = length of cycle (period over which a maximum flow rate is determined)

2.3.2.6 Bimodal distribution method

When observed traffic flow includes intensities at about the point of capacity of the road, a bimodal distribution may be observed. Minderhoud et al. (1997) combined measurements at a single cross-section with many low intensities during the night (curve I) and with high intensities during day and night (curve II) in Figure 22. Data collected during the day can be depicted as a Gaussian curve. The general form for a compound probability density function can be used to estimate the capacity. Its value is estimated as the expectation of the mean by probability density function b(q):

$$f(q) = \varphi \cdot g(q) + (1 - \varphi) \cdot b(q) \qquad \text{Eq. 4}$$

where

ϕ = fraction of the probability density function representing traffic demand below capacity
g(q) = probability density function representing the traffic demand below capacity
b(q) = probability density function representing the capacity state

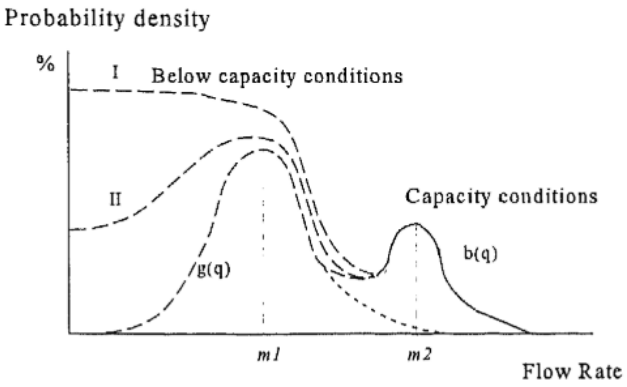


Figure 22 - Bimodal distribution of intensities (Minderhoud et al., 1997)

The downside of this method is the below-capacity probability density function. Although capacity can be estimated with a normal, Gaussian-type distribution, the free flow distribution can be doubted and depends on the observation period chosen.

2.3.2.7 Kaplan-Meier approach (parametric and non-parametric)

The Kaplan-Meier approach combines under-capacity measurements with capacity measurements to estimate the capacity. The method, based on capacity distribution function estimation, is proposed by Brilon, Geistefeldt, and Regler (2005). It builds upon the stochastic concept of capacity and takes into consideration that traffic is more likely to break down when the flow is high than when the flow is low. So, low-capacity observation bias is removed by this method. Therefore, the mean queue discharge rate can be better approximated with the Kaplan-Meier approach instead of the queue discharge distribution method.

Brilon et al. (2005) proved for German highways that empirical distribution of capacity, at the 5-minute level measured over many sites and several months, clearly shows a Weibull distribution. This was identified using the non-parametric method to estimate the survival function, the so-called Product Limit Method of Kaplan-Meier:

$$F_c(q) = 1 - \prod_{i:q_i \le q} \frac{k_i - d_i}{k_i}; i \in \{B\} \qquad \text{Eq. 5}$$

where

- $F_c(q)$ = distribution function of capacity c
- q = traffic volume
- q_i = traffic volume at interval i
- k_i = number of intervals with a traffic volume of $q \geq q_i$
- d_i = number of breakdowns at a volume of q_i
- (B) = set of breakdown intervals (see below)

Using this equation, each observed traffic volume q is classified according to

B: Traffic is fluent in time interval i, but the observed volumes cause a breakdown, i.e. the average speed drops below the threshold speed in the next time interval i + 1.

F: Traffic is fluent in interval i and the following interval i + 1. This interval i contains a censored value. Its information is that the actual capacity in interval i is greater than the observed volume q_i.

C1: Traffic is congested in interval i, i.e. the average speed is below the threshold value. As this interval i provides no information about the capacity, it is disregarded.

C2: Traffic is fluent in interval i, but the observed volume causes a breakdown. However, in contrast to classification B, traffic is congested at a downstream cross-section during interval i or i – 1. In this case, the breakdown at the observation point is supposed to be due to a tailback from downstream. As this interval i does not contain any information for the capacity assessment at the observation point, it is disregarded.

However, a complete distribution function could not be estimated since the highest q-values observed were not followed by a breakdown. A Weibull distribution was tested positive. To estimate the parameters of the distribution function, the maximum likelihood technique was used. The likelihood function is given by:

$$L = \prod_{i=1}^{n} f_c(q_i)^{\delta_i} \cdot [1 - F_c(q_i)]^{1-\delta_i} \qquad \text{Eq. 6}$$

where

- $f_c(q_i)$ = statistical density function of capacity c
- $F_c(q_i)$ = cumulative distribution function of capacity c
- n = number of intervals
- δ_i = 1, if uncensored (breakdown of classification B)
- δ_i = 0, elsewhere

The likelihood function must be maximized to calibrate the parameters of the distribution function. The Weibull distribution function is:

$$F(x) = 1 - e^{-\left(\frac{x}{\beta}\right)^{\alpha}} for\ x \geq 0 \qquad \text{Eq. 7}$$

where

- α = shape parameter
- β = scale parameter

The mean and variance of the Weibull distribution are given by:

$$\mu = \beta * \Gamma\left(1 + \frac{1}{\alpha}\right) \qquad \text{Eq. 8}$$

$$\sigma^2 = \beta^2 \cdot \left\{ \Gamma\left(1 + \frac{2}{\alpha}\right) - \left[\Gamma\left(1 + \frac{1}{\alpha}\right)\right]^2 \right\} \qquad \text{Eq. 9}$$

where Γ(x) equals the Gamma function at point x.

If the measurements only contain capacity observations, one will find similar results for the non-parametric Kaplan-Meier estimation and the queue discharge rate method. However, difficulties arise due to the capacity drop. It should be noted that simply applying the Kaplan-Meier approach to the two-capacity case does underestimate the queue discharge flow (Daamen, 2010). The standard deviation of the capacity estimation is also found to be very large. The queue discharge rate method yields values much closer to the actual values in that case. To overcome this problem, Regler (2004) developed a method comparable to the Product Limit technique to obtain a capacity distribution in queue discharge flow. This method is based on the following assumptions:

- Flow rates during congestion do not represent the maximum possible flow of congested traffic.
- Capacity can directly be measured in queue discharge flow, i.e. in the last 5-minute interval before recovery of traffic flow (uncensored data)
- Flow rates during congestion are lower than capacity in queue discharge flow (censored data).
- Flow rates in free flow are not relevant for the capacity of queue discharge flow.

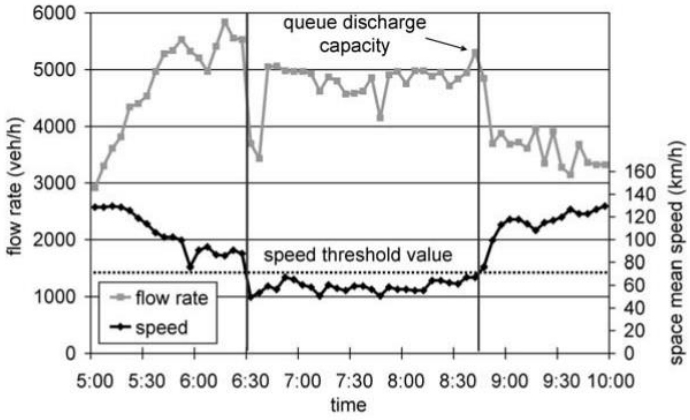


Figure 23 – Flow rate and speed time series during congestion (freeway A5, 5-minute intervals) (Regler, 2004)

Based on these considerations, the observed flow rates in each time interval can be classified as follows:

B*: Traffic recovers from congestion to free flow, i.e. the average speed exceeds the threshold value from time interval i to interval i + 1.

F*: Traffic is congested in intervals i and i + 1, i.e. the average speed is lower than the threshold value in both intervals. This interval i contains a censored value.

C*: Traffic is fluent in interval i, i.e. the average speed is above the threshold value. This interval is not relevant.

After this classification, a capacity distribution for queue discharge flow according to $F_c(q)$ where (B) is to be replaced by (B*) can be computed from empirical data. This capacity level turned out to be always lower than the capacity before breakdown obtained from $F_c(q)$. The difference between both distributions (Figure 24), represented by the median value was regarded as the capacity drop.

An overview of data needs, traffic states identified, capacity measured, and type of capacity (free flow and/or queue discharge) is visualized in Table 6.

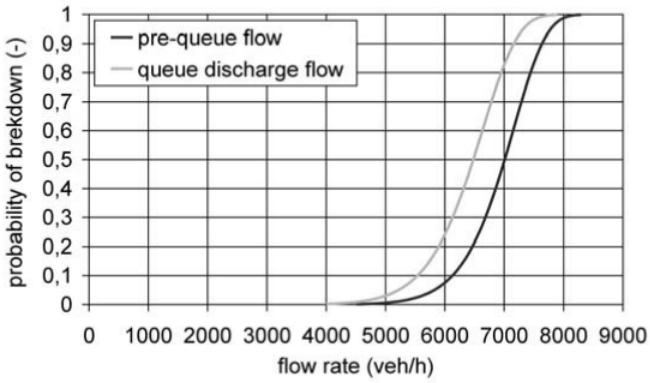


Figure 24 - Capacity distribution for pre-queue and queue discharge flow (freeway section A5-7, 11.8% average truck percentage, 5minute intervals) (Regler, 2004)

Table 6 - Comparison of capacity estimation methods (extended Minderhoud et al. (1997))

Method	Data Needs				Traffic State[1]		Capacity		Type[2]
	h	q	v	k	(Q)	(C)	q_c	F(q)	
Composite Headway Models	Yes				Yes		Yes		1
Fitting a Fundamental Diagram		Yes	Yes	Yes	Yes		Yes		d
Queue Discharge Distribution Method		Yes	Yes			Yes		Yes	2
Cumulative Curves		Yes	Yes		Yes	Yes	Yes		d
Selected Maxima Method		Yes			Yes	Yes		Yes	m
Bimodal distribution		Yes			Yes	Yes		Yes	m
Product Limit Method (Kaplan-Meier)		Yes	Yes		Yes	Yes		Yes	m, d

[1] *(Q) represent free flow intensities, (C) represents congested flow intensities*

[2] *Type 1 denotes a capacity value estimation representing the maximum free flow intensity, Type 2 denotes a capacity value estimation representing the maximum congested flow intensity, m stands for type 1 and type 2 mixed into one capacity estimate and d stands for the dependency with the study set up (either type 1 or 2 is possible)*

To compute the capacity drop, both the PQF and QDF should be computed. Therefore, two methods seem to be viable for this specific analysis. The use of cumulative curves can compute a deterministic value for PQF and QDF, with the use of just one method. The Kaplan-Meier (PLM) approach is also feasible for computing the PQF and QDF and does this in a stochastic way but requires two different approaches.

2.3.3 Site Selection Criteria

2.3.3.1 Measurement locations

Based on the computation methodology used, measurements as their location are required. Daamen (2010) describes quantitatively how to compute the capacity based on either traffic volumes or headways. Optimally, four measurements locations are needed when computing the capacity based on traffic volumes, viz:

1. Upstream of the bottleneck that will not be reached due to overloading of the bottleneck
2. Close to the bottleneck that will be reached if congestion sets in
3. Inside the bottleneck
4. Downstream of the bottleneck

An overview of how the measurement stations should be located is visualized in Figure 25. Stations at possible discontinuities such as on- or off-ramps are useful to estimate the ramp flow and compute the mainline – ramp flow ratio. Depending on the methodology, more or fewer detectors are needed to quantify the traffic state. It should be noted that the activation of the bottleneck should not be caused by spillback of earlier bottlenecks, which is measured by the downstream detector. In that case, the bottleneck should be neglected for capacity computation.

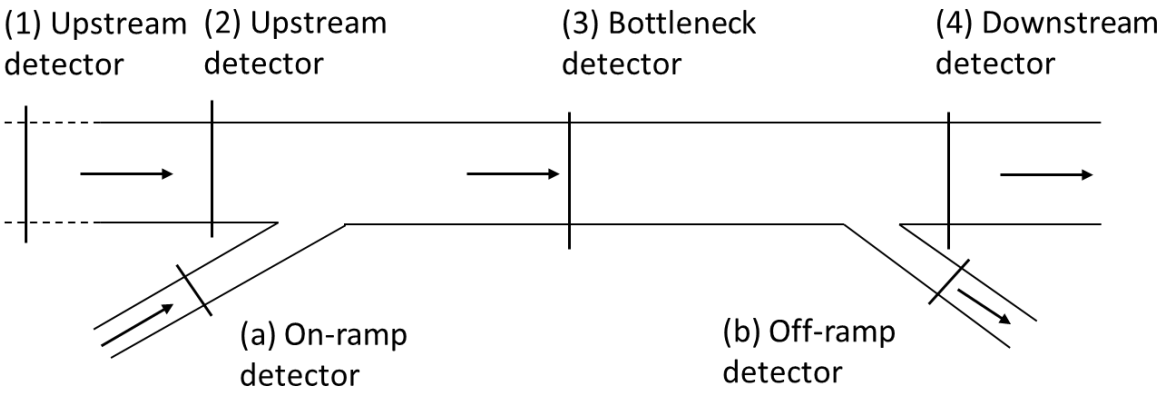


Figure 25 - Overview measurement locations for capacity

The location of the loop detectors is highly important. To determine the impact of a single bottleneck it is required to have detectors close to the activated bottleneck. To illustrate this, L. Zhang and Levinson (2004) only considered 27 freeway segments appropriate while hundreds of candidates were available. Loop detectors appropriate are usually spaced out approximately every 800 meters, with additional detectors at merging and diverging areas so that uniform flow characteristics from individual segments could be measured.

2.3.3.2 Minimization external factors

The period of data collection also varied among the studies. Some researchers limited their measurement period to only a few rush periods (Chung et al., 2007; Leclercq, Laval, & Chiabaut, 2011), while others used more extensive measurement periods up to 60 days, without excluded data (Banks, 2006). Studies containing data from multiple years are scarcer in the literature. However, more interesting are the long-term analysis studies regarding the exclusion of other variable systematic factors influencing freeway capacity (Geistefeldt, 2009; Shiomi et al., 2019). The following impacts were accounted for in comparable studies:

- Weather conditions
- Work zones
- Accidents and incidents
- Driver population

The inclusion of weather conditions varied per study. Most researchers looked at rainfall and removed days that contained precipitation. Geistefeldt (2009) also removed temperatures below the freezing point, as icing might occur. Intervals measured during darkness were omitted from the data so that the samples only represent dry and daylight conditions.

As for short measurement studies, road work-related could be easily removed. However, for long-lasting studies, this data is often not available. Therefore, unusual circumstances were removed such as reduced speeds or closed lanes. Also, traffic breakdowns at low volumes (pre-breakdown flow of less than 1200 veh/h) were excluded (Geistefeldt, 2009). In case traffic detectors measured implausible values, that specific data was also neglected (Shiomi et al., 2019).

Accounting for the driver population is a more difficult task. Shiomi et al. (2019) did this after the analysis by evaluating the holiday months in Japan – August, and December. Whereas Geistefeldt (2009) did this in a more systematic way by classifying data as either morning peak (Monday through Friday, 5:00 – 9:00) evening peak (Monday through Friday, 15:00-19:00), or as weekend and holiday combined (8:00 – 20:00). In this way, different driver populations could be separated.

The impact of technological advancements on traffic flow has been described in Appendix F. An exposition of impact from in-vehicle and infrastructure perspective due to new implemented technologies is described there.

2.3.4 Summary and Discussion

Many methods to compute capacity based on the purpose (design, strategic and operational) exist. Operational capacity methods are most relevant in this case and summarized by Minderhoud et al. (1997). To compute the capacity drop, a clear distinction should be made between the pre queue flow (PQF) and the queue discharge flow (QDF), as the difference between the two flows indicates the drop rate. Therefore, only methods that are classified as a 'Type d' estimation methodology will be appropriate. An overview of these and other capacity computation methods can be found in Table 6.

This narrows down the number of appropriate methodologies to three: Fitting a fundamental diagram, the Use of Cumulative Curves, and the Product Limit Method. For the first one, a fundamental diagram can be fitted onto the empirical data based on the assumption of a fundamental shape. The second uses occupancy during breakdown plotted against time, either with regular cumulative curves (L. Zhang & Levinson, 2004), transformed cumulative curves (Cassidy & Rudjanakanoknad, 2005), or re-scaled cumulative curves (Oh & Yeo, 2012). A threshold density should be estimated to classify the breakdown event and separate the PQF from the QDF. A deterministic capacity drop value can subsequently be computed. The third one, The Product Limit Method combines under capacity estimations with capacity estimations to come to a stochastic capacity distribution (Brilon et al., 2005). Regler (2004) extends the methodology by defining specific traffic states so that pre queue flow could be distinguished from queue discharge flow. Multiple variables should be estimated such as the speed-threshold, as well as the distribution form in case of a parametric estimation. For the Weibull distribution to fit the non-parametric distribution, the α and β parameters should be estimated.

An optimal measured observation location is visualized in Figure 25. This includes two upstream detectors, bottleneck detector, downstream detector, and possibly ramp detectors. It should be assured that spillback is non-existent as this will influence the activation of the bottleneck. The main exclusion of other variables in long-term studies (Geistefeldt, 2009; Shiomi et al., 2019) where should be focussed on are weather conditions, work zones, accidents and incidents; and driver population.

As illustrated, several computation methods exist to identify capacity and subsequently the capacity drop. For this study, one of the three methods will be picked as the basis for the Calculation Framework. As for the first method, empirical data often shows no clear distinctive shape and requires visual validation. As for the scale of the data analysis, this methodology is not appropriate.

Now, two methods remain which both can be used for large data analysis: The Use of Cumulative Curves and the Kaplan-Meier Product Limit Method. The stochastic approach of the Kaplan-Meier Product Limit Method fits the chaotic properties of capacity estimation better in comparison to the use of cumulative curves. The main problem lies in properly distinguishing the PQF from the QDF. As the PQF depends on both censored and uncensored measurements, the estimation might underestimate the flow rate (Daamen, 2010). The QDF on the other hand seems to accurately represent the discharge flow values. The precision in estimating the capacity drop from the Kaplan-Meier Product Limit Methods seem to be contradicted by researchers (Daamen, 2010; Regler, 2004) but is proven to be appropriate for estimating the pre queue capacity distribution and the queue discharge flow distribution in large-scale data analysis (Geistefeldt, 2009). Therefore, the Kaplan-Meier Product Limit Method will be the basis for the Calculation Framework (3.4) in the Empirical Analysis (3).

2.4 Overview Simulation Models

Multiple microscopic traffic flow models have been developed, each with a different underlying car-following and lane change model to describe the longitudinal and lateral driving behaviour respectively (Broekman, 2017). First, current models will be elaborated regarding longitudinal and lateral behaviour in 2.4.1. Subsequently, the models will be compared based on 2.4.2. Lastly, the results will be summarized in 2.4.3 and a choice for the simulation model will be made which will be used in chapter 4.

2.4.1 Microscopic Traffic Flow Simulation Models

Microscopic Traffic Flow Simulation models are used to describe traffic flow, based on estimating behaviour of single vehicles (Maciejewski, 2010). The purpose of such a model in this study is to fill the gaps of the data analysis as external factors can at most only be reduced within the empirical data. Simulation models differ mainly on the underlying longitudinal behaviour and lateral behaviour models implemented. An overview of criteria and functionalities was created by Van Beinum et al. (2020) to compare different simulation models, see Table 7. The following categories: manoeuvres, microscopic behaviour, macroscopic effects, and general model criteria were investigated for the most widely used models: AIMSUN, CORSIM, MITSIM, MOTUS, PARAMICS, VISSIM, and OpenTrafficSim.

Table 7 - Overview of criteria and functionalities (Van Beinum et al., 2020)

No.	Category	Criteria and functionalities
1	Manoeuvres	The locations of lateral movements around ramps and in weaving segments are based on latent tactical plans by the driver, which incorporate pre-allocation and gap selection
2		The lateral movement is anticipative
3		The lateral movements are influenced by the rule to keep right
4	Microscopic behaviour	The urge to change lanes increases as the remaining length for changing lanes decreases
5		The longitudinal/lateral interactions include cooperation and synchronization
6		The longitudinal/lateral interactions include relaxation
7		Maximum acceleration and deceleration related to vehicle limitations
8	Macroscopic effects	Desired speeds for each vehicle are distributed realistically
9		Lane selection is based on the desired speed
10	General model criteria	The number of model parameters for calibration preferably small
11		Model is open source so that simulated behaviour can be explained

However, only the models VISSIM, MOTUS, and its successor OpenTrafficSim are accessible for this research. Therefore, the main aspects of these models will be shortly elaborated.

VISSIM uses a psycho-physical model developed by Wiedemann (1991) for its longitudinal driving behaviour. This model is characterised by different driver states e.g. free driving, closing in, car-following, and emergency which determines the drivers' reaction (Broekman, 2017). The lane change model is based on the work of Willmann (1978) and Sparmann (1979) and the location of a vehicle is affected by two variables: the distance of the lane change and the distance of an emergency stop. The car following behaviour can be adjusted with 59 parameters, which makes the calibration rather complex (Van Beinum et al., 2020). Besides, the source code is not published in an open-source way which leads to less transparency in the exact interaction of the parameters.

MOTUS uses the IDM+ longitudinal model (Treiber, Hennecke, & Helbing, 2000), which estimates the acceleration nor deceleration of the driver based on the difference in speed between the leader and the difference between the current and desired speed. The lane change model implemented, is the LMRS model

(Schakel, Knoop, & Van Arem, 2012). An underlying lane change desire is hypothesised and the execution is subsequently adapted with synchronization and/or relaxation (Van Beinum et al., 2020). The acceleration of the driver also plays an important role in gap acceptance. The source code is open-source and contrary to VISSIM is suitable for any adjustments as desired.

OpenTrafficSim is also an open-source microscopic traffic flow simulator, which is currently developed by the TU Delft. It is based on identical driver models implemented in MOTUS but includes social interactions between road users, which leads to more realistic headway distributions during car-following and increases the number of lane changes (Van Beinum et al., 2020).

2.4.2 Comparison of Models

A comparison of the different criteria described in Table 7 has been made by Van Beinum et al. (2020), see Table 8. The decisive criteria of OpenTrafficSim are not there yet but are similar to MOTUS in its current state. The difference between VISSIM, MOTUS, and OpenTrafficSim lies within the number of parameters and adaptability of the software. For this research, the simulation package should be easy to adapt on the following criteria: percentage of heavy vehicles, acceleration of heavy vehicles, and length of the heavy vehicles. The goal of the simulation stage is not to create new behavioural models, but to complement the gaps of the data analysis and investigate the impact of the different circumstances on the capacity drop.

Table 8 shows that these models do not differ in criteria 1 – 9. However, based on a study on the advantages and limitations of these simulation packages, VISSIM appears to be the least favorable option because many features are hidden for the user, the complexity is the highest and the platform is not open source (Broekman, 2017). Between MOTUS and OTS, OpenTrafficSim is the more extensive software of the two but requires more programming knowledge, lacks an extensive manual, and is not finished as of yet. However, after some practical experience, it is believed that OTS offers easy adaptability in the variables that will be adapted. The lack of an extensive manual is compensated by the help of a PhD student who is willing to help. Therefore, OpenTrafficSim is chosen as a simulation package in chapter 4.

Table 8 - Overview of microscopic simulation software package criteria (Van Beinum et al., 2020)

No.	Criteria	AIMSUN	CORSIM	PARAMICS	VISSIM	MITSIM	MOTUS
1	Latent tactical plans	Yes	No	-	No	Yes	No
2	Anticipation	No	No	-	No	No	No
3	Keep right	Yes	No	-	Yes	Yes	Yes
4	Lane change desire	Yes	Yes	-	Yes	Yes	Yes
5	Synchronization	No	No	No	Yes	Yes	Yes
6	Relaxation	No	No	Yes	Yes	Yes	Yes
7	Acceleration	Yes	-	Yes	Yes	Yes	Yes
8	Desired speed	Yes	Yes	Yes	Yes	Yes	Yes
9	Lane selection	Yes	Yes	-	Yes	Yes	Yes
10	Number of parameters	Small	Small	Small	Large	Large	Small
11	Open Source	No	No	No	No	Yes	Yes

2.4.3 Summary and Discussion

Microscopic Traffic Flow Simulation models are used to describe traffic flow, based on estimating the behaviour of single vehicles (Maciejewski, 2010). The way this behaviour is estimated differs per simulation package and can be divided into longitudinal and lateral behaviour models. Some of the most used models are AIMSUN, CORSIM, MITSIM, MOTUS, PARAMICS, VISSIM, and OpenTrafficSim. However, not all models are accessible for this research and will not be considered as viable options for the simulation in chapter 4. The three accessible models are VISSUM, MOTUS, and OpenTrafficSim, whereby OTS is built upon identical driver models as MOTUS and can be seen as the unfinished successor. A study investigated how these software packages fulfil important criteria and functionalities, e.g. manoeuvres, microscopic behaviour, and macroscopic effects. The three simulation packages are found to be equivalent in performance (Van Beinum et al., 2020). However, the high complexity, black-box appearance, and lack of open source are to the detriment of VISSUM. OTS on the other hand requires programming knowledge, lacks an extensive manual, and is not finished yet. However, the possibility of guidance in OTS from a PhD student within the ToGrip project pushes the decision to use OTS for this research.

The main purpose of the simulation stage is not to create or adapt existing driver models, but to fill the gaps of the data analysis stage. This can be done by simulating different scenarios whereby the penetration of heavy vehicles, operational characteristics, and length of the heavy vehicles are changed and subsequently research the impact on the capacity drop. As it is impossible to diminish all side effects in the empirical data, these effects can be diminished in a simulation. Moreover, different less realistic scenarios can be tested, e.g. a very high heavy vehicle percentage, to see how this would impact the capacity drop. The combination of empirical data analysis and modelled simulation data should give insight into how heavy vehicles impact the capacity drop, and how changing circumstances will affect this.

3 EMPIRICAL ANALYSIS

First, the structure of the empirical analysis will be explained in 3.1. Next, a set of site selection criteria is described, and the sites chosen described in 3.2. The collection and modification of empirical data are elaborated in 3.3. Subsequently, the calculation framework is designed step-by-step in 3.4. The results of the empirical data in the calculation framework are shown in 3.5. Finally, the main conclusions are discussed in 3.6

3.1 Methodological Approach

The main goal of this chapter is to investigate the impact of freight traffic on the capacity drop. Therefore, the following research question was asked:

Research Question 2: **How and to what extent does freight traffic impact the capacity drop on highways currently?**

The activities related to the Data Analysis research stage are displayed in Figure 26 Based on the Site Selection Criteria (2.3.3) specific requirements are added based on data availability in 3.2. Also, the sites selected are described. The loop-detector data from the Netherlands collected by NDW and processed and modified by TNO are prepared (3.3) and used in the Flow Type Classification Model, as the first step in the Calculation Framework (3.4). Subsequently, the data is analysed via the Kaplan-Meier Product Limit Method based on the methodologies used by Brilon et al. (2005) and Geistefeldt (2009). From this, the capacity drop can be computed. Besides, the Passenger Car Equivalency has been estimated for the selected sites. Finally, statistical tests are proposed to test the results. Finally, the results can be observed in 3.5.

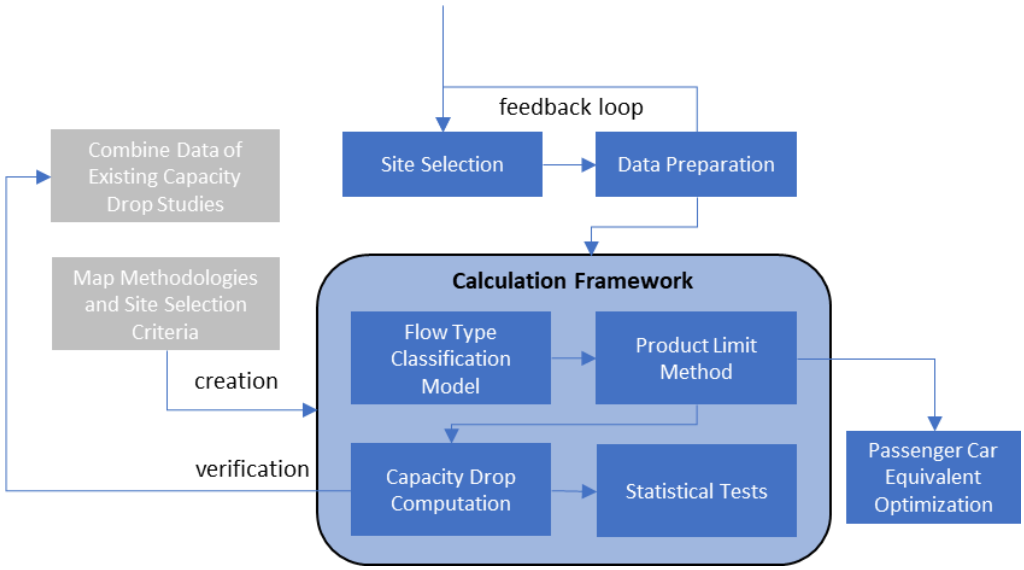


Figure 26 – Partial Research Framework (Data analysis stage)

3.2 Site Selection

In section 2.3.3, site selection criteria used in previous studies were elaborated. Important findings were the spatial location of the loop detectors, and exclusion of external factors, e.g. weather conditions, work zones, accidents, and driver population. Therefore, a decisive set of selection criteria used in this study is determined in 3.2.1. This is separated in selection criteria to get comparable roads in terms of conditions and criteria to minimize the effects of external factors. Hereafter, the selected sites are visualized and described in terms of conditions.

3.2.1 Selection Criteria

To select freeway sections for analysis, site conditions are characterized such that a set of comparable sites is selected. Multiple comparisons can be made to test how specific conditions affect the capacity drop, e.g. by varying the number of lanes (Oh & Yeo, 2012). For this study, the focus will be on picking roads with different traffic compositions. Therefore, it is necessary to know the road conditions when selected different sites. For example, a bottleneck at a road with 5-lanes is expected to have a different traffic flow phenomenon occur compared to a road with a slope and a 2-lane bottleneck. To categorize these conditions, three types of conditions are distinguished for the identification of the sites:

- Geometric conditions
- Traffic conditions
- Control conditions

The geometric conditions tell something about the location of the loop detectors, the number of lanes, the type of bottleneck, and the grade of the highway. For a location to be suitable, at least one loop detector is needed to measure the actual flow and speed at the bottleneck, while a nearby downstream detector is needed to measure whether back-propagating congestion occurs. These detectors should not have other bottlenecks in between. The number of lanes is investigated before (Oh & Yeo, 2012), and ranges from 2 to 5 lanes for Dutch highways. Different lane configurations are chosen for the freeway sections. However, it turned out that highways with 3-5 lanes do often not have loop detectors to measure the number of heavy vehicles at each lane. Besides, way less congestion was observed compared to some of the analysed 2-lane roads which would base the estimation of capacity drop at only a few single days. Therefore, the focus will be on 2-lane roads, while only locations with sufficient congestion were added. Two main types of bottlenecks exist, namely on-ramps and off-ramps. As most of these are located close together, it was inevitable to pick a few bottlenecks which consist of both on-ramps and off-ramps. Finally, the grade of the highway will be taken into consideration as heavy vehicles do lose operational capabilities as the gradient becomes higher.

The traffic conditions are highly important as they contain the speed limit, the average percentage of heavy vehicles, and the total number of traffic breakdowns. The higher the speed limit is, the larger the speed differences between cars and heavy vehicles become. Different speed limits between 100 – 130 km/h are selected. Next, the average number of heavy vehicles on the road is one of the main factors on which bottlenecks are selected. As for the capacity drop computation, the main factor is the share of heavy vehicles during peak hours, which is on average way less compared to off-peak hours, as illustrated in Figure 12. Here, the average percentage of heavy vehicles during the morning peak was about 10%, while this value decreases to about 7.5% for the evening peak. Therefore, the exact percentage of heavy vehicle share during breakdowns could only be estimated after classifying the different breakdowns. Also, the total number of traffic breakdowns indicates the severity of the bottleneck and shows the number of breakdowns as well as recoveries that were classified. As mentioned, it is important that systematic breakdowns occur at these bottlenecks such that sufficient data can be collected.

Lastly, the control conditions are separated into static and dynamic conditions. The static conditions indicate the driving rules, such as an overtaking ban for freight traffic while the dynamic conditions inform about the control schemas that are present, e.g. ramp-metering or variable speed limits. Specific truck driving bans are inventoried, e.g. overtaking bans for heavy vehicles (Rijkswaterstaat, 2020b).

Besides the conditions of the roads which are used to characterize each road, it is important to also minimize the effects of external factors. The effort taken differs marginally in the existing study as highlighted in 2.3.3.2. For this study, the following conditions are taken into consideration for evaluation of the data:

- Accidents
- Driver population

To remove accidents, a minimum flow to consider a breakdown a breakdown will be assumed. Only breakdowns above 1200 veh/h/lane will be considered as breakdowns. They are considered the result of oversaturation of the flow. Although accidents might just occur when the flow is high, it is impossible within this study to filter those accidents. However, this could be done by considering the type of congestion that follows a breakdown. Furthermore, only recoveries above 900 veh/h/lane are considered, as lanes might be closed due to an accident for example, which would lead to an undersaturated recovery.

Regarding the driver population, separation will be made for the morning and evening peak, respectively from 6:00-11:00 and 14:00-19:00. In the analysis of Geistefeldt (2009), it was found that different capacity distributions were found based on assumingly different driver populations, at either the morning or evening peak. In that study, also a distinction between weekdays and weekend days was made, which eventually lead to a shortage of traffic breakdowns during the weekend for accurate analysis. The breakdown flow is lower for weekdays compared to weekend days (Calvert et al., 2016) and should preferably be removed. However, in this study, it is decided to not remove weekend days from the analysis, because of the lack of return-effort yield.

The weather conditions and working zones were not considered for this analysis. The removal of raining days would remove the low-flow observed breakdowns, assumingly decreasing the standard deviation of the breakdown intensity. However, as the study period is such extensive, and the sites are located relatively close to each other, the effect will be minimal. Here again, there was a lack of effort-yield. As for the working zones, the study periods were simply too large as was the number of selected sites to manually remove working zone days.

3.2.2 Description of Sites

In this section, different locations have been selected based on the requirements set in 3.2.1. As shown in 2.3.3.1, both a bottleneck detector (3) and a downstream detector (4) are at least required for capacity estimation. In this analysis, only loop-detectors that measure all vehicle categories are used.

It was observed that although these loop detectors are placed on tactical locations, the downstream detector sometimes is not close enough to rule out spillback from another bottleneck. Often, regular loop detectors that do not measure different vehicle classes were located at those locations, but due to difficulty in combining the datasets, these were not considered. Moreover, the geometric, traffic, and control conditions are important in selecting locations, illustrated in 3.2.1. Sites have been checked manually for locations that have systematic breakdowns. Bottlenecks that showed systematic breakdowns were selected. After analysing the data, it was observed that geometric conditions changed over time for some locations. Therefore, the study period is in these cases reduced to ensure the geometric properties are equal for the entire study period. The static control properties regarding overtake bans have been added according to the information provided by Rijkswaterstaat (2020b) but sometimes also changed within the study period. During visualization of the road sections with Google maps, no modifications to the truck driving bans were noticed within the respective study periods.

The selected locations have been summarized in Table 9 with information about the geometric, traffic, and control conditions. Although the data preparation and utilization will be discussed in the next chapters, 3.3 and 3.4 respectively, info regarding conditions is already shown. The observed number of heavy vehicles mentioned in Table 9, is the percentage of heavy vehicles during all classified breakdowns, and not the average percentage of heavy vehicles for an entire day. Several dynamic control systems were observed during the

selection process and are abbreviated as Variable Speed Limits (VSL), Peak Hour Lanes (PHL), and Lane Changing Bans (LCB).

A visualisation of the investigated road sections can be found in Figure 27. The bottleneck detector and the downstream detector, including their location, have been added to show that the breakdowns find their origin in the selected bottleneck. An extended overview, including a description per road with pictures of the real situation and other information regarding the analysis, has been added in Appendix B.

A total of 11 sites have been selected, located at 9 different main roads, namely the A1, A2, A4, A6, A12, A15, A27, A50, and A58. All the sites are located below Amsterdam, ranging from the A15 near the port of Rotterdam to the border between the Netherlands and Germany on the west side.

The main characteristics of the roads will be discussed. Mainly 2-lane roads have been selected. The A2 is the only site to have 5 lanes, while the A6 consists of two separated 2-lane roads in one direction. Almost all measurements collected were at the evening peak, except for the A27 and A58, where the morning peak data was used. The A1 and A12 solely consist of an off-ramp, while the A2, A6, A50, and A58 solely consist of an on-ramp. At the A4 and A15 sites, the on- and off-ramp detectors were located before the ramps, while the downstream detector was after both ramps. The A27 does have not have a bottleneck in the form of a ramp, but with the narrowing of a bridge and a barrier. All roads had a relatively flat surface, while the A15b had a gradient of more than 5%. The number of classified breakdowns ranges from 381 at the A15c to 4556 at the A12. Only the A2 and A15a did not have an overtaking ban for heavy vehicles during the measurement period. Finally, most roads had Variable Speed Limits (VSL) except for A15b, A15c, and A50. The A1 has a Peak Hour Lane (PHL), while the A50 and A58 have Lane Changing Bans (LCB) at the ramps, such that the left traffic was not able to change lanes.

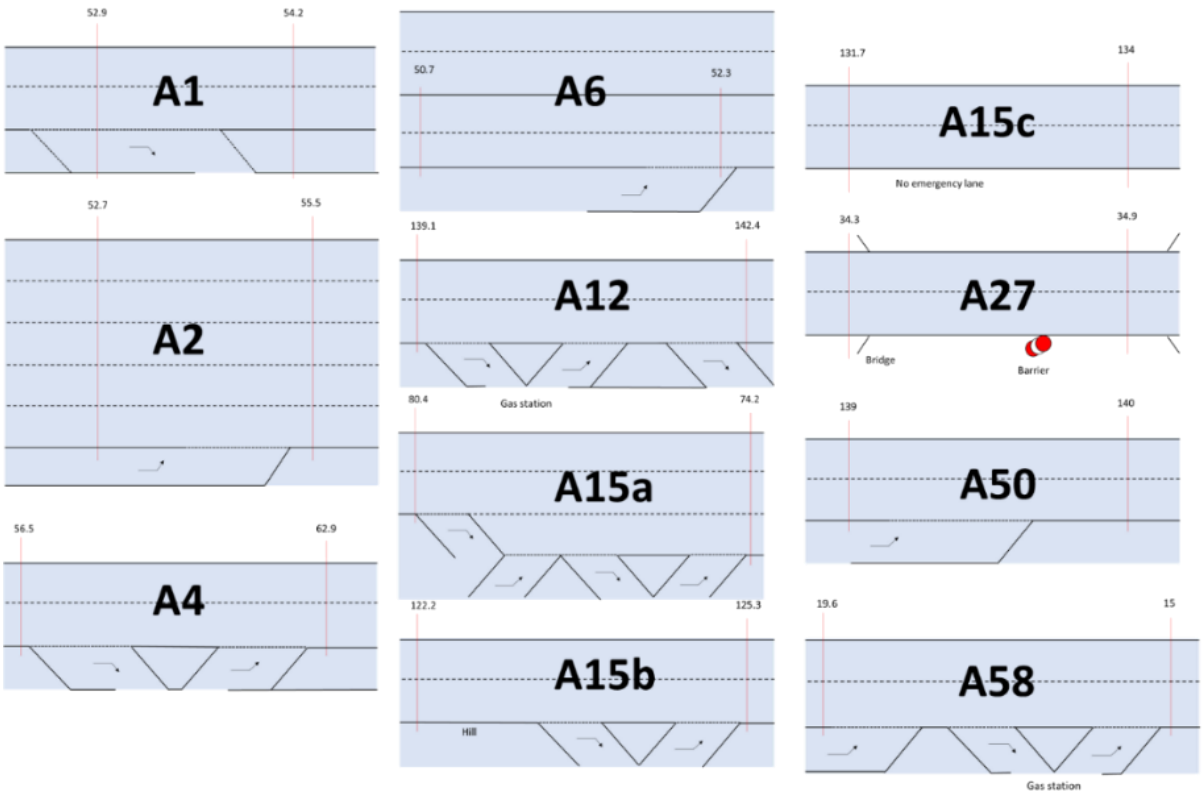


Figure 27 - Visualisation of analysed freeway sections

Table 9 - Description of analysed freeway sections

Free-way	Location	Study period	Geometric			Traffic			Control	
			No. of lanes	Type of bottleneck	Slope (%)	Speed limit	Share of heavy vehicles	No. of traffic breakdowns	Static	Dynamic
A1 R 52.9	Barneveld	2013 - 2020 (E)*	3	Off-ramp	>5	100	2-3%	864	Overtake ban 6:00-19:00	VSL + PHL
A2 R 52.7	Amsterdam – Utrecht	2014 - 2019 (E)*	5	On-ramp	>5	100	3-4%	2307	-	VSL
A4 R 56.6	Delft-Schiedam	2016 - 2019 (E)*	2	Off- and On-ramp	>5	100	4-6%	1490	Overtake ban 6:00-19:00	VSL
A6 R 50.7	Amsterdam – Almere	2013 - 2016 (E)*	2	On-ramp	>5	100	0-3%	432	Overtake ban 6:00-19:00	VSL
A12 R 139.1	Arnhem – Germany	2015 - 2020 (E)*	2	Off-ramp	>5	130	5-8%	4556	Overtake ban 6:00-19:00	VSL
A15a L 80.4	Sliedrecht	2013 - 2018 (E)*	2	Off- and On-ramp	>5	120	6-10%	1225	-	VSL
A15b R 122.2	Gelder-malsen	2014 – 2018 (E)*	2	Off- and On-ramp	<5	130	8-11%	1486	Overtake ban 15:00-19:00	-
A15c R 131.7	Tiel	2014 – 2018 (E)*	2	No emer-gency lane	>5	130	7-11%	381	Overtake ban 15:00 – 19:00	-
A27 R 34.3	Utrecht – Breda	2014 – 2017 (M)*	2	Bridge	>5	100	4-8%	563	Overtake ban 6:00-19:00	VSL
A50 R 139	Ravenstein	2013 – 2017 (E)*	2	On-ramp	>5	130	5-8%	1560	Overtake ban 6:00-19:00	LCB
A58 L 19.6	Moeigestel – Oirschot	2013 – 2018 (M)*	2	On-ramp	>5	120	4-7%	1611	Overtake ban 6:00-19:00	VSL + LCB

*(E) = Evening (14:00-19:00), (M) = Morning (6:00-11:00)

3.3 Data Preparation

In this section, the preparation of the data is elaborated. The Dutch loop-detector data from the Netherlands collected by NDW will be the primary source of data. Data analysts at TNO have restructured the raw data to improve machinability, which is used and further build upon in this study.

3.3.1 Description of the Data

The datasets used contain measurements of the speeds (km/h), intensities (veh/h), and the number of available lanes aggregated per minute per detector. The detectors are most often spaced out every few hundred meters, depending on the placement of the detectors by road authorities. In the dataset, information about the loop detectors such as location, road number, hectometre, and driving direction is also stored. In total, 1925 road sections are distinguished in the data each road containing at least a single loop detector. Main highways have more loop detectors compared to other roads, e.g. the A1 has a total of 443 loop detectors in one direction over a length of about 140 kilometres, divided over multiple lanes. However, not all detectors can determine vehicle classes in terms of flow rate based on the length of the vehicles passing by, as explained in 2.2.1. Therefore, only detectors that can distinguish vehicle classes will be used in this analysis. Detector data has been collected since 2008 and is available up to the current data. However, TNO only possesses data from 2013-2020, which is available for this analysis. However, some years do miss files, e.g. 2018 misses about 150 days wherefrom most in the second half of the year, while 2017 and 2019 miss approximately 5% of the data.

All measurements collected on a single day from a single-vehicle class are stored in one .mat file. Due to the indexing of the roads in the file, it is possible to combine the flow rates of the different vehicle classes of a single day of a single loop detector. The process of combining the separate data files is performed in the Flow Type Classification Model (3.4). For example, the classification of the data of the A1 required scanning through approximately 8000 files for possible breakdowns and recoveries. The .mat data files are processed into Python with the help of h5py and SciPy.io depending on the binary format of the data. The processing had to be optimized in such a way that looking for the desired loop detector was not a step-wise process of looking for the proper index due to the scale of the data. Therefore, an interpolation search algorithm was used.

3.3.2 Data Filtering & Smoothening

Raw loop detector data is far from flawless due to irregularities in the measurement accuracy, measurement errors, and spatial location of the detectors. Therefore, based on the scientific research executed, raw data is often filtered to reconstruct incomplete information. TNO's internal data expert, Taoufik Bakri, has filtered the raw data according to the extended and generalized Treiber-Helbing filter to fuse multiple sources based on the recursive Kalman filter-based approach (M. Treiber & Helbing, 2002). The so-called "adaptive smoothing method" reconstructs spatiotemporal traffic data from complex traffic dynamics. The spatiotemporal pattern depends on individual freeway flow and bottleneck configuration and therefore smoothens fluctuations out, such as the highest flows measured. These extreme flows are what is needed to find the capacity according to the Product Limit Method as the maximum flow measured before breakdown indicates the temporal capacity. However, filtering out peak flows will lead to underestimated capacities. Therefore, the filtered data is only used during the development of the algorithm as a comparison.

Another possibility is to smoothen the data, such that the measured data at a single point in time is averaged with the measurements just before and after. Since the speed and intensity measurements from the raw data are very volatile, it is necessary to somehow smoothen the data. Both Brilon et al. (2005) and Geistefeldt (2009) used an aggregated temporal resolution of 5-minutes for the speed and flow values. After comparing several resolutions, it was decided to use a 3-minute moving average. This is done so that the used values are as close to the real measurements as possible, while the smoothing disregards the high fluctuation. Moreover, the moving-average technique does keep the number of measurements almost identical to the unsmoothed data, while the aggregation technique used in Geistefeldt (2009) reduces the number of measurements multiplied by the chosen resolution. A detailed explanation can be found in Appendix A including visualisation and results of the different temporal resolutions of both the raw and filtered data.

3.4 Calculation Framework

The Calculation Framework has as goal to process the raw data measured by the detector data, consisting of flow, speed, and the number of available lanes, and identify the relationship between the capacity drop and the share of heavy vehicles. The framework consists of 4 different stages and is designed in such a way that it can be used for the analysis of the simulation data as well. To establish this, the input data should be generic without modification.

An overview of the calculation, with visually made in- and output data, can be observed in Figure 28. The Calculation Framework is developed based on the methodologies discussed in 2.3.2. The main methodologies used are the Kaplan Meier Product Limit Method to estimate breakdown probability distributions (Regler, 2004), extended with recovery probability distributions (Brilon et al., 2005) which are at the basis of this framework. The following computational steps are executed:

- Flow Type Classification Model (3.4.1)
- Product Limit Method (3.4.2)
- Capacity Drop Computation (3.4.3)
- Statistical Tests (3.4.5)

The data at the bottleneck detector is aggregated from the vehicle category files, including the actual number of lanes to exclude lane closures. Furthermore, the downstream detector speed measurements are used collected to identify spillback from other bottlenecks. The Flow Type Classification Model classifies the measured flows, as well as relevant road circumstances, e.g. the share of heavy vehicles. Second, non-parametric Kaplan Meier curves are drawn from these data and subsequently, Weibull distributions are fitted. Besides, the Passenger Car Equivalency is computed to identify the spatial effect on traffic flow of heavy vehicles (Geistefeldt, 2009). Next, the capacity drop is computed based on the difference between the observed breakdown and recovery distribution functions. Finally, statistical tests are performed on the Kaplan Meier curves as well as on the individual time-dependent breakdown-recovery detected flows to test correlation with the share of heavy vehicles.

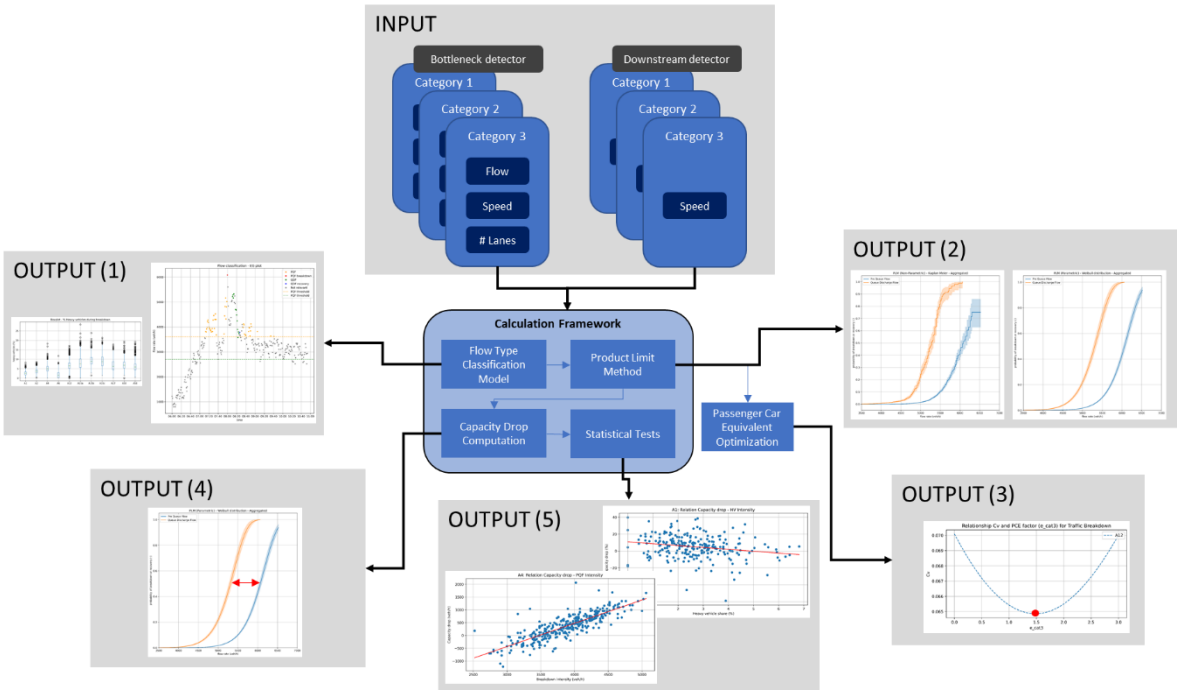


Figure 28 - Calculation framework

3.4.1 Flow Type Classification Model

After the data is ready for computation and the sites have been selected, the breakdown and recovery moments should be identified. The Flow Type Classification model aims to quantify, from speed and flow measurements, at what flow the breakdown of traffic occurs used to estimate the PQF capacity. The flow and speed measurements before breakdown are collected, as they indicate at which flows traffic does yet not break down and together are used to estimate the probability of a breakdown at a certain flow. If an event happens, in this case, a breakdown, the time-dependent event will be classified uncensored. Contrary, when the breakdown has not happened yet, the previous time-dependent measurements close to breakdown will be classified as censored. Simultaneously, the model classifies breakdown recoveries at the moment congestion resolves to estimate the QDF capacity. Measurements before recovery will be classified censored, while the breakdown recovery flow will be classified uncensored. Traffic breakdown and traffic breakdown recovery are both regarded as failure events so that lifetime data analysis methodology can be used for the estimation of respectively the breakdown capacity and the breakdown recovery capacity.

However, this theoretical approach needs clarification before applying it in practice as classifying breakdown and breakdown recovery is not as trivial as illustrated. The classification illustrated in 2.3.2, the Kaplan-Meier approach (parametric and non-parametric), shows how a breakdown is determined according to different traffic stages (Brilon et al., 2005). Four different stages are classified regarding the PQF: B, F, C1, and C2.

B: Traffic is fluent in time interval t, but the observed volumes cause a breakdown, i.e. the average speed drops below the threshold speed in the next time interval t + 1.

F: Traffic is fluent in interval i and the following interval t + 1. This interval i contains a censored value. Its information is that the actual capacity in interval t is greater than the observed volume $q_{i.}$

C1: Traffic is congested in interval t, i.e. the average speed is below the threshold value. As this interval t provides no information about the capacity, it is disregarded.

C2: Traffic is fluent in interval t, but the observed volume causes a breakdown. However, in contrast to classification B, traffic is congested at a downstream cross-section during interval t or t – 1. In this case, the breakdown at the observation point is supposed to be due to a tailback from downstream. As this interval t does not contain any information for the capacity assessment at the observation point, it is disregarded.

A practical application is demonstrated in Geistefeldt and Brilon (2009) including an elaborated set of criteria to detect a traffic breakdown at a single time interval. Their criteria have been adjusted such that the smoothening method of the input data has been taken into account. The simplest method to detect a breakdown is by choosing a speed threshold v_t, such that breakdown is recognised when speed v drops below v_t from time interval t into time interval t+1. However, multiple breakdowns, shockwaves, could be recognised when the speed fluctuates around the speed threshold. Therefore, the following set of sophisticated criteria were adjusted to classify breakdowns and minimize shockwave classification:

$$v(t-1) \geq v_t \qquad \textit{Eq. 10}$$

$$v(t) \geq v_t \qquad \textit{Eq. 11}$$

$$v(t+1) \leq v_t \qquad \textit{Eq. 12}$$

$$v(t+2) \leq v_t \qquad \textit{Eq. 13}$$

$$\sum_{i=-5}^{0} \frac{v(t)}{5} - \sum_{i=1}^{10} \frac{v(t)}{10} \geq 10\ km/h \qquad \textit{Eq. 14}$$

$$I(t) > 1200\ veh/h/lane \qquad \textit{Eq. 15}$$

The first four equations assure that the speed is above the threshold value in the two intervals before, and below the threshold speed below in the two intervals afterwards. Eq. 14 checks whether the speed five minutes before breakdown shows at least a drop of 10 km/h average compared to the ten minutes after breakdown to prevent relatively small shockwaves to be classified as breakdown. Different speed threshold values were tried but the same speed threshold value used in Geistefeldt and Brilon (2009), v_t = 70 km/h, appeared to show proper results. If all criteria are met, the interval is classified as stage B. However, if only

$$v(t) \geq v_t \quad \textit{Eq. 16}$$

$$v(t+1) \geq v_t \quad \textit{Eq. 17}$$

Eq. 16 and Eq. 17 are satisfied, the measurement is regarded as a censored value (stage F). This classification however does not yet consider the second loop detector. In case the congestion is caused by a spillback from a downstream bottleneck, the measurement should be disregarded (stage C2). Therefore, the second loop detector downstream is used to check Eq. 11 for the respective detector. If a speed below the threshold is measured, the stage is classified as C2 and disregarded from the data. Moreover, all future measurements that day will be disregarded as the congestion is not initialized by the analysed bottleneck.

Besides classifying breakdown, the model is also able to classify recovery from a breakdown. This method is based upon Brilon et al. (2005) for the definition of the traffic states. Unfortunately, their exact set of criteria were not presented within the paper. Therefore, the following set of traffic states were created:

B*: Traffic recovers from congestion to free flow, i.e. the average speed exceeds the threshold value from time interval t to interval t + 1

F*: Traffic is congested in intervals t and t + 1, i.e. the average speed is lower than the threshold value in both intervals. This interval t contains a censored value.

C*: Traffic is fluent in interval t, i.e. the average speed is above the threshold value. This interval is not relevant.

The following set of criteria have been created for the classification of a recovery:

$$v(t-1) \leq v_{t,2} \quad \textit{Eq. 18}$$

$$v(t) \leq v_{t,2} \quad \textit{Eq. 19}$$

$$v(t+1) \geq v_{t,2} \quad \textit{Eq. 20}$$

$$v(t+2) \geq v_{t,2} \quad \textit{Eq. 21}$$

$$\sum_{i=-5}^{0} \frac{v(t)}{5} - \sum_{i=1}^{10} \frac{v(t)}{10} \leq -5\ km/h \quad \textit{Eq. 22}$$

$$I(t) > 900\ veh/h/lane \quad \textit{Eq. 23}$$

It can be observed that the set of recovery criteria are very similar but opposite to the breakdown classification criteria. The only exceptions are the lower average speed gain of only 5 km/h and the lower required intensity value. The latter is done because of the decreased intensity during recovery. The speed threshold for recovery is set slightly below the speed threshold for breakdown, at $v_{t,2}$ = 60 km/h. This is done to exclude shockwaves as much as possible from the recovery identification. Concluding, the interval is regarded as state B*, including an uncensored value when Eq. 18 - Eq. 23 are satisfied. However, if only

$$v(t) < v_{t,2} \quad \textit{Eq. 24}$$

$$v(t+1) < v_{t,2} \quad \textit{Eq. 25}$$

Eq. 24 and Eq. 25 are satisfied, the measurement is regarded as F*. Moreover, the state is only recognized as F* if in the previous four time-intervals the stage is not regarded as state B. The reason is that a short period after the breakdown, traffic might become very unstable, and a short cool-down period needs to be realized. A similar cooldown approach was applied for the breakdown recognition (state F), which was also only classified t if the previous four time-intervals were not classified as a state B*.

To illustrate the effect of the breakdown and breakdown recovery detection, the algorithm will be visualized with an example of an analysed bottleneck. In Figure 29, the A1 near junction Barneveld (hectometre 52.9 at Wednesday 23-03-2016 during the morning peak) is visualized. The site consists of a three-lane highway with an off-ramp as the bottleneck. The black line shows the speed including the speed threshold (70 km/h), while the grey line shows the intensity and the intensity threshold (3600 veh/h) according to the number of lanes. The data used for this example was prepared with the 3-minute smoothening technique. A steep reduction in speed around 8:00 can be observed, followed by a decline in intensity. Around 8:25, the speed recovers back above the threshold speed and the intensity seems to decline onwards indicating that the morning peak is over.

Figure 30 shows the speed contour plot for the entire day for the A1. The y-axis shows the location while the x-axis shows the time. The colours of the heatmap represent the measured speed, the lighter the colour the higher the measured speed. The green and purple lines indicate the hectometre at which location the measurement loop detector and downstream detector are located. So, the intensities and speeds observed in Figure 29 are derived from the morning peak at the level of the green line (hectometre 52.9). The off-ramp ends at hectometre 53.8, therefore the detector at hectometre 54.2 is used to check for downstream spillback, which is not the case. One can observe a more severe traffic breakdown in the evening which is not originated at the off-ramp but a bottleneck downstream. As the spillback detector did not satisfy Eq. 11 the classification of flows stopped at the evening peak after the downstream congestion occurs.

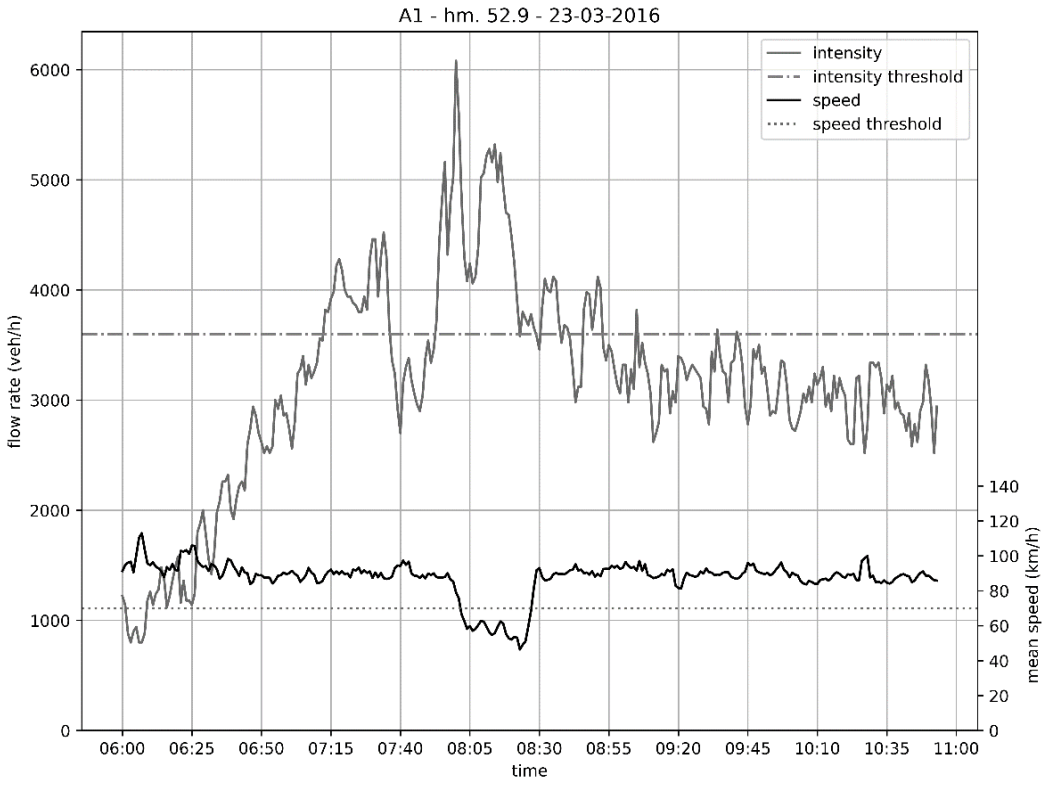


Figure 29 – A1: Flow and speed plot (Q(t) and U(t) plot)

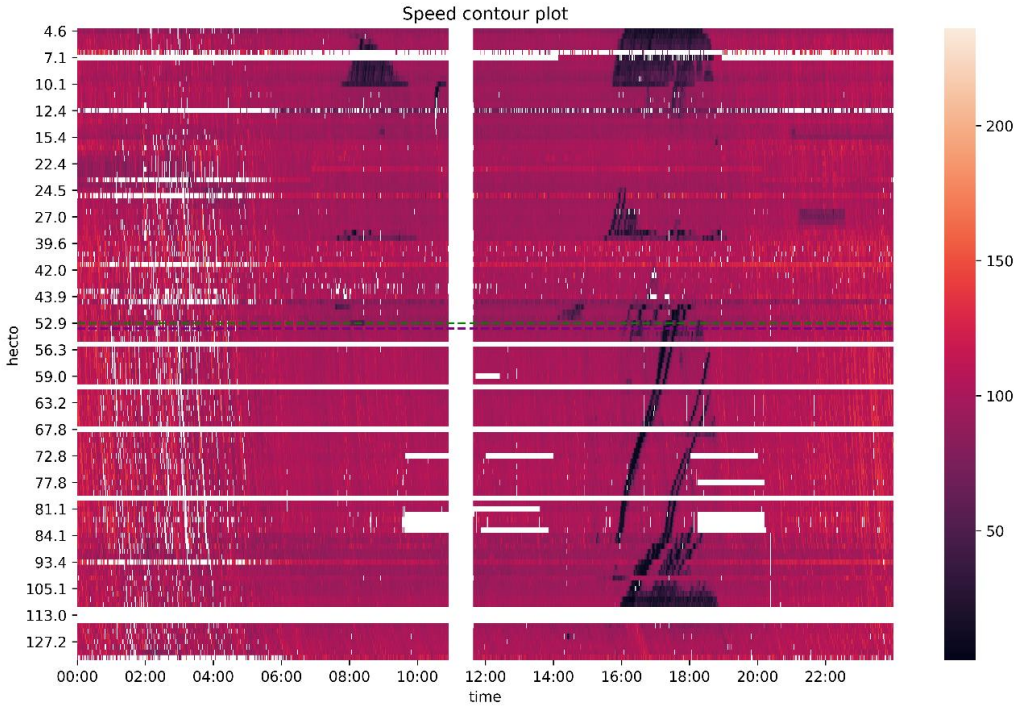


Figure 30 - A1: Speed contour plot

The classified flows of the Flow Type Classification Model for the morning peak can be observed in Figure 31 and Figure 32. First, the intensity-time plot including the classification is shown in Figure 31. The measurements before 7:10 are disregarded as Eq. 15 is not satisfied. Thereafter, the intensity increases up to the breakdown intensity at 08:00, as indicated in the overview of uncensored measurements in Table 10. However, the speed at t=08:00 is not below the threshold speed v_t of 70 km/h. This is because when Eq. 10 - Eq. 15 are satisfied at t, the model searches the highest intensity value in a short time window (t-4:t) and classifies that timestep as uncensored. The reason is that the raw speed and flow still show quite some volatility at each timestep as the shown data is only smoothened with a 3-minute MA. The maximum intensity reached in the four minutes before the breakdown likewise caused an overflow of vehicles and subsequently the breakdown. This is different in Regler (2004), where the data was aggregated per five minutes. One of the problems with their methodology was the lack of uncensored measurements, which could likely be caused by the fact that the 5-minute aggregation did not capture breakdowns properly within a timeframe. The applied technique of picking the highest intensity with a non-aggregated 3-min moving averaging results in less flattening of the highs while being able to precisely capture the flow causing the breakdown. However, the volatility within the data remains high as illustrated in Figure 31.

The next ten timesteps are disregarded from the data due to the instability of traffic. Some of the following measurements satisfy Eq. 24 - Eq. 25 and are classified as censored QDF points. Finally, the breakdown recovery criteria (Eq. 18 - Eq. 23) are met and in the time window (t-3:t+1), the timestep with the highest intensity is classified as uncensored. This appears to be at 08:27, see Table 10.

Also, the speed of the upstream detector is checked and appeared to be above v_t at any time. If the upstream detector would measure a speed below v_t, no further measurements will be classified and the previous classified measurements will be disregarded.

Table 10 - Uncensored classified measurements

	SPEED	SPEED_UPSTREAM	INTENSITY	UNCEN-SORED	CONGESTION TYPE	CAT1	CAT2	CAT3
08:00	79.17	87.03	6080.0	1	PQF	5800.0	220.0	60.0
08:27	69.56	98.81	3780.0	1	QDF	3680.0	20.0	80.0

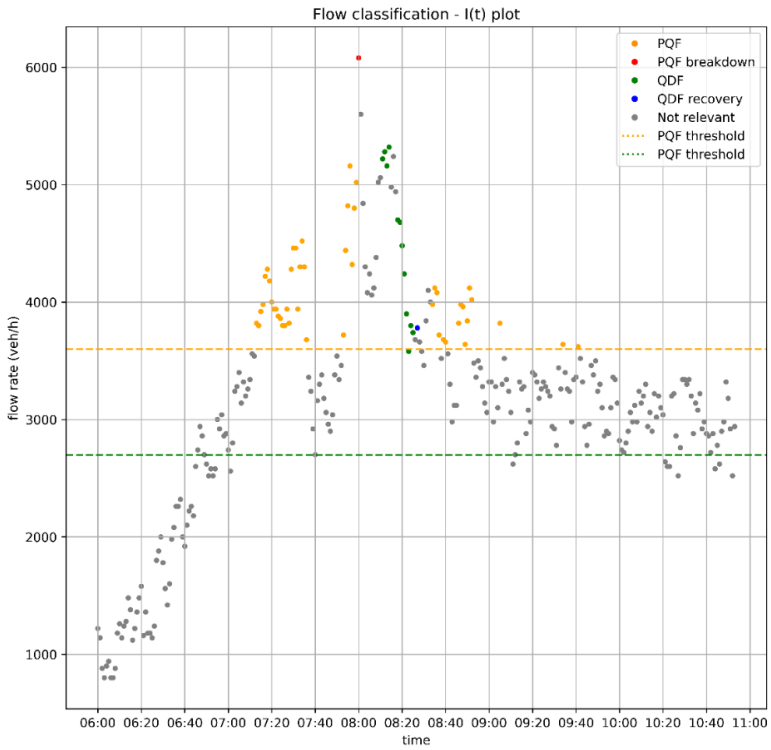


Figure 31 – A1: Flow Classification (I(t) plot)

The Flow classification visualized in the speed-flow plot can be observed in Figure 32. To improve visualization, the data points are connected with lines to see the dependency between measurements. The PQF breakdown and QDF breakdown recovery breakdown classified in Figure 31 are again marked with red and blue dots in Figure 32. The methodology to select the highest intensity in a 4-minute window before breakdown becomes clear as a clear drop in both intensity and speed can be observed after Eq. 10 - Eq. 15 are satisfied. However, the measurement at exactly t would not show the intensity that caused the breakdown, but a point slightly below the red dot. The same holds for the breakdown recovery, where the highest intensity 5 minutes before the measurement that satisfies Eq. 18 - Eq. 23 is selected.

All classified measurements, either censored or uncensored, are combined in a data frame including the flow per vehicle class also with the 3-minute MA value (see Table 10). The process of classification is subsequently scaled for the entire study period. This data frame with flow classifications will be the input for the Product Limit Method, which will be elaborated in 3.4.2.

The main problem observed during scaling the flow type prediction model is the detection of stop-and-go-waves. One of the benefits of smoothing data is that these abrupt differences are smoothened, and traffic states could be better distinguished. However, the model still sometimes detects stop-and-go waves as either a breakdown or recovery value, although Eq. 14 and Eq. 22 are there to prevent stop-and-go-waves from being classified as multiple breakdowns. The weakness of the model lies within properly detecting stop-and-go-waves, which could be diminished by enlarging the resolution. However, this goes hand in hand with flatting the tops and subsequently leads to an underestimation of the capacity. Therefore, it has been decided to keep this method to approximate the actual breakdown and recovery flow values in the best possible way, while collecting more observations due to falsely detected stop-and-go waves as breakdowns or recoveries.

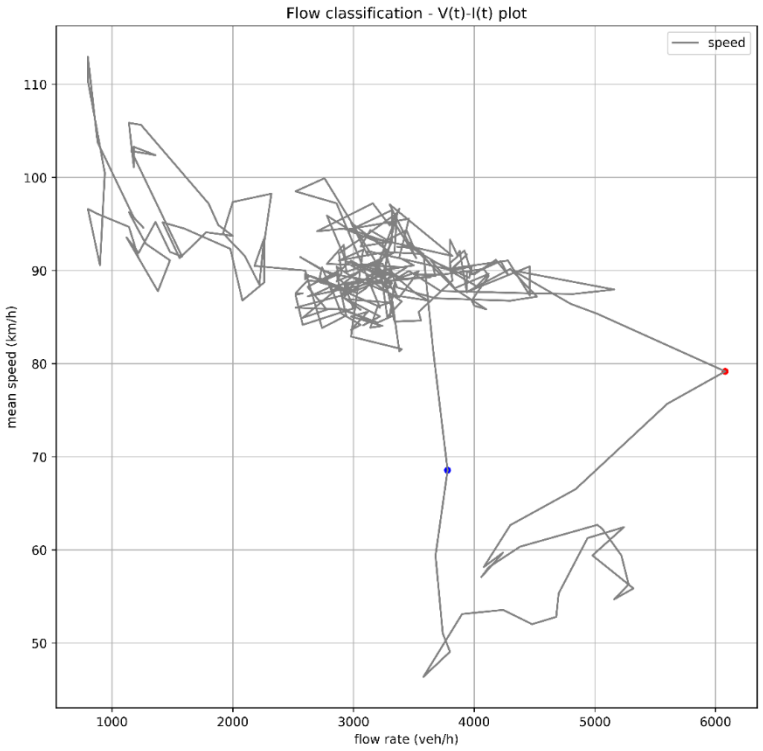


Figure 32 – A1: Flow Classification (V-I plot)

3.4.2 Product Limit Method

After the traffic flow classification at the bottleneck for the chosen study period, the collected data points are used as input to compute the survival function. To explain the methodology, the site at the A1 will be used as an example. An overview of the classified data points distinguished as censored or uncensored measurements can be found in Figure 33 and Figure 34 from 2013-2017 for respectively the PQF and QDF. It appeared that the road has both a configuration of 2 lanes and 3 lanes in the data set. Therefore, only 2013-2017, the situation where the road has a 3-lane configuration, with the third lane being the peak-hour lane, will be used for the analysis of this road section. The difference in capacity between those years due to the difference in the number of lanes measured will be elaborated at the end of this chapter, but can already be observed in Figure 39.

The censored values are coloured in light blue, while the uncensored values are coloured on a scale ranging from yellow to dark blue based on gaussian kernel density estimation (Scott's Rule) to visualize the density of the measurements. As can be observed in Figure 33, the maximum measured flows were not classified as uncensored. At this site, 262630 measurements were classified as censored, while only 864 breakdowns were detected. As can be observed, most detected breakdowns were between 5000 and 5500 veh/h with a mean speed between 75 and 85 km/h.

Regarding the QDF classification in Figure 34, 11907 measurements were classified while only 355 breakdown recoveries were observed. The breakdown recoveries are more scattered with intensities between 4800 and 5200 veh/h while the mean speed varied between 65 and 72 km/h. The main difference is that the maximum observed flow values were uncensored in this case, which is significantly different in establishing a survival function. The reason that many observed uncensored flows are above 60 km/h, the recovery speed threshold ($v_{t,2}$), Besides, many measurements were classified uncensored above the speed threshold. This is due to the short time window, 4 minutes, to compensate for the one-size-fits-all speed threshold approach. In this case, the mean recovery speed value is around 70 km/h, substantially higher than the 60 km/h threshold.

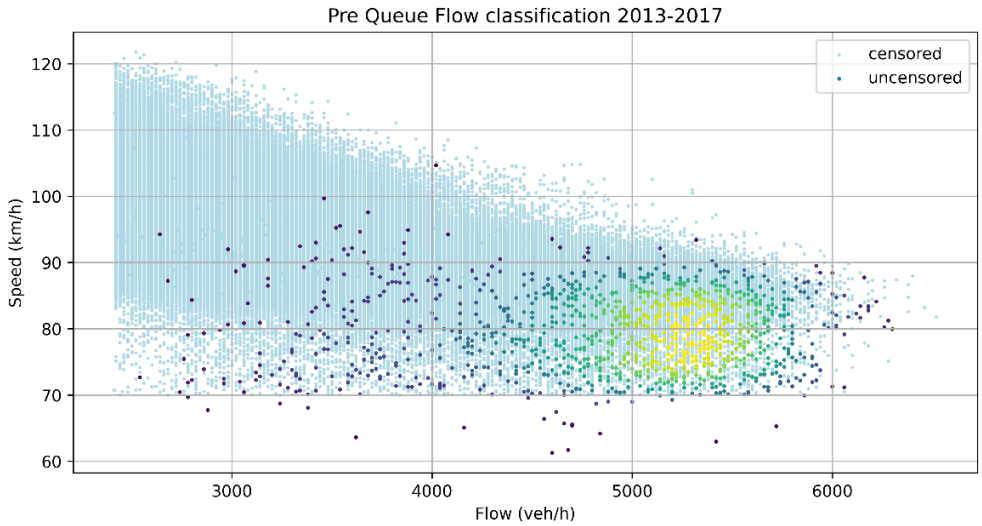


Figure 33 – A1: Flow Classification PQF (2013-2017)

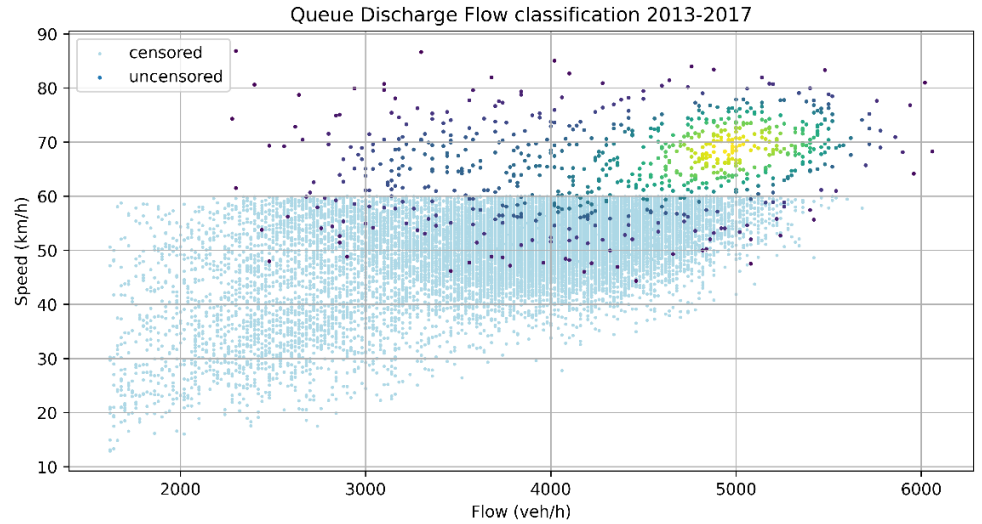


Figure 34 - A1: Flow Classification QDF (2013-2017)

Subsequently, the capacity distribution can be estimated according to the Kaplan-Meier approach (Eq. 5). Both the PQF and QDF classified measurements will be separately inserted into [B]. Next, the non-parametric Kaplan-Meier distribution is drawn with the help of the Lifelines package in Python. However, the distribution could not be entirely estimated as the maximum observed intensity values were not classified as breakdown nor even lead to a breakdown. Therefore, Brilon and Zurlinden (2003) tested different function types such as Weibull, Normal, and Gamma distribution to fit the capacity distribution, and subsequently estimate the entire distribution function. The Weibull function (Eq. 7) turned out to be the best fit and is called the parametric Kaplan-Meier distribution. Both the parametric and non-parametric distribution of the A1 highway can be observed in respectively Figure 35 and Figure 36.

$$F_c(q) = 1 - \prod_{i:q_i \leq q} \frac{k_i - d_i}{k_i}; i \in \{B\} \qquad \text{Eq. 5}$$

$$F(x) = 1 - e^{-\left(\frac{x}{\beta}\right)^{\alpha}} for\ x \geq 0 \qquad \text{Eq. 7}$$

In Figure 35, the orange and blue lines represent the breakdown and recovery distribution respective of Eq. 5. The shade behind the lines represents the confidence intervals based upon the exponential Greenwood confidence interval of 0.05, which is the standard in the Lifelines package. The QDF distribution displays a full distribution, while the PQF line only reaches about the 75th percentile. The maximum capacity found for this three-lane road corresponds to the rule of thumb capacity of 2100 veh/h/lane.

The fitted Weibull distribution of the breakdown and recovery data are displayed in Figure 36, while the corresponding estimated parameters to fit the Weibull distribution are displayed in Table 11. The values for the estimated parameters were in the same range in comparable studies (Brilon et al., 2005; Geistefeldt, 2009). However, the number of breakdowns (1155) appears to be considerably larger compared to the road with the most breakdowns (464) in Geistefeldt (2009). Although their exact study period is not known – between 5 months and 6 years – the results suggest that the removal of aggregation and lower moving average detects breakdowns more often, including the expected overestimation due to stop-and-go-waves.

Table 11 - Example: A1 R hectometre 53.0 - Estimated parameters

CLASSIFICIATION	CENSORED VALUES	UNCENSORED VALUES	A	B (PCU/H)	Σ (PCU/H)	μ (PCU/H)
BREAKDOWN (PQF)	289298	1155	14.51	6407	503	6180
RECOVERY (QDF)	18321	751	12.73	5531	488	5312

Comparing the results of the Kapan-Meier approach for estimating a capacity distribution (Figure 35) to the Gaussian density approach to highlight the average breakdown and recovery intensity (Figure 33 and Figure 34) shows some interesting results.

First, as expected, the PQF distribution does not reach a probability of 1 due to some censored measurements with higher intensities compared to the uncensored measurements. This is not the case for the QDF distribution, which likewise reaches a probability of 1.

Second, the most common breakdown and recovery intensity are respectively between 5000-5500 veh/h and 4800-5200 veh/h. The breakdown probability for those intensities is estimated between 0.02 – 0.1, while a recovery probability between 0.1 – 0.4 is estimated. This shows that this method founds its estimated distribution mainly on the maximum observed measurements which appear to be higher than the more common breakdown intensity.

Furthermore, does the fitted Weibull distribution very well fit the non-parametric Kaplan Meier distribution. The shape, even more for the QDF, does capture the distribution seamlessly. Therefore, it is decided to not test different distributions but use the methodology as Brilon and Zurlinden (2003). The Weibull fitted distributions in this figure are plotted until the Greenwood confidence interval of 0.05 is reached. However, it is possible to estimate the complete distribution function as the Weibull parameters are computed.

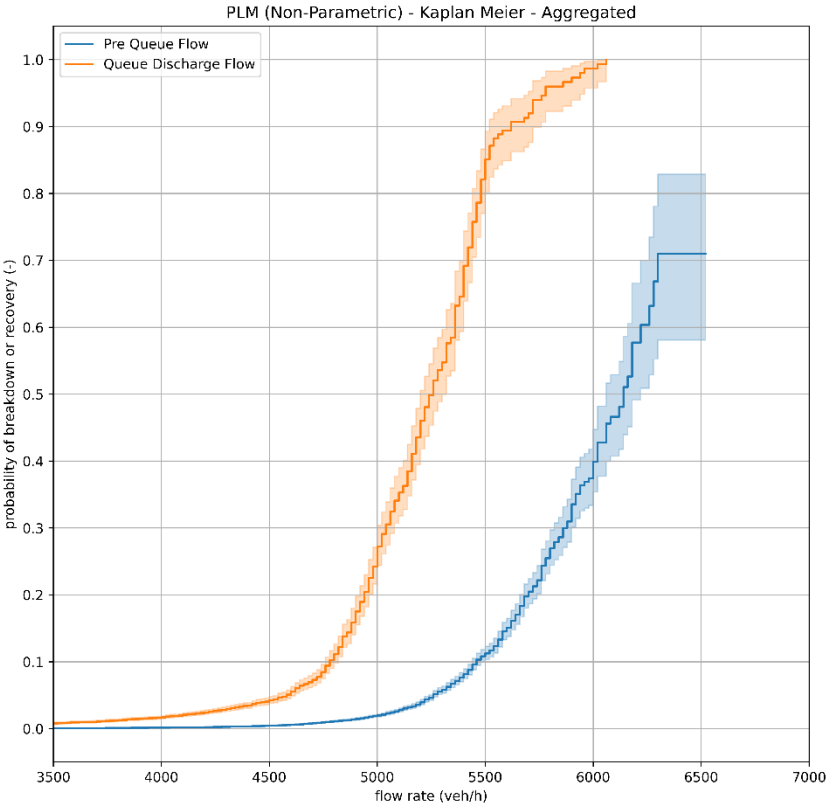


Figure 35 – A1: PLM (Non-Parametric)

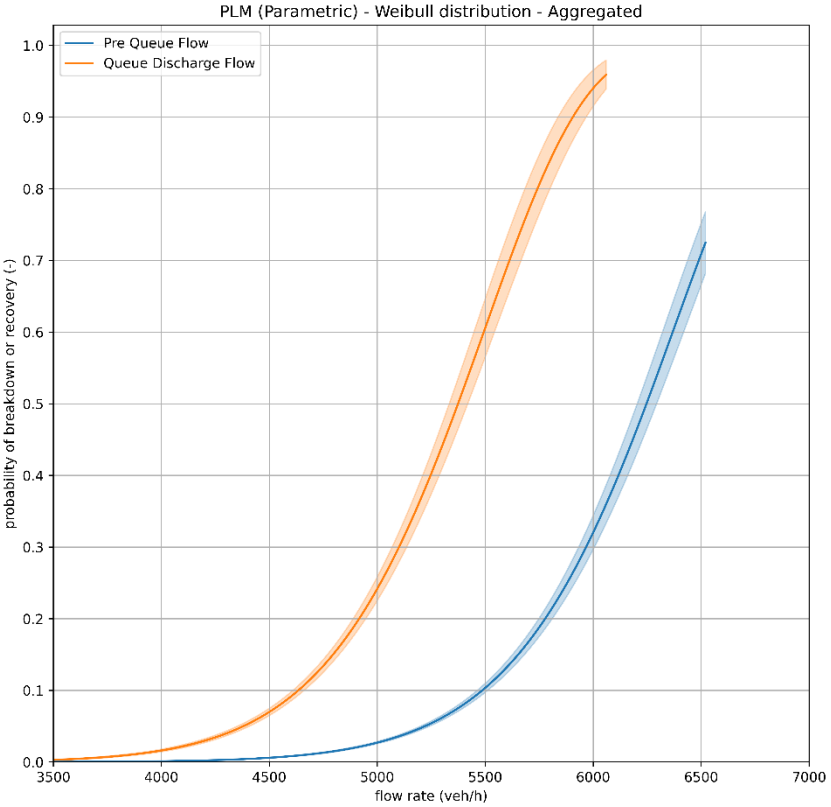


Figure 36 – A1: PLM (Parametric)

As mentioned, the percentage of heavy vehicles is captured during the Flow Type Classification Model, allowing to measure the percentage of heavy vehicles during particularly breakdowns or recoveries. In Figure 37 and Figure 38, the uncensored intervals show the heavy vehicles share when a breakdown/recovery was detected while the censored intervals show the intervals of all measurements classified within the Flow Type Classification Model. The share of heavy vehicles consists of measurements from only vehicle category 3 (Table 4). In this example, it can be observed that the share of heavy vehicles depletes as the intensity increases, which is in line with the results found in Figure 12. An overview of the heavy vehicle shares at each site can be found in 3.5.1.

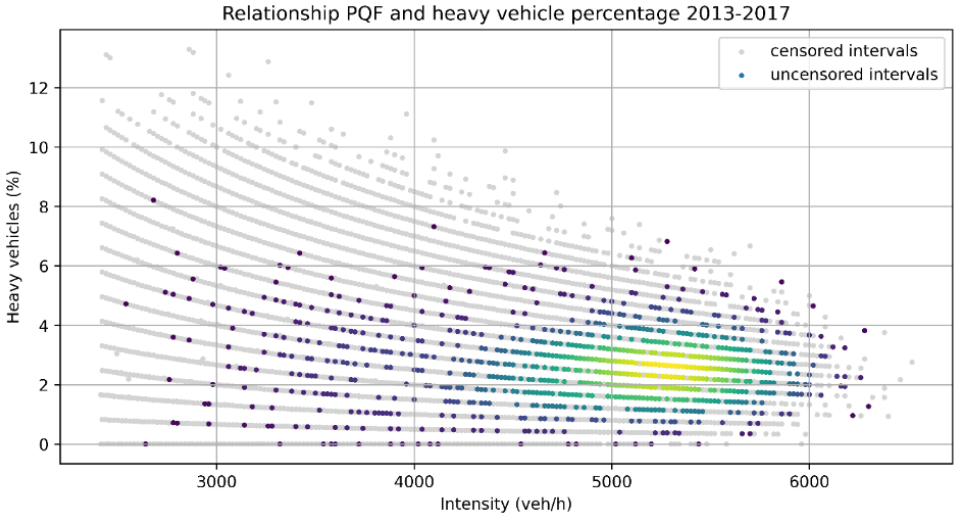


Figure 37 – A1: Relationship PQF and heavy vehicle percentage

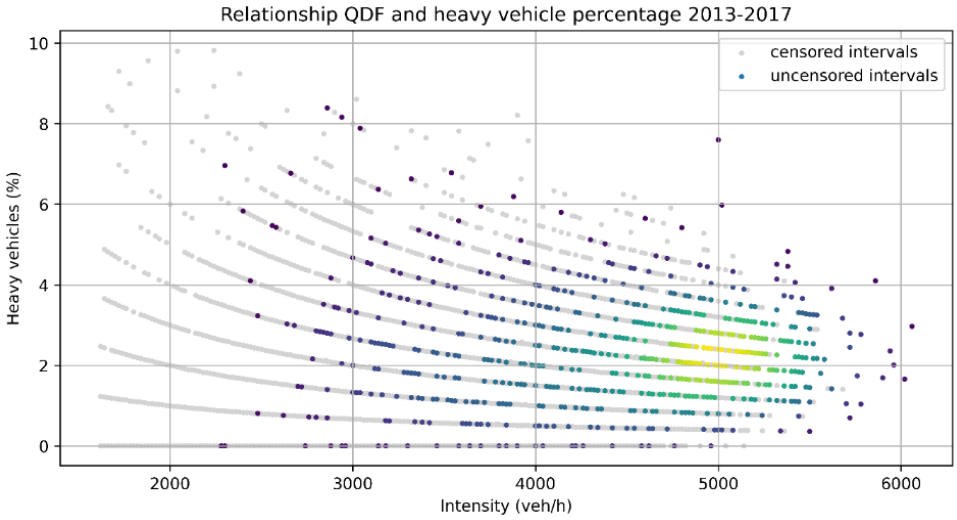


Figure 38 - A1: Relationship QDF and heavy vehicle percentage

3.4.2.1 Between-year results

Besides the aggregated parametric Product Limit Method computation for the entire study period, a yearly Product Limit Method has also been computed to see how the distribution changes between years and check for changes in the geometric road properties, which were not identified manually due to the scale of data. Although these properties have been checked visually within Google Maps, loop detectors measurements might change over time due to a changing location or a change in geometric properties of a road within the data. When observing Figure 39, the capacity observations do lower significantly for the years 2018-2020

compared to the 2013-2017 period. As mentioned, the number of measured lanes between 2013-2017 changed from three lanes to two lanes in the dataset. Therefore, the analysis period in the aggregated Product Limit Method has been changed to only 2013-2017 to keep constant road circumstances. The distributions have been fitted along the probability axis up until a Greenwood confidence interval of 0.05, but without the shade as in Figure 36 to maintain the readability of the graph. An overview of the parameters of all estimated distribution functions can be found in Appendix E.

Besides the goal of finding unobserved irregularities within the data, the PLM per year also allows for comparing yearly distributions of a single analysed freeway section. These observations allow for a comparison of the capacity drop each year for a specific site. Also, these yearly distributions could be combined with the average truck percentage per year to investigate how the in general the increasing number of heavy vehicles impact the capacity or recovery flow, which can be found in the result section (3.5).

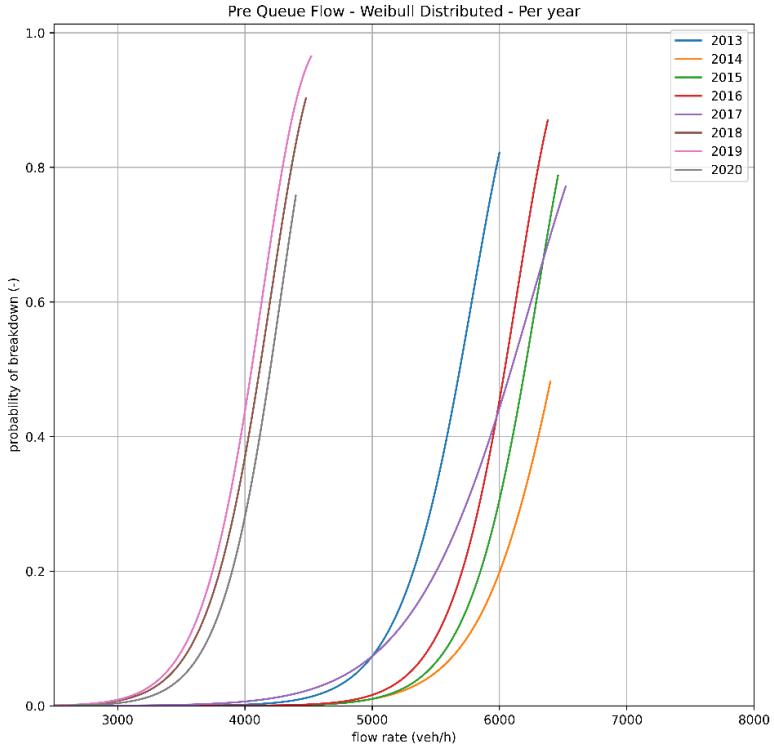


Figure 39 - A1: PLM (Parametric) per year

3.4.3 Capacity drop computation

After the PQF and QDF Weibull distributions have been estimated, the capacity drop is computed. The capacity drop is considered the difference between the breakdown flow rate and recovery flow rate. Hence, the difference between these two distributions results in a capacity drop. The median of both distributions will be used to compute the difference in flow rate and subsequently the percental difference in flow rate (Eq. 26).

$$Capacity\ Drop = \frac{PQF_{median} - QDF_{median}}{PQF_{median}} * 100\% \qquad \textit{Eq. 26}$$

This results in a capacity drop of 14% for the A1. Results of the other sites are displayed in the Results section (3.5). This method was also proposed in Brilon et al. (2005) as capacity drop estimation.

3.4.4 PCE Optimization

In an attempt to investigate the chaotic properties of the capacity breakdown rate, Geistefeldt (2009) estimated the passenger car equivalent based on capacity variability for different cross-sections. A similar method will be used to investigate the spatial impact of heavy vehicles on traffic flow in terms of passenger cars. Extending on that study, the heavy vehicle equivalency factor for both the breakdown capacity and the recovery capacity will be estimated. The data is different as three vehicle classes are distinguished, while only two vehicle classes are distinguished in Geistefeldt (2009): passenger cars and commercial traffic.

The Product Limit Method aggregates different flow rates of the different vehicle classes into a single flow rate value within the chosen temporal resolution. Eq. 27 shows how this is done with the equivalency factors kept 1, such that every vehicle contributes equally to the effective flow rate q.

$$q = q_{cat1} + e_{cat2} * q_{cat2} + q_{cat3} * e_{cat3} \qquad \textit{Eq. 27}$$

In which

$$\begin{aligned} q &= \ effective\ flow\ rate\ (PCU/h) \\ q_{cat1,\ cat2,\ cat3} &= \ flow\ rates\ for\ the\ different\ vehicle\ categories\ (veh/h) \\ e_{cat2,\ cat3} &= \ PCE\ for\ the\ different\ vehicle\ categories\ (PCU/veh) \end{aligned}$$

Generally, heavy vehicles are accounted for different compared to passenger cars in capacity estimations as both the physical and operational characteristics are different for these vehicle classes. Instead of assuming an equivalency factor, in this case, e_{cat3}, this method optimizes the Passenger Car Equivalency factor by minimizing the coefficient of variation of the capacity distributions (C_v), see Eq. 28, for both the breakdown and recovery distributions. In this study, this is done similarly as in Geistefeldt (2009) by varying the e_{cat3} between 0 and 3 with step size 0.1.

To reduce the computational complexity, e_{cat2} will not be varied such that only category 3 vehicles are estimated in terms of passenger car equivalents. Smoothened values will be used for the individual flows, such that the sum of these flows equals the initial flows classified as in the Flow Type Classification Model according to the 3-minute resolution.

$$C_v = \frac{\sigma}{E} \qquad \textit{Eq. 28}$$

In which

$$\begin{aligned} C_v &= \ Coefficient\ of\ variation\ (-) \\ \sigma^2 &= \ Variance\ of\ the\ sample\ (-) \\ E &= \ Expectation\ of\ the\ sample\ (veh/h) \end{aligned}$$

The mean and the standard deviation of the sample can be found with the use of the Weibull parameters:

$$E = \beta * \Gamma\left(1 + \frac{1}{\alpha}\right) \qquad \textit{Eq. 29}$$

$$\sigma^2 = \beta^2 \cdot \left\{\Gamma\left(1 + \frac{2}{\alpha}\right) - \left[\Gamma\left(1 + \frac{1}{\alpha}\right)\right]^2\right\} \qquad \textit{Eq. 30}$$

In which

$$\begin{aligned} \alpha &= \ Weibull\ shape\ parameter\ (-) \\ \beta &= \ Weibull\ scale\ parameter\ (veh/h) \\ \Gamma(x) &= \ Gamma\ function\ at\ point\ x \end{aligned}$$

As the Weibull parameters are already estimated for the parametric Kaplan-Meier distribution, the Expectation (E) and Variance (σ^2) of the distribution can subsequently be determined to compute the coefficient of variation C_V. This computation is repeated for different equivalency factors such that the minimal C_v value indicates the corresponding PCE value for the respective road section. This process has been visualized for the PQF in Figure 40 and the QDF in Figure 41 for the A12. It can be observed that the minimum C_v leads to a PCE of 1.5 for the PQF and a similar PCE of 1.5 for the QDF.

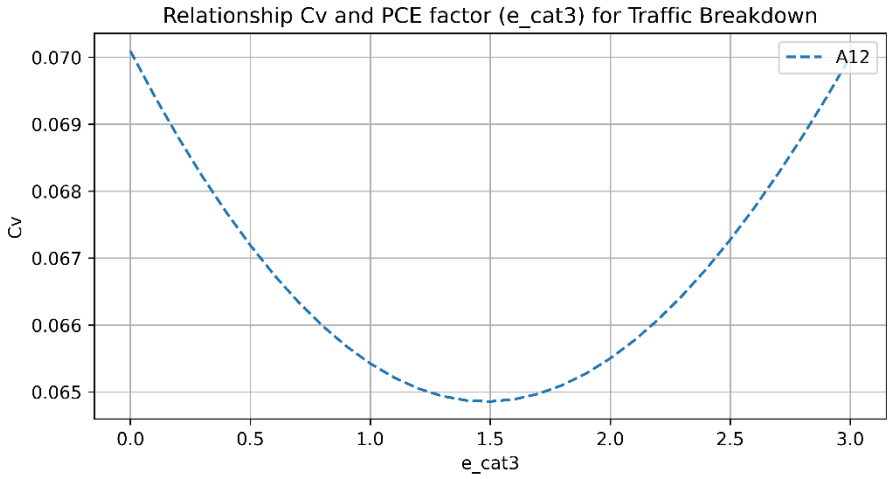


Figure 40 - A1: Relationship C_v and PCE factor for Traffic Breakdown

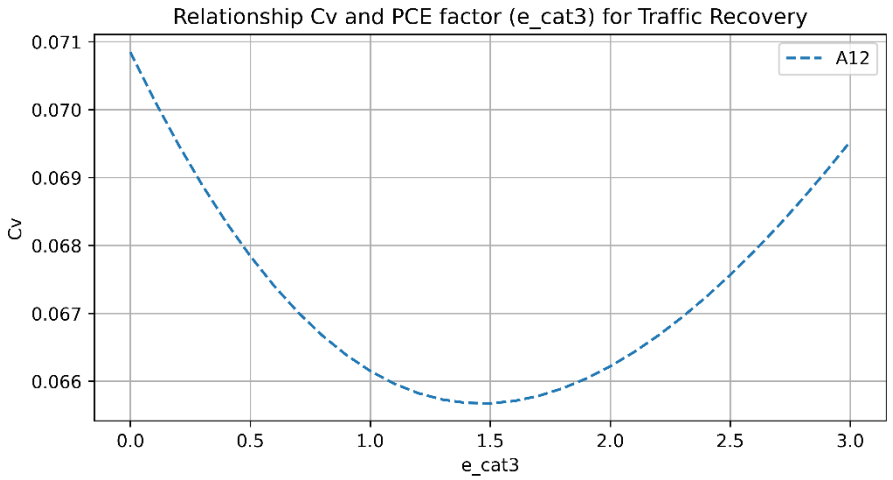


Figure 41 – A1: Relationship C_v and PCE factor for Traffic Recovery

3.4.5 Statistical Tests

To conclude how and to which extent heavy vehicles impact the capacity drop, it is insightful to test for correlation between variables regarding capacity drop and heavy vehicle share. The Pearson correlation test can be used to test for linear correlation between two variables. However, the following three conditions should be satisfied to perform a Pearson correlation test (Daamen, 2010):

- Both variables should have at least had an interval measurement scale
- The sampling distribution should be normally distributed
- A linear relationship should be assumed

Subsequently, a t-test is used to test if the correlation coefficient is significantly different from zero and hence that there is evidence of an association between the variables. The test statistic (t_r) follows the t-distribution with n – 2 degrees of freedom and is computed by:

$$t_r = \frac{r_{X,Y}\sqrt{n-2}}{\sqrt{1-r_{X,Y}^2}} \qquad \textit{Eq. 31}$$

In which:

$$\begin{aligned} t_r &= \ Test\ statistic \\ r_{X,Y} &= \ Correlation\ coefficient \\ n &= \ Sample\ size \end{aligned}$$

However, if the normality assumption is not met, the non-parametric test alternative, Spearman correlation, should be used. In this case, it is required that the two variables have a monotonic relation. The test is similarly computed as in Eq. 31, but with replacing the actual values by their ranks instead.

To test for normality, the Shapiro-Wilko test has been used to test whether the sample comes from a normally distributed population. The test statistic can be computed by:

$$W = \frac{(\sum_{i=1}^{n} \alpha_i X_{(i)})^2}{\sum_{i=1}^{n}(X_i - \overline{x})^2} \qquad \textit{Eq. 32}$$

In which:

$$\begin{aligned} W &= \ Test\ statistic \\ \alpha &= \ Shapiro-Wilk\ constant \\ \overline{x} &= \ Sample\ mean \\ X_i &= \ i-th\ value\ of\ variable\ X \\ X_{(i)} &= \ i-th\ smallest\ number\ in\ the\ sample \end{aligned}$$

These values are compared to the critical value, depending on the degrees of freedom, with an alpha level of 0.05. If the found p-value is less than the alpha value, the null-hypothesis is rejected and the data tested is not normally distributed, However, if the sample size is sufficiently large, the test can be significant even when it only shows small deviations from the normal distribution. Therefore, it is recommended to do a second test by establishing a Q-Q plot to check if the normality assumption is falsely rejected. However, since checking multiple Q-Q plots per site takes too much time, it is decided to execute both tests.

The variables that will be tested for correlation can be found in Table 12 and are motivated below.

Table 12 - Relationships statistically tested

	VARIABLE 1	**VARIABLE 2**
INTRA-LOCATION	Aggregated heavy vehicle flow at PQF	Kaplan-Meier estimated Capacity drop
INTRA-DAY	PQF CAT3	Capacity drop
	PQF	Capacity drop

The Product Limit Method allows for testing aggregated correlations between the different freeway sections. As the average heavy vehicle share is known at the moment of breakdown and recovery. This number can be tested for correlation with the capacity drop estimated from the Kaplan-Meier distribution curves. Although each site has different characteristics nor was measured at different bottleneck sections, a rough estimation of the correlation between the heavy vehicle flow at the PQF and the capacity drop can be estimated.

Besides the intra-location methodology, the data gathered from the Flow Type Classification Model also allows for intra-day analysis of each site. As both the capacity drop and the heavy vehicle share show some variability from time-to-time between days. Therefore, the question arises whether the number of heavy vehicles during a breakdown does relate to the capacity drop. To establish this test, the PQF breakdown value should be time-dependent on the QDF value. This method uses a deterministic approach rather than a stochastic approach used to test between locations.

The dataset created in the Flow Type Classification Model will be the basis of the analysis. As both the data and time have been collected for each classified flow, the PQF and QDF within a respectable time window will be assumed dependent. However, the Flow type classification model does not recognize dependence between the PQF and QDF. It only recognizes a drop between a certain speed as a PQF and a recovery above a certain speed as QDF. Therefore, the dataset must be modified slightly to be useable. For simplification, if multiple PQF or QDF have been recognized at a single date, only the first recognized values are kept in the dataset. The reason is that the following speed drops or recoveries are likely caused by stop-and-go waves and are difficult to match as subsequent PQF and QDF measurements. The next requirement is that the PQF time and QDF time should be within a time window of 10 to 90 minutes, such that the dependency assumption holds. Subsequently, the capacity drop has been computed as the PQF minus the QDF and the average share of heavy vehicles as the breakdown intensity of heavy vehicles. A subset of a dataset can be found in Table 13.

Table 13 - Example of time-dependent PQF – QDF dataset (A4)

DATE	PQF (VEH/H)	PQF CAT3 (VEH/H)	QDF (VEH/H)	QDF CAT3 (VEH/H)	PQF TIME (HH:MM)	QDF TIME (HH:MM)	CAP-DROP (VEH/H)	DELTA TIME (HH:MM)
…	…	…	…	…	…	…	…	…
190916	3180.0	180.0	3260.0	260.0	16:04	16:26	-80.0	00:22
190917	4280.0	420.0	3980.0	220.0	15:36	16:56	300.0	01:20
190918	3480.0	140.0	3120.0	220.0	16:02	16:17	360.0	00:15
190919	3900.0	280.0	3620.0	240.0	15:28	16:20	280.0	00:52
190925	3600.0	280.0	3520.0	200.0	16:02	16:21	80.0	00:19
…	…	…	…	…	…	…	…	…

Now, several statistical tests can be executed to test for correlation between variables. First, the prescribed normality test will be performed on the capacity drop and PQF heavy vehicle flow (PQF CAT3). Q-Q plots of some of the tested variables are drawn, to investigate the normality assumption.

As is observed in the Q-Q plots, the PQF CAT3 on the A1 in Figure 42 and the Capacity Drop on the A1 in Figure 43, a rather straight line with some deviations from the normal distribution can be observed in the tails. For this example, the Shapiro-Wilk test for PQF normality shows $p > 0.05$, while the HV shows $p < 0.05$. However, it can be concluded from both Q-Q plots that both samples are normally distributed.

The same process has been repeated for other locations. Subsequently, similar Q-Q plots were observed as in the examples shown. Only some small deviations in the tails were observed. After checking multiple probability plots, only normally distributed assumptions were accepted. However, since not all variables were visually tested, both the t-test and spearman correlation test will be used to test for both normally distributed and non-normally distributed values.

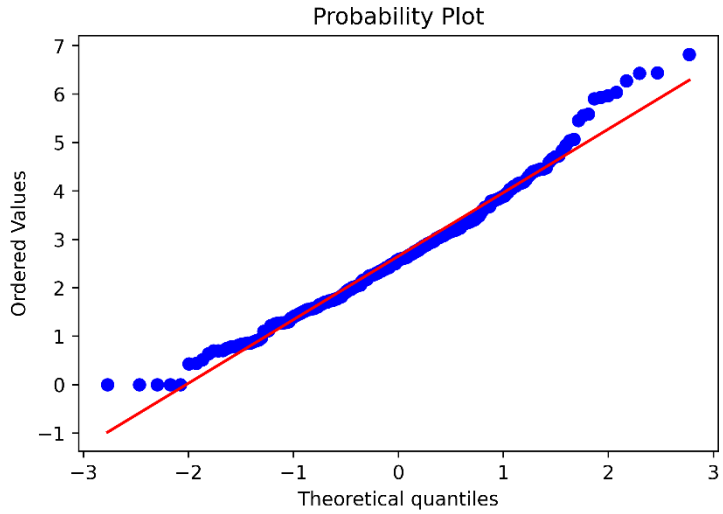


Figure 42 - Q-Q plot PQF CAT3 (%) - A1

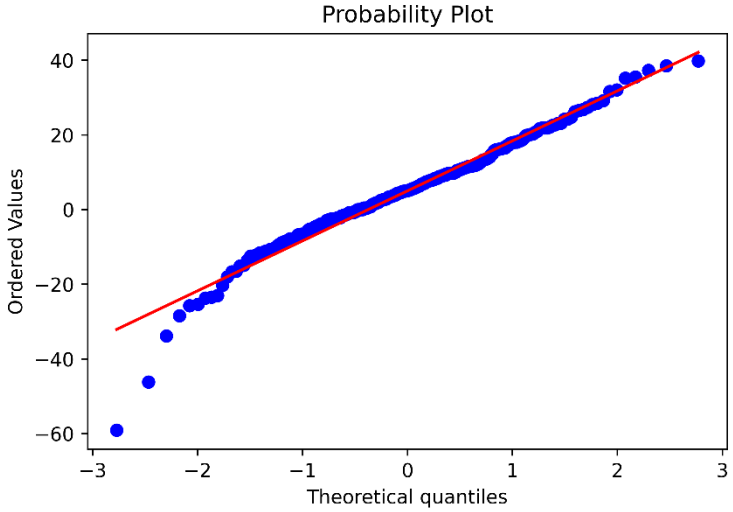


Figure 43 - Q-Q plot Capacity Drop (%) - A1

3.5 Results

The methodological investigation of the empirical dataset was carried out in the chapters above. First, the results of the Product Limit Method will be presented. Insightful observations between the probability curves will be highlighted. Aggregated results of the measurement periods will be compared for different locations. The results on a year-to-year basis will also be analysed for certain locations. Second, the passenger car equivalency factor of heavy vehicles was computed for each freeway section. Third, the statistical testing procedure was performed as elaborated in 3.4.5. The statistical tests will be performed both intra-location as intra-day. As the overall heavy vehicle share does differ significantly per location on aggregate, as well as along the day, insightful information about correlations is provided.

3.5.1 Product Limit Method Results

The elaborated Product Limit Method in chapter 3.4.2, explained based on the A1 site has been executed for all sites described in Table 9. The results of the estimation of the Kaplan-Meier curves for both the breakdown and recovery distributions can be observed in Table 14. On each row, the selected freeways are listed, including information about the type of bottleneck in brackets. Furthermore, the parameters have been estimated for both the PQF and QDF. In the rows, the number of uncensored measurements observed by the Flow type classification model is displayed. Based on the number of uncensored values, the Kaplan-Meier curves were drawn. Next, the A and B parameters of Eq. 7 are listed. The A parameter represents the shape parameter, while the B represents the scale parameter. It turns out that the shape parameter of the PQF ranges from 11.5 to 19.4 with an average of 15.3, while the shape parameter for the QDF ranges from 12.0 to 19.8 with an average of 15.6. The scale parameter B ranges between 6570 to 4693 for the PQF excluding the 5-lane freeway A2. For the QDF, this value ranges between 5422 to 3559.

After estimation of the Weibull parameters, the mean (μ) and standard deviation (σ) of the fitted Weibull curves are computed. The mean is slightly lower than the scale parameter B and ranges between 4531 to 6282 for the PQF and 3428 to 5259 for the QDF when excluding the A2 site. The standard deviation ranges from 256 to 597 for the PQF, while this ranges between 243 to 446 for the QDF, also excluding the A2 site. The capacity drop computed resulted in values ranging from 14% to 29.5% with an average of 22.2%.

The two sites (A1, A12) with a bottleneck only consisting of an off-ramp had an average capacity drop of 14.8%. The capacity drop for the four sites (A2, A6, A50, and A58) with only an on-ramp bottleneck had a considerably higher capacity drop of 21.6%. The four sites (A4, A15A, A15B, and A15C) where both an off-ramp and on-ramp were observed had an average capacity drop of 25.1%.

Table 14 - Estimated parameters empirical data of capacity distributions

FREEWAY		UNCENSORED VALUES	A	B (VEH/H)	Σ (VEH/H)	μ (VEH/H)	CAPACITY DROP (%)
A1	PQF	1155	14.5	6407	**503**	**6180**	14.0
(OFF-RAMP)	QDF	751	12.7	5531	**488**	**5312**	
A2	PQF	2307	15.9	10784	**790**	**10432**	19.8
(ON-RAMP)	QDF	1971	17.4	8634	**575**	**8375**	
A4	PQF	1490	11.5	5290	510	5062	29.9
(BOTH)	QDF	932	15.4	3681	274	3558	
A6	PQF	401	19.4	5147	311	5006	22.4
(ON-RAMP)	QDF	322	19.8	3991	237	3884	
A12	PQF	4556	18.4	4936	314	4794	15.5
(OFF-RAMP)	QDF	2360	18.1	4172	268	4052	
A15A	PQF	1225	15.5	5453	404	5272	22.4
(BOTH)	QDF	571	14.9	4238	326	4091	
A15B	PQF	1486	16.6	4991	348	4834	21.4
(BOTH)	QDF	451	16.0	3924	283	3796	

A15C (BOTH)	PQF	381	12.3	6570	**597**	**6282**	26.6
	QDF	368	12.0	4809	**446**	**4608**	
A27 (BARRIER)	PQF	563	13.1	5528	256	4593	29.9
	QDF	869	13.4	3907	316	3769	
A50 (ON-RAMP)	PQF	1560	15.7	4748	347	4592	25.3
	QDF	781	13.8	3559	292	3428	
A58 (ON-RAMP)	PQF	1611	14.9	4693	361	4531	19.1
	QDF	1429	18.2	3780	243	3671	

The estimated breakdown distribution functions can be observed in Figure 44 and the estimated recovery distribution functions can be observed in Figure 45. The confidence intervals of the graphs can be observed in Appendix G. It can be observed that the A2 has a much higher flow rate for both breakdown and recovery as it is a 5-lane road compared to the other 2-lane roads. The rule of thumb capacity of 2100 veh/h/lane matches the median of this distribution function, at about 10500 veh/h. A higher distribution function can also be observed for the A1 and A15C, as their median matches the 3-lane road rule of thumb capacity of 6300 veh/h. The median of the other roads ranges between 4500 to 5500, which matches the expected capacity of a 2-lane road. The standard deviation is higher for roads with a higher estimated capacity, e.g. the A1, A2, and A15C. The standard deviation for the 2-lane capacity roads also varies considerably as multiple distribution functions cross each other.

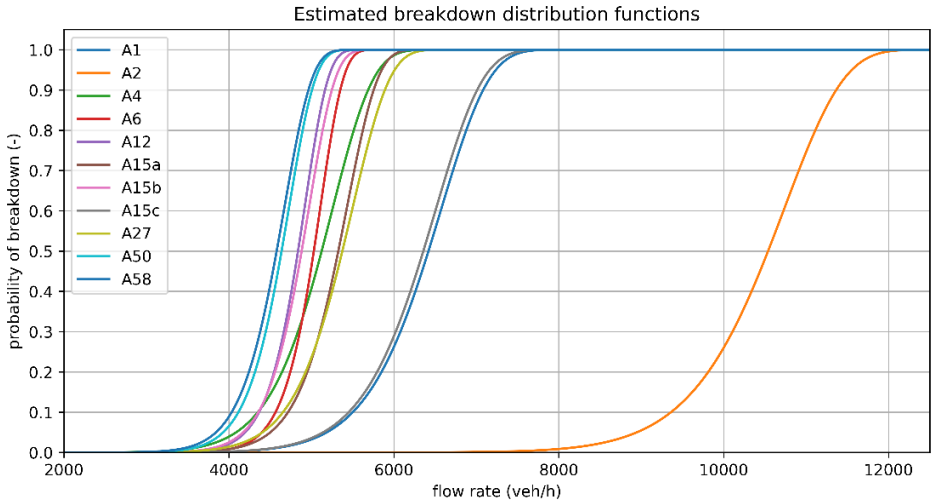


Figure 44 - Estimated breakdown distribution functions (empirical data)

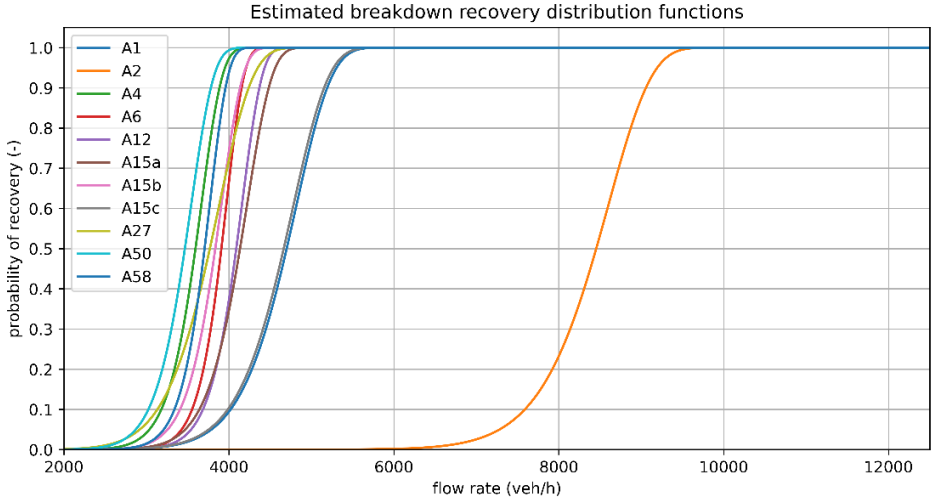


Figure 45 - Estimated breakdown recovery distribution functions (empirical data)

As observed in the capacity drop computation, the estimated breakdown distribution functions have higher flow rates than the estimated recovery distribution function for all sites, which results in the observed capacity drop. The standard deviation of the recovery functions is lower compared to the breakdown functions. In Figure 46 and Figure 47 the 2-lane road observations are zoomed in to get insight into the shape of the distribution functions and how they relate to each other. In the breakdown distribution functions, it can be observed that the A4 and A27 have a more oblique shape compared to the other functions, i.e. a higher standard deviation. This also appears for the A27 in the recovery distribution function where again the shape appears to be oblique compared to the other functions, which are almost parallel. The A4 and A27 were also the two sites where the highest capacity drop was measured. Furthermore, after comparing the order of the lines in both distribution functions, it can be noted that the order is not equal which shows how seemingly similar roads do differ considerably in terms of capacity.

An overview of all classified measurements including detected breakdown and recovery measurements, comparable to the example figures of each site shown in 3.4.2, can be found in Appendix C.

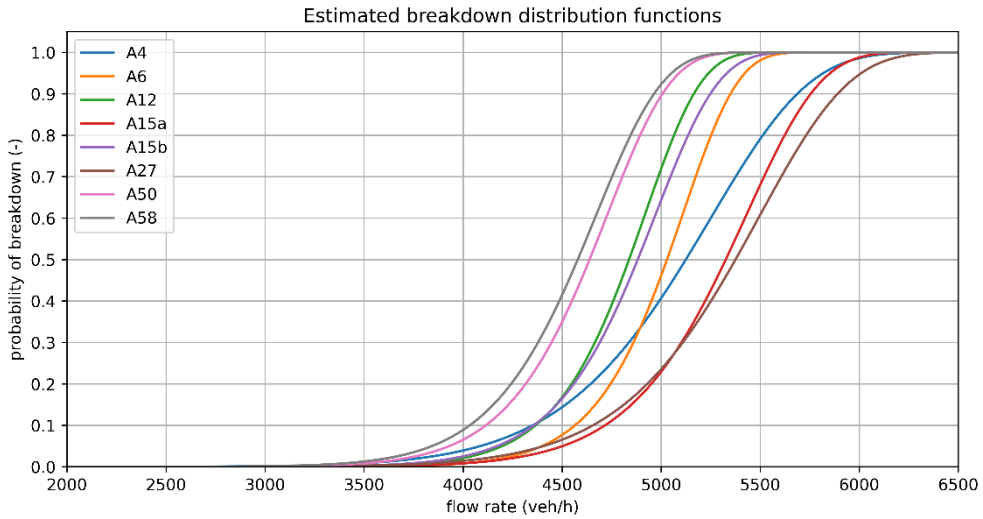


Figure 46 - Estimated breakdown distribution functions (zoomed)

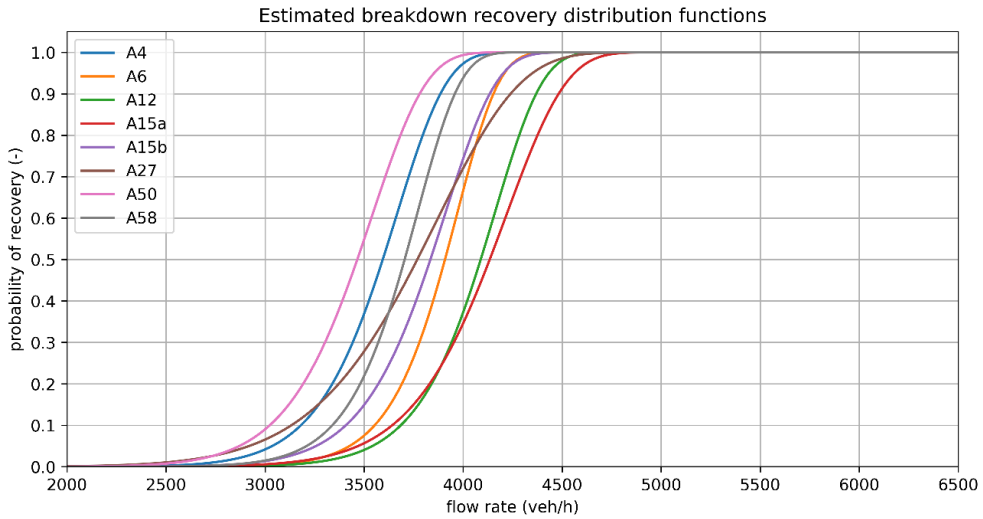


Figure 47 - Estimated breakdown recovery distribution functions (zoomed)

Boxplots have been created to identify the mean, standard deviation, and outliers of the breakdown and recovery flow and can be found in Appendix D. These boxplots match the estimated distribution functions. To get more insight into the traffic composition at each of the sites, boxplots indicating the percentage of

category 3 vehicles are displayed in Figure 48 and Figure 49 respectively for breakdown and recovery. As can be observed in Figure 48, the mean of the percentage of heavy vehicles at breakdown ranges from 2% at the A6 to 9% at the A15. The outliers are mainly at the upside, as most of the minimums reach 0%. The percentage of heavy vehicles during recovery boxplots in Figure 49 is almost identical to the boxplots in Figure 48. The main difference can be observed in the outliers. Therefore, it can be assumed that the traffic flow composition during breakdown and recovery stays equal after aggregation, which is not a certainty the case for individual time-dependent measurements. Furthermore, it should be noted that the measured flow composition is an aggregation of the 3-minute mean at which the breakdown or recovery is classified. The boxplots with the percentage of category 2 vehicles can be found in Appendix D. The main observation is that all sites have on average 4% to 6% category 2 vehicles, while the A4 has considerably more category 2 vehicles at about 26% of the total traffic flow. The speeds during breakdown and recovery are also visualized in boxplots and can also be found in Appendix D. A considerable difference in speed at congestion and recovery between the sites can be observed. On average, the breakdown speed is at about 80 km/h, while the recovery speed is at just below 70 km/h.

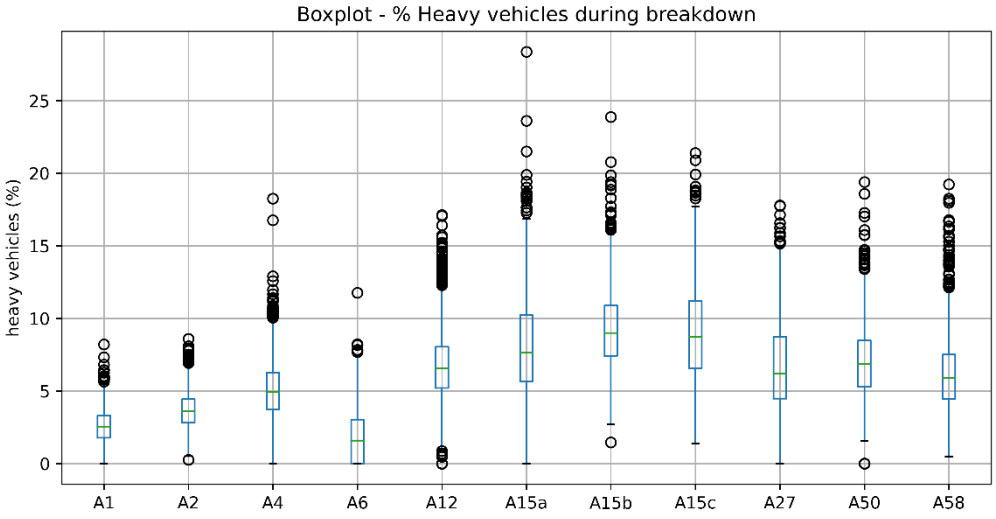


Figure 48 - % Heavy vehicles during breakdown

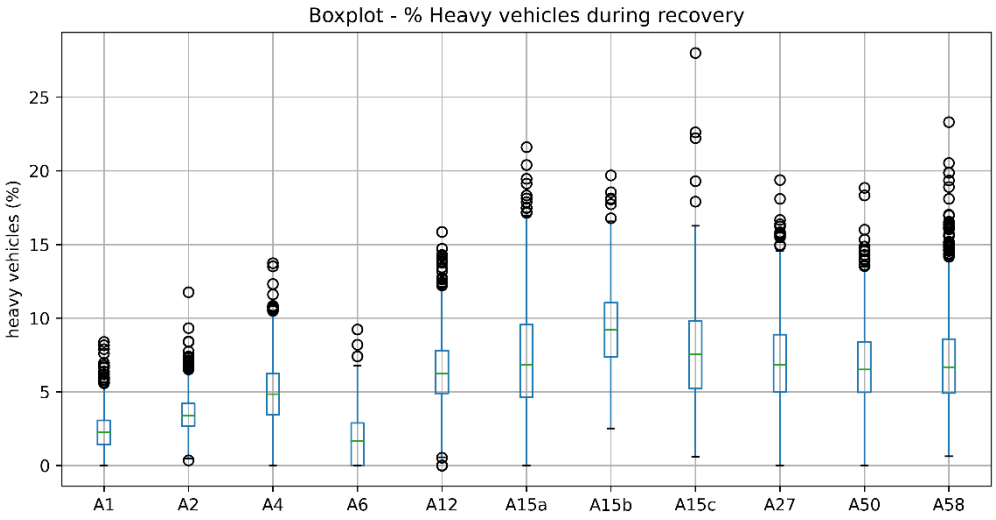


Figure 49 - % Heavy vehicles during recovery

3.5.1.1 Between-year analysis

Besides aggregating the results over the entire study period for each site, the results have also been aggregated per year as explained in 3.4.2. An overview of the results can be found in Appendix E. Here, the number of detected breakdowns and recoveries are displayed. The median of the estimated distribution function, the standard deviation, and the aggregated heavy vehicle flow for both the PQF and the QDF. Also, the capacity drop has been computed per year.

The number of breakdowns or recoveries shows a considerably lower value in some of the first or last measured years. This is because the analysed years did not have appropriate loop detector data for the entire year. Thus, lower values than expected are probably caused by the analysis not containing a year full of data. Therefore, the years with substantially fewer measurements should be considered cautiously.

There appears to be a decreasing trend in the capacity for the analysed sites. In Figure 50, four roads that show a clear decreasing capacity over the years can be observed. The medians of the PQFs are used to create a scatter plot with linear trendlines. For each of the trendlines, the respective R-squared is displayed near the line. Notice that the R-squared values are between 0.86 and 0.90, which is a good fit of the line to the data. Most other roads do also show a decreasing capacity but not as significant as the ones displayed. The slope shows a decreasing trend between 28 to 90 veh/h per year.

The flow of heavy vehicles at a breakdown has also been measured yearly. Again, the same four roads can be found in Figure 51 showing the flow of heavy vehicles during a breakdown. The R-squared values vary between 0.74 and 0.98, which therefore do slightly less fit the data. An upward trend in heavy vehicle flow for selected sites can be observed. The slope shows an increasing trend between 10 to 14 veh/h per year.

The standard deviation of both the PQF and QDF of the estimated distribution functions does not show a trend in general but stays rather within a certain margin. The values do slightly fluctuate throughout the years. However, some deviating values can be observed. There does not seem to be a relation between those outliers across different sites in the same year.

Regarding the QDF median and QDF of heavy vehicles, similar trends can be observed as explained above. A decreasing trend in recovery capacity can be observed from the QDF median, while an increasing trend in the QDF of heavy vehicles can be observed. This is again the case for most roads, however, not all do show a clear similar trend.

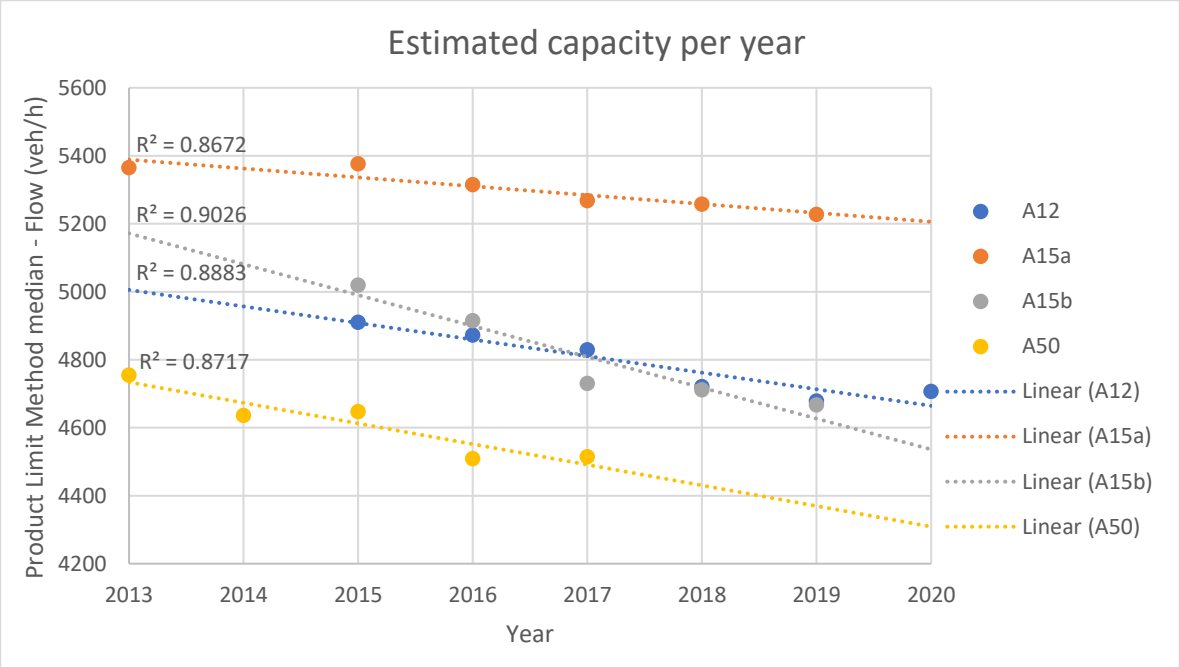


Figure 50 - Estimated capacity per year

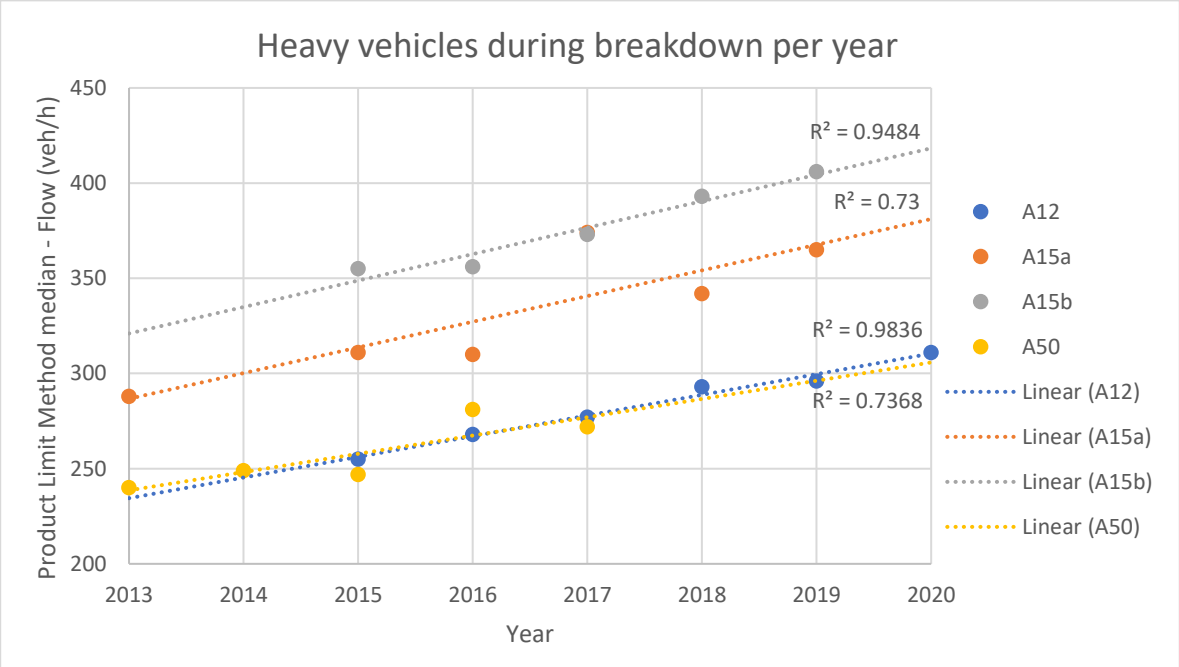


Figure 51 - Heavy vehicle flow during breakdowns per year

3.5.2 Passenger Car Equivalent Estimation

The method for estimating the PCEs on the estimated distribution functions is elaborated in 3.4.4. The optimization of the PCE for category 3 vehicles is displayed for the breakdown flow in Figure 52 and the recovery flow in Figure 53. On the Y-axis, the covariance (Cv) is displayed, which is the ratio of the standard deviation divided by the expected value of the distribution function, see Eq. 28. The covariance was computed for different PCE values, ranging from 0 to 3 with steps of 0.1 for each site. Thus, by varying the PCE of heavy vehicles, the flow changes, which then leads to a different distribution function for each of the varying PCE

values. When looking at the shape of the curves, the A15C is different compared to the other curves. An only increasing covariance factor when increasing the equivalency factor can be observed. This means that for increasing the equivalency factor, the standard deviations increase relatively more than the expected value of the distribution function. The other sites have a similar relationship, with an almost flat curve and horizontal close to the vertex. The covariance varies between 0.06 and 0.12 for both the breakdown and recovery. Furthermore, the order of the curves does seem to be completely different for both the breakdown and recovery.

The minimum of the curves has been determined and those values have been combined in Table 15. When excluding the A15C, the PCE values vary between 0.7 and 2.0 for the breakdown flow and between 1.1 and 2.25 for the recovery flow. Except for the A4 and A15A during a breakdown, all sites have an optimized PCE above 1 as expected. When comparing the equivalency factor for individual sites at both breakdown and recovery, a few things are noted. For the A1, A2, A6, A15B, A27, and A58 a lower equivalency factor for the breakdown was found compared to during recovery, with both factors being above 1. Only in the case of the A15A and A50, the equivalency factor during breakdown is higher than the equivalency factor during recovery.

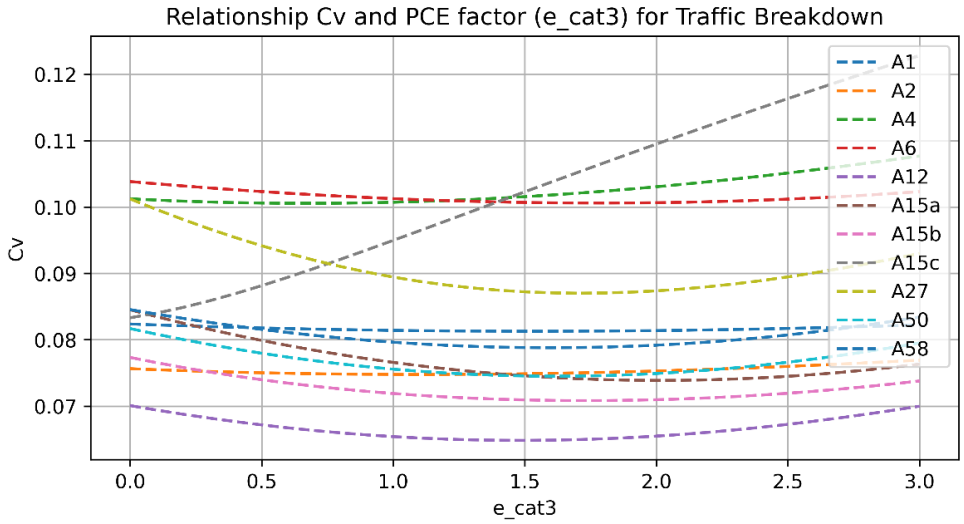


Figure 52 - Relationship Cv and PCE factor for traffic breakdown

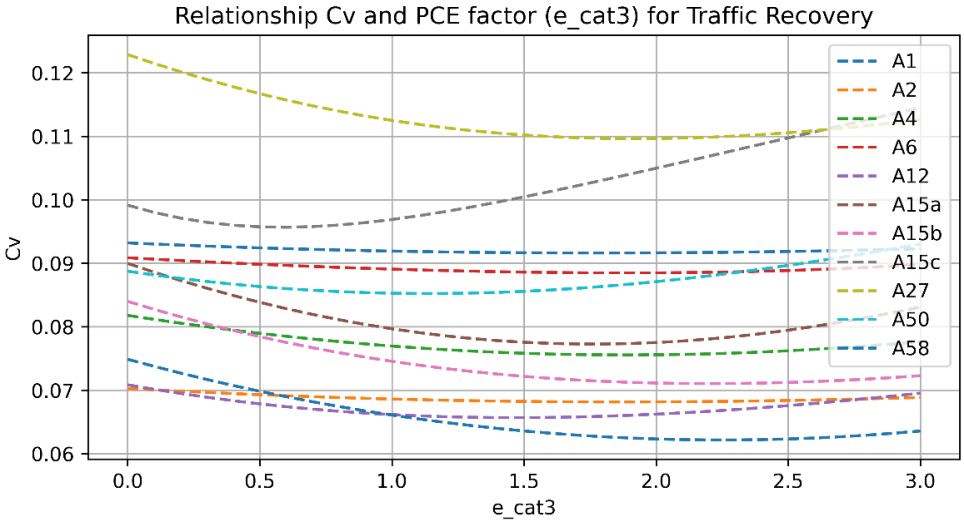


Figure 53 - Relationship Cv and PCE factor for traffic recovery

Table 15 - Minimized Cv and corresponding equivalency factor

CLASSIFICATION		A1	A2	A4	A6	A12	A15A	A15B	A15C	A27	A50	A58
PQF	Minimal C_v	0.081	0.075	0.101	0.101	0.065	0.074	0.071	0.083	0.087	0.075	0.079
	e_cat3	1.5	1.1	0.7	1.8	1.5	2.0	1.7	-	1.7	1.6	1.6
QDF	Minimal C_v	0.091	0.068	0.076	0.088	0.066	0.077	0.071	0.096	0.110	0.085	0.062
	e_cat3	1.8	1.9	1.9	1.9	1.5	1.8	2.2	0.6	1.9	1.1	2.3

Besides the computation of the equivalency factor of category 3 vehicles, it is also possible to compute the equivalency factor for category 2 vehicles. See Table 2 for the exact difference between the two vehicle classes. To get insight into the relationship between both vehicle categories in terms of equivalency factor, the covariance has been computed for both the e_cat2 and c_cat3 from 0 to 3 with steps of 0.1. This made the computational work increase by the power of 2, increasing the number of computations from 30 to 900. Since this requires quite some computational power as each computation requires estimating a new distribution function, it was decided to perform this computation only for the breakdown flow of the A12.

The result has been visualized in Figure 54 and shows a 3D plot with on the z-axis the e_cat2 equivalency factor, on the y-axis the e_cat3 equivalency factor and on the x-axis the computed covariance. A 3D plane has been drawn, which has been coloured dark blue for the lowest Cv value and red for the highest Cv value. In the case of the A12, increasing the e_cat2 value leads to a higher covariance compared to increasing the e_cat3 value. The minimum covariance is found at an e_cat2 factor of 1.2 and an e_cat3 factor of 1.5. This indicates that the passenger car equivalent of category 2 vehicles is lower compared to category 3 vehicles.

However, it should be kept in mind that this visualization is nothing more than an exploration into the ratio between the equivalency of category 2 and category 3 vehicles. The results have only been computed for the A12, which shows expected equivalency factors. Roads with substantially different flow composition, e.g. A4 would be interesting for comparison of results.

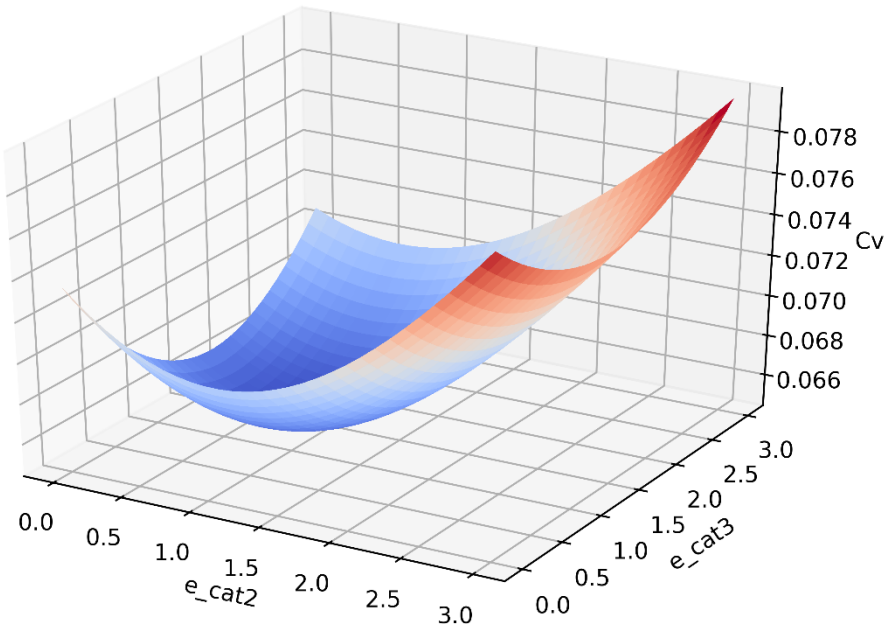


Figure 54 - 3D plot A12 PQF PCE e_cat2 and e_cat3

3.5.3 Statistical Tests

In the last part of the empirical analysis results, statistical tests are performed. The relationship between the variables in Table 12 is tested with different statistical tests. The tests will focus on the intra-location relation between the analysed sites and the intra-day relation for individual sites.

3.5.3.1 Intra-location

First, the relation between the capacity drop and the share of heavy vehicles is plotted against each other in Figure 55. The aggregated capacity drop percentage is put on the y-axis, while the aggregated heavy vehicle share is put on the x-axis. The decision is made to use the percentages here rather than the absolute values as this allows for trend exploration. For each location, the aggregated variables are plotted for their respective measurement period in the graph. Roads with relatively few heavy vehicles such as the A6, A1, and A2 can be found on the left side of the graph while more truck-dominant sites such as the A15 can be found on the right side of the graph. Hereby, it should be noted that the heavy vehicle share presented here is the value that is measured at the moment of breakdown, and matches the mean of the boxplot presented in Figure 49. The capacity drop value is identical to the value presented in Table 21, rather than aggregating the yearly results presented in Appendix E.

For these measurements, a linear trend line has been computed. The trend line does not fit the data very well, as can be observed and because the R-squared of the line only has a value of 0.1, however an upward-going trend can be observed. The main outliers of the trendline are the A4, A12, and A27. In the case of the A4, the share of category 3 vehicles is relatively low, while the amount of category 2 vehicles is considerably higher. The trendline is currently very sensitive to each location, removing either measurement would change the slope of the line. Additional low percentage heavy vehicle share locations should verify the observed upward-going trend.

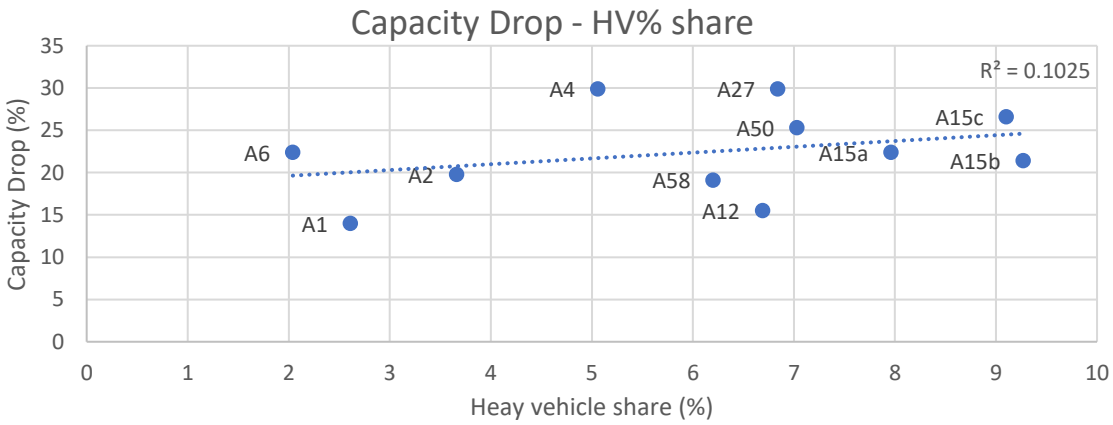


Figure 55 - Correlation capacity drop - HV% share (intra-location)

3.5.3.2 Intra-day

Second, the intra-day results are statistically tested on correlation per location. To perform the Pearson correlation, the variables tested should be normally distributed. This has been tested with the Shapiro-Wilk test and Q-Q plots as the sample sizes are substantially large. As checking each variable visually takes a lot of time, it is decided to both perform the Pearson correlation test and Spearman test, in case one of the samples is not normally distributed without visual check. The small portion of Q-Q plots checked were all normally distributed similarly to the two examples shown in Figure 42 and Figure 43.

In Table 16, an overview of the sample size and the Shapiro-Wilk test p-value for both the capacity drop and heavy vehicle share can be observed. As explained in chapter 3.4.5, only the breakdown and recovery of a single day are used for the capacity drop computation. Therefore, the sample size shrank considerably compared to the uncensored values obtained in Table 14, only existing of assumingly related breakdown-recovery pairs.

Regarding the capacity drop variable, the A15b, A27 and A50 all have p-values above 0.05, while the other sites have p-values below 0.05. Regarding the A2, A15a and A27 all have p-values above 0.05, while the other sites have p-values below 0.05. After checking Q-Q plots of some of the sites, it was concluded that the checked samples were normally distributed.

Next, both the Pearson and Spearman correlation tests have been executed to test for correlation between the capacity drop and the flow of heavy vehicles at breakdown. The correlation coefficient can range in value from -1 to +1 and indicates the strength. Higher strengths have a value that deviates more from 0. The direction of correlation is indicated by the sign. A negative sign indicates a negative correlation while a positive sign indicates a positive correlation. The Pearson correlation examines the linear relationship between both variables, while the Spearman correlation test assesses whether the two variables can be described using a monotonic relationship function. A p-value higher than 0.05 is not statistically significant and indicates a strong evidence for the null-hypothesis.

Only the p-values of the Pearson test of the A1, A15a, A15c, and A58 are below the alpha value of 0.05 as can be observed in Table 16. Only the p-values of the A1. A15c and A58 showed significance for the Spearman test. The correlation is considered weak as the factor only ranges between -0.16 and -0.22. The A15c showed a moderate correlation of -0.48.

Table 16 – Empirical Analysis statistics

	A1	A2	A4	A6	A12	A15A	A15B	A15C	A27	A50	A58
SAMPLE SIZE	247	656	333	165	587	166	150	163	106	199	445
SHAPIRO CAPACITY DROP P-VALUE	< .001	< .001	< .001	.018	< .001	< .001	.103	< .001	.218	.187	< .001
SHAPIRO PQF CAT3 P-VALUE	< .001	.333	< .001	.001	< .001	.561	.013	< .001	.505	.005	< .001
PEARSON CORRELATION	-0.22	-0.04	0.02	0.13	-0.07	-0.16	-0.08	-0.48	-0.05	-0.12	-0.17
PEARSON P-VALUE	< .001	.267	.682	.107	.078	.043	.335	< .001	.608	.102	< .001
SPEARMAN CORRELATION	-0.22	0.01	0.05	0.12	-0.05	-0.15	-0.03	-0.37	-0.02	-0.13	-0.14
SPEARMAN P-VALUE	< .001	.718	.363	.088	.218	.051	.686	< .001	.868	.063	.004

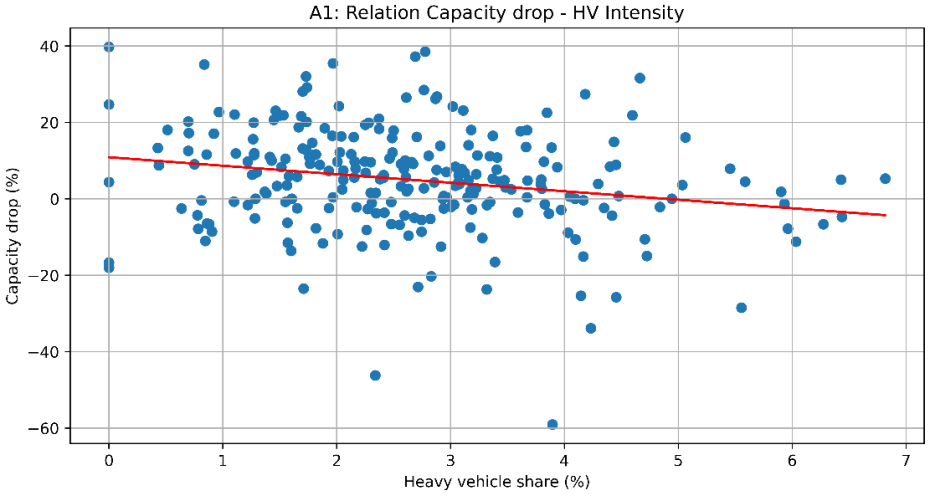


Figure 56 - Relation Capacity drop (%) - HV intensity (%) at the A1

In Table 17, the results of the statistical analysis on the capacity drop – pre queue flow can be observed. First, the pre queue flow is tested for normality. After inspection of the Q-Q plots, it was concluded that all sites had a normally distributed PQF.

Subsequently, the Pearson and Spearman correlation test values and p-values can also be observed in Table 17. It turned out that all sites had p-values below the threshold of 0.05, such that the null-hypothesis is rejected. The Pearson correlation factors obtained were rather high, ranging from 0.37 to 0.78. The Spearman correlation factors were almost identical, confirming the linear relationship.

The samples were also tested for correlation of only positive capacity drop values. However, the normality assumption of the PQF's and Capacity drops did not hold. Therefore, it was decided to not remove the negative capacity drop values.

In Figure 57, an example of the correlation between the absolute capacity drop and breakdown intensity of the A4 can be observed. The linear relationship becomes clear as the measurements all are located around the trendline. The wide variance of pre queue flow compared to the queue discharge flow becomes clear. When the capacity drop is close to 0 veh/h, the breakdown flow and queue discharge flow are equal at roughly 3500 veh/h. When the capacity drop is about 1000 veh/h, the pre queue flow is roughly 4500 veh/h, meaning that the queue discharge flow remains equal for different breakdown flows at the A4. This phenomenon does not occur at all sites. The A12 does have a 1:2 recovery-breakdown ratio. So, when the breakdown intensity is about 1000 veh/h higher, the recovery intensity is roughly 500 veh/h higher.

Table 17 - Capacity Drop - PQF intensity statistics

	A1	A2	A4	A6	A12	A15A	A15B	A15C	A27	A50	A58
SAMPLE SIZE	247	656	333	165	587	166	150	163	106	199	445
SHAPIRO PQF P-VALUE	< .001	< .001	< .001	.075	< .001	.835	.007	< .001	< .001	.189	< .001
PEARSON CORRELATION	0.37	0.77	0.84	0.78	0.71	0.73	0.64	0.68	0.59	0.73	0.71
PEARSON P-VALUE	< .001	< .001	< .001	< .001	< .001	< .001	< .001	< .001	< .001	< .001	< .001
SPEARMAN CORRELATION	0.38	0.78	0.85	0.77	0.72	0.73	0.64	0.70	0.65	0.72	0.71
SPEARMAN P-VALUE	< .001	< .001	< .001	< .001	< .001	< .001	< .001	< .001	< .001	< .001	< .001

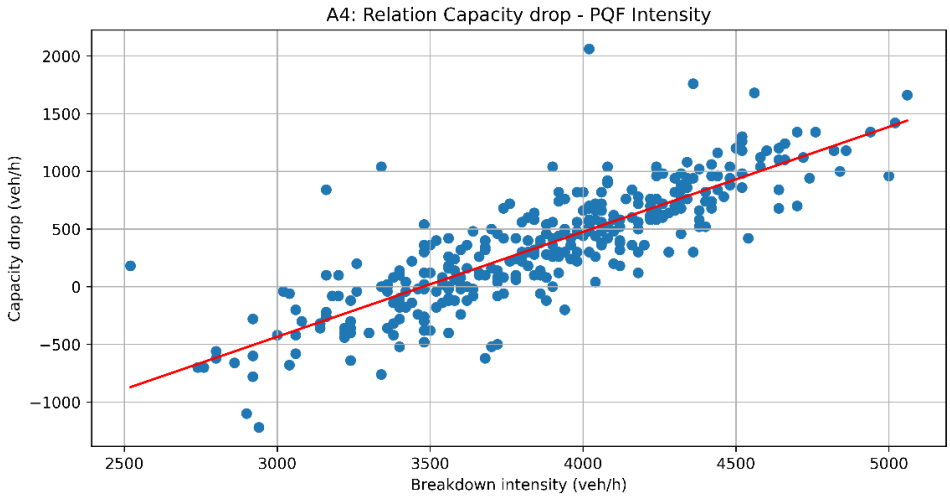


Figure 57 - Relation capacity drop - PQF intensity at the A4

3.6 Conclusions and Insights

A methodology was presented to empirically investigate the impact of heavy vehicle share on the capacity drop. First, the loop detector had to be prepared such that it could be used for the detector of breakdown and recovery events. It appeared that the 3-minute moving average technique gets sufficiently reduces volatility within the data. Compared to the 5-minute aggregation used in comparable studies (Brilon et al., 2005; Geistefeldt, 2009), the precise capture of breakdown is independent of the chosen temporal-resolution. Besides, the lower 3-minute resolution grasps more realistically the raw measured data, with less flattening of the peak flows.

The number of suitable bottlenecks for the analysis was restricted. This was mainly due to the fact that the methodology designed demanded that two loop detectors, which are able to measure separately flows of different vehicle categories, were present at both the bottleneck and downstream. Mainly 2-lane sites were selected, which limits the possibility in comparing the capacity drop with respect to the number of lanes. The number of lanes is a crucial factor in the overtaking possibilities of passenger cars and the utilization of space.

As part of the Calculation Framework, the Flow Type Classification Model was developed. This model classified measurements prior to a breakdown or recovery event as censored, and the breakdown or recovery event itself as uncensored. In classifying the uncensored events, it turned out that searching for the maximum flow value in a time window of four minutes when the critical speed was hit, resulted in more accurately selecting the maximum flow as visualized in Figure 32. The classifications in Appendix C show the difference in average breakdown speeds of the sites.

The fitted Weibull distribution onto the non-parametric Kaplan-Meier Product Limit Method showed a good fit, equivalent to comparable studies (Brilon & Zurlinden, 2003). For the estimation of the capacity drop, the estimated capacity value at the 0.5 probability was used for both the breakdown and recovery distribution functions. The distribution functions showed, for the Green confidence interval, distribution functions up to a probability of 0.5 to 1, except for the A27. The recovery distributions functions reached up to a probability above 0.95. This secures that the estimated capacity drop at the 0.5 probability level is within the confidence interval of the estimated distribution functions such that the methodology can be considered solid.

Interesting results were obtained from the Kaplan-Meier Product Limit Method. The estimated capacity distribution functions showed, as expected, an increasing mean and standard deviation as the number of lanes increases for the selected sites. However, both the A4 and A27 had substantially higher standard deviations compared to locations with 2-lanes. At both these locations, the highest capacity drop was observed at 29.9% while the share of heavy vehicles was average at about 5 – 6% at the uncensored breakdown event. At the other locations, a capacity drop between 14 – 26.6% was observed, averaged at 20.7%. The share of heavy vehicles during breakdown ranged from approximately 2% at the A6 to 9% at the A15. No clear relationship was yielded between the observed mean or standard deviation of the PQF and QDF or the capacity drop and any of the geometric, traffic and control conditions of the sites.

The intra-location statistical tests showed a non-correlated relation between the aggregated share of heavy vehicles and the capacity drop. Although a slight uptrend could be observed, removing any individual measurement would lead to a different slope. The intra-day Pearson correlation test only showed significance for the A1, A15a, A15c, and A50 with correlation factors between -0.16 and -0.48. Therefore, it can be concluded that the extent to which the magnitude of the capacity drop is attributable to the influence of heavy vehicle percentages is neglectable. The drawbacks of the methodology and the exclusion of external factors will be discussed in the Discussion in chapter 6.

The correlation between the pre queue flow and the capacity was tested similarly and showed significance at all locations. For each site, a positive relation was yielded, indicating that the higher the pre queue flow breakdown value, the higher the resulting capacity drop. The wide range of breakdown intensities observed during speed drops, seem to recover with a rather constant queue discharge flow intensity as the speed

increases above the recovery speed threshold. This implies that capacity drop can be roughly predicted when the breakdown intensity is observed.

Passenger Car Equivalency factors were computed by minimizing the covariance of the distributions for PCE values ranging between 1 to 3. This was done for both the estimated breakdown and recovery distributions. A PCE-value below 1 was observed for the PQF at the A4. From a theoretical perspective, this would imply that heavy vehicles have less spatial impact than passenger cars during breakdown. This is practically very unlikely and the reason for this result could be that the share of category 2 vehicles is considerably higher compared to the other sites. At most other sites, a PQF PCE between 1.5 – 2 was observed, while a QDF PCE between 1.5 – 3 was observed. For most sites, a higher PCE was observed at queue discharge compared to pre queue, which implies that heavy vehicles have more spatial impact during the recovery event compared to the breakdown event.

The distribution functions were also computed annually for the selected sites. For a selection of four sites; the A12, A15a, A15b, and A50; a clear increase in heavy vehicles during breakdown was observed annually. The sites had an increase of 10 to 14 veh/h per year. This went hand in hand with a decrease in capacity between 28 to 90 veh/h per year. The decreasing capacity trend at these sites had a minimum of $R^2 = 0.87$, while the increasing heavy vehicles had an $R^2 = 0.73$. The decrease in capacity was greater than the increase of heavy vehicles even after multiplication of the computed PCE values for the respective sites. This implies that besides the increase in heavy vehicles, another phenomenon is responsible for the decline of breakdown capacity on highways.

4 SIMULATION

The simulation study phase focuses on filling the gaps in the empirical analysis. This is done by exploring several scenarios that could not be explored with empirical data. First, the methodological approach, including the overall goals, will be discussed in 4.1. Subsequently, the model preparation will be described in 4.2 including the different scenarios tested. The results will be presented in 4.3 Lastly, the main conclusions will be elaborated in 4.4.

4.1 Methodological Approach

As discussed in the Introduction, the simulation study extends on the literature study and empirical data analysis stage to answer the final research questions:

Research Question 3.1: **How do different traffic flow compositions and operational characteristics of freight traffic impact the capacity drop?**

Research Question 3.2: **Which measures can be taken to reduce the drop rate on truck-dominant highway sections?**

The relationship with previous research activities has been visualized in Figure 58. These questions extend the empirical investigation of the relation between the capacity drop and heavy vehicles. Empirical data includes systematic influence of external factors, which are embedded into the geometric, traffic, and control situations. Therefore, the simulation stage aims to investigate the impact on capacity drop while keeping external factors out of the equation. The purpose is to investigate the impact of future scenarios with increased freight traffic and with adjusted physical and operational characteristics of freight traffic investigated in 2.2.1. The results of the simulation will be compared to the empirical capacity drop estimations. Finally, measures to reduce the capacity drop will be explored.

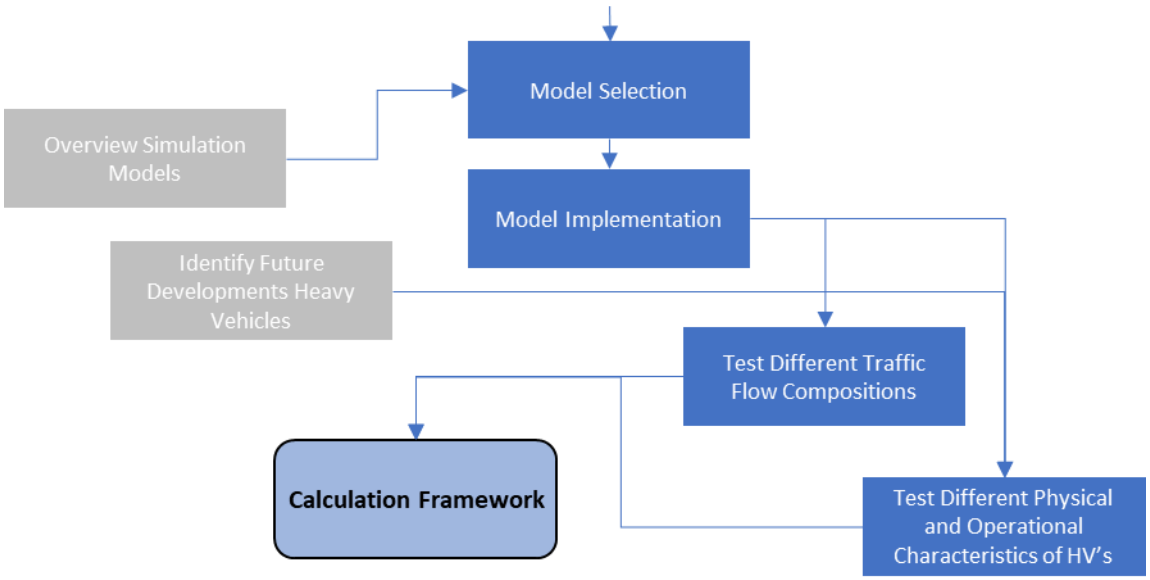


Figure 58 – Partial research framework (Simulation stage)

4.2 Model Implementation

The model that has been selected is the OpenTrafficSim simulation model. The main reason is that despite all viable models (VISSUM and MOTUS) have similar performance, the adaptability of the models is different and most favourable for OTS. The extended elaboration of the decision and comparison between different simulation models can be found in 2.4 Overview Simulation Models. In this chapter, the implementation of the OTS environment will be elaborated in 4.2.1, followed by the scenario's investigated elaborated in 4.2.2.

4.2.1 OpenTrafficSim

The lane change model implemented, consists of multiple on-ramps and off-ramps for a 2-lane road along the A20 (Schakel et al., 2012). The model has been calibrated and validated in both congested and free-flow conditions and sensitivity analysis showed that the model is appropriate for both scenarios. However, the model fit in the congested scenario shows some variable output results as this is highly dependent on the input demand. Since congested conditions are the main research point in this study, the demand should be well considered. Furthermore, to fit the current speed limit in the Netherlands, the speed limit was lowered from initially 130 km/h to 100 km/h.

To simplify the existing scenario, it was decided to only use a single on-ramp, the Ter Heide on-ramp, of the existent study. This leads to the simulation environment depicted in Figure 59. The main road starts out-of-bounds in this figure and has a total length of 2000m up to the start of the merge. The merging section has a length of 320m, while the on-ramp has a total length of roughly 620m.

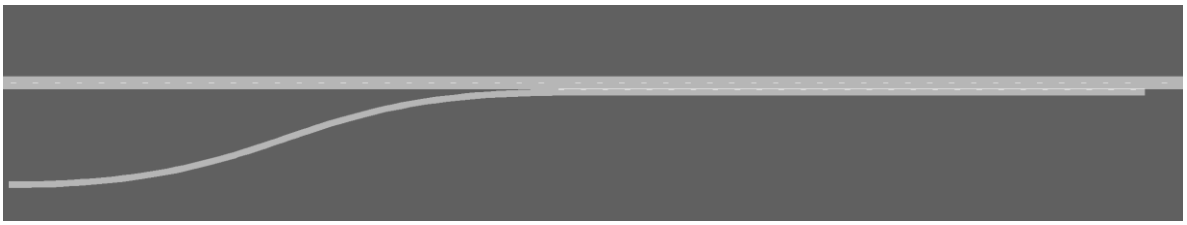


Figure 59 - Simulation environment

The calibrated and validated parameters of the lane changing behaviour, car-following behaviour, and other driver models in the model can be found in Schakel et al. (2012). These values will not be touched since this would require new calibration and validation of the scenario. The only parameters that need to be set up are the mainline and ramp demand. The goal of each simulation run is to create a congested traffic state, which is solved as the demand decreases. The model is set up in such a way that the demand can be changed every 15 minutes, with a total of 6 periods. Demands will be chosen in such a way, that they can be used for all the different scenarios, satisfying mentioned criteria for all the different traffic compositions. The main criterion is that each simulation run starts in free flow but becomes congested as demand increases which allows a breakdown measurement. This congested state is followed by a free flow traffic state which allows for a recovery measurement. The flow type classification model created in 3.4.1 will be used to detect these breakdowns and recoveries. The main drawback of the classification model is that the performance decreases when stop-and-go-waves occur around 60 – 70 km/h. Therefore, the demand will be set in such a way to limit these stop-and-go-waves. Furthermore, the more truck-dominant scenarios will need more time to throughput all generated traffic after the bottleneck has resolved. Therefore, the last two periods will be without a demand. Different demands have been tested to the criteria, and the final demand pattern that fits these can be observed in Table 18.

Table 18 - Demand pattern simulation

	0 – 15 MIN	15 – 30 MIN	30 – 45 MIN	45 – 60 MIN	60 – 75 MIN	75 – 90 MIN
MAIN ROAD	1500	2500	4000	3000	1500	1000
ON-RAMP	100	200	500	400	200	100

As mentioned, the flow type classification model will be used for the analysis of breakdowns and recoveries. To re-use this model and compare the results with the empirical data, similar output data should be created. Therefore, a loop detector has been located within the bottleneck, 100 meters downstream from the start of the merging section. This detector measures the 1-minute aggregated speed, flow, and share of heavy vehicles of the traffic stream similarly to Dutch loop detectors. However, a single simulation run would only provide a single breakdown and recovery. Therefore, multiple runs must be performed to accumulate sufficient data. Each run will generate slightly different vehicle configurations, so-called seeds, to accumulate multiple breakdowns and recoveries for each scenario.

After testing the more truck-dominant scenarios, the phenomenon occurs that the number of right-keeping heavy vehicles was too high. This led to congestion on only the inner lane due to vehicles merging between a platoon of heavy vehicles while the vehicles on the outer lane did not face congestion. Therefore, the V_{gain} parameter which motivates the desired anticipation speed difference for a full lane change has been lowered to increase the spread of heavy vehicles across all lanes. In the calibrated model, this value has been estimated to be 50 km/h. However, after testing several lower values, it was decided to change the V_{gain} for passenger cars to 20 km/h and the V_{gain} for heavy vehicles to 30 km/h such that the phenomenon was not observable anymore.

4.2.2 Scenario Analysis

After the model has been set up, the different scenarios that are tested will be elaborated. First, the number of simulation runs for each scenario must be decided. This depends on the indicators that need to be computed for the evaluation of the performance. Identical methods for the empirical data evaluation will be used, which require statistical testing combined with the Product Limit Method. Therefore, a minimum of 20 runs per scenario is decided. This will enable sufficient breakdowns and recoveries to estimate the Product Limit Method and besides test for correlation between the heavy vehicle share during breakdown and the capacity drop.

4.2.2.1 Traffic flow composition scenarios

First, traffic flow compositions with increasingly more heavy vehicles will be simulated. An overview of the different scenarios can be observed in Table 19. Scenario A1 to A4 are currently realistic scenarios, while A5 and A6 are more extreme scenario scenarios. The fraction of both cars and heavy vehicles can be changed individually for the mainline and the on-ramp. It is decided to set the share of heavy vehicles for the mainline and on-ramp both accordingly to each traffic flow composition scenario.

Table 19 - Traffic flow composition scenarios

Scenario	Share freight traffic (%)
S1	0%
S2	5%
S3	10%
S4	15%
S5	20%
S6	40%

Furthermore, the configuration of heavy vehicles and cars is important, as discussed in 2.2.2, and is defined in this model with a reversible Markov chain. In case the configuration would be highly correlated, the heavy vehicles (HV) and cars (C) would be simulated as (C, C, C, C, C, HV, HV, HV, HV, HV). If the correlation is set close to 0, a non-correlated configuration would be created, e.g. (HV, C, C, HV, C, C, C, C, HV, C). The equation used to compute the configurations can be observed in Eq. 33. In the model, a correlation c_i of 0.4 was found after calibration and will be used in these scenarios.

$$p_{ij} = ss_i + (1 - ss_i) * c_i \qquad \text{Eq. 33}$$

where

p_{ij} = the probability state i returns after state i
ss_{ii} = the steady-state probability of state i
c_i = correlation of state i

4.2.2.2 Changed physical and operational characteristics scenarios

Second, different operational and physical characteristics of heavy vehicles will be simulated in line with the developments explored in 2.2.1. The main changes are the decreased heavy vehicle length and the increased acceleration capability due to the electrification of vehicles. Currently, the length of heavy vehicles has been set at 15 meters within OpenTrafficSim. In the future scenario, a length of 10 meters has been chosen. Furthermore, the maximal desired car following acceleration of heavy vehicles has been fixed to 0.4 m/s^2, while the car acceleration has been calibrated at 1.25 m/s^2. Based on the current acceleration capabilities of electric vehicles, the acceleration of cars will be increased to 1.5 m/s^2 while the acceleration of heavy vehicles will be increased to 0.65 m/s^2.

The parameters for each scenario can be observed in Table 20. Scenario B1 and B2 only take into consideration changed physical or operational characteristics under current occurring traffic compositions. Scenario B3 is a combination of both these scenarios. Finally, scenarios B4 and B5 are similar to scenario B3 but with an increased share of freight traffic.

Table 20 - Heavy vehicle characteristics scenarios

Scenario	Share freight traffic (%)	Heavy vehicle length (m)	Acceleration cars (m/s^2)	Acceleration heavy vehicles (m/s^2)	Configuration correlation (-)
F1	10%	**10**	1.25	0.4	0.4
F2	10%	15	**1.5**	**0.65**	0.4
F3	10%	**10**	**1.5**	**0.65**	0.4
F4	20%	**10**	**1.5**	**0.65**	0.4
F5	40%	**10**	**1.5**	**0.65**	0.4

4.2.3 Calculation Framework

The performance indicators that will be used to evaluate the impact of the different scenarios on the capacity drop are similar to the empirical data analyses and can be divided into two sub-categories:

1. Product Limit Method estimation
2. Statistical tests for correlation

4.2.3.1 Product Limit Method

The Product Limit Method methodology applied has been described in chapter 3.4.2. As the data has been prepared in an identical way as the empirical data, it was possible to reuse the method without changes. The only difference is the amount of data that is used for computing the Kaplan-Meier curve. The empirical data consists of data of at least multiple years, while the simulation data always represents 20 situations comparable to a rush hour.

The yearly Product Limit Method could not be applied, since the simulation is not measured time dependent. Therefore, aggregating over simulation runs does not provide useful information. The capacity drop computation is again similar to the methodology used and can be found in chapter 3.4.3.

4.2.3.2 Statistical Tests

The statistical tests described in 3.4.5 will also be used for the simulation data, and the relations tested can be observed in Table 21. It turned out that the sample size was not large enough to test the intra-day correlation coefficient. Therefore, only the intra-location correlation will be investigated.

Table 21 – Relationships statistically tested

	VARIABLE 1	VARIABLE 2
INTRA-LOCATION	Aggregated heavy vehicle flow at PQF	Kaplan-Meier estimated Capacity drop

4.3 Results

The simulation results follow from the executed methodology elaborated in chapter 4. First, the estimated Product Limit distribution functions for the different traffic flow composition scenarios will be elaborated. As mentioned, the methods used are similar to those used in chapter 3, the Empirical Analysis, which becomes useful in comparing the results. Second, the results of the changing operational and physical characteristics scenarios of heavy vehicles are elaborated. These scenarios give insight into long-term changing characteristics of traffic flow.

4.3.1 Product Limit Method Results

4.3.1.1 Traffic flow composition scenarios

The changing traffic flow composition scenarios can be found in Table 19, and simulate breakdowns for 5% stepwise increasing heavy vehicle flow. The acquired loop-detector measurements for the 20 runs per scenario were classified similarly as in 3.4.1 for their flow type. Subsequently, the Product Limit Method was applied, and the results can be observed in Table 22. The confidence interval of the estimated Kaplan Meier curves can be observed in Appendix G.

The probability of breakdown within the confidence interval was between 0.6-0.8, except for the S4 (15% HV) scenario. Here, only a probability of just above 0.4 was established. As the 0.5 probability is used for the computation of the PQF and the capacity drop, the displayed values for the S4 scenario are an estimation based on the fitted Weibull distribution.

The number of uncensored values ranges between 6 and 38. As the simulation demand has been prepared with the 10% heavy vehicle scenarios, the scenarios with considerably more or less heavy vehicle share do show less uncensored values as the demand pattern stays equal for all scenarios. The shape parameter A ranges between 5.66 and 26.43 for the PQF, with a decreasing value as the share of heavy vehicles increases. A similar observation is made for the shape parameter of the QDF, which ranges between 10.88 and 29.25. The scale parameter B ranges between 3214 and 4361 for the PQF and seems to keep decreasing as the share of heavy vehicles increases, except for the 15% scenario. The same holds for the scale parameter of the QDF, which ranges between 2449 and 3642 but decreases continuously.

The standard deviation of the estimated Weibull distribution curve of the PQF first increases for the first scenarios from 198 to 439, subsequently 529. The 15% scenario has a considerably higher standard deviation of 715, while the higher scenarios again have a standard deviation of about 560. The standard deviation of the QDF raises gradually, while the 20% and 40% scenarios have a similar standard deviation. The mean decreases from 4272 to 2338 for the PQF with again a considerably higher value at the 15% scenario of 4289. For the QDF, the mean decreases as the share of heavy vehicles increase from 3575 to 2338. The capacity drop value ranges between 16.4% for the 0% heavy vehicle scenario to 34% for the 15% heavy vehicle scenario.

Table 22 - Estimated parameters traffic flow composition scenarios simulated data of capacity distributions

FREEWAY		UNCENSORED VALUES	A	B (VEH/H)	Σ (VEH/H)	μ (VEH/H)	CAPACITY DROP
S1	PQF	11	26.43	4361	198	4272	16.4
(0% HV)	QDF	6	29.25	3643	150	3575	
S2	PQF	29	10.88	4334	439	4137	22.2
(5% HV)	QDF	27	20.88	3316	187	3232	
S3	PQF	32	8.36	4174	529	3939	24.0
(10% HV)	QDF	36	18.90	3096	192	3009	
S4	PQF	25	6.54	4602	**715**	**4289**	34.0
(15% HV)	QDF	38	14.69	2946	228	2843	
S5	PQF	23	7.14	3844	556	3600	24.4
(20% HV)	QDF	32	12.99	2838	246	2728	
S6	PQF	11	5.66	3214	562	2972	21.4
(40% HV)	QDF	15	10.88	2449	248	2338	

The estimated breakdown distributions can be observed in Figure 60. As observed in the mean and standard deviation per Weibull distribution, a decreasing mean and increasing standard deviation can be observed for increased heavy vehicle flow compositions. However, S4 shows a deviating trend with a considerably higher standard deviation and mean than the other scenarios.

Up to the 15% HV scenario, the breakdown capacity of the road fits the rule of thumb 2100 veh/h/lane capacity rather well. The main difference is the standard deviation which increases with the increase of the heavy vehicle share. However, as the share of heavy vehicles further increases in the last scenarios, the breakdown capacity drops rather substantially.

The estimated recovery distribution functions can be observed in Figure 61. As the heavy vehicle share increases, the recovery distribution difference gradually decays. The recovery capacity decreases a bit less as the share of heavy vehicles stepwise increases by 5 per cent. This seems to be opposite to the breakdown capacity, where the capacity increasingly decreases with each 5 per cent heavy vehicle stepwise increase.

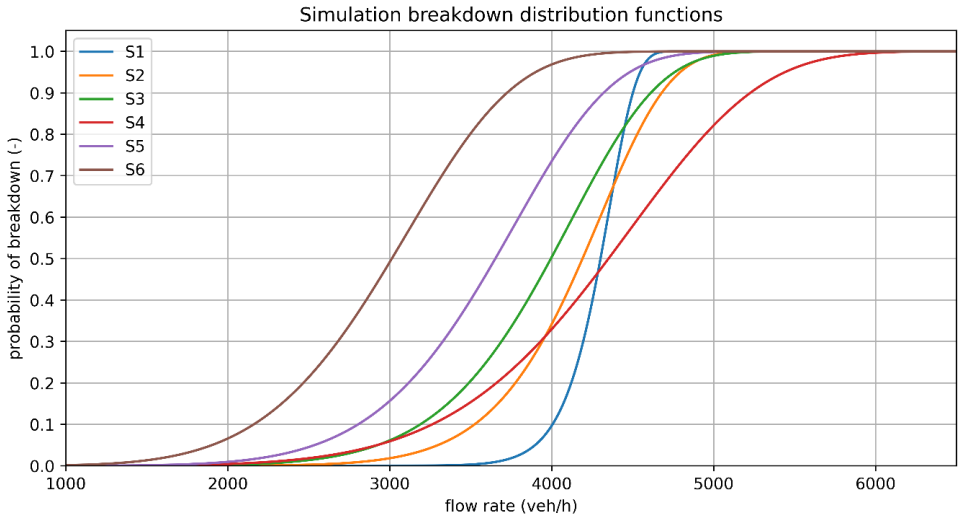


Figure 60 - Estimated breakdown distribution functions (traffic flow composition scenarios)

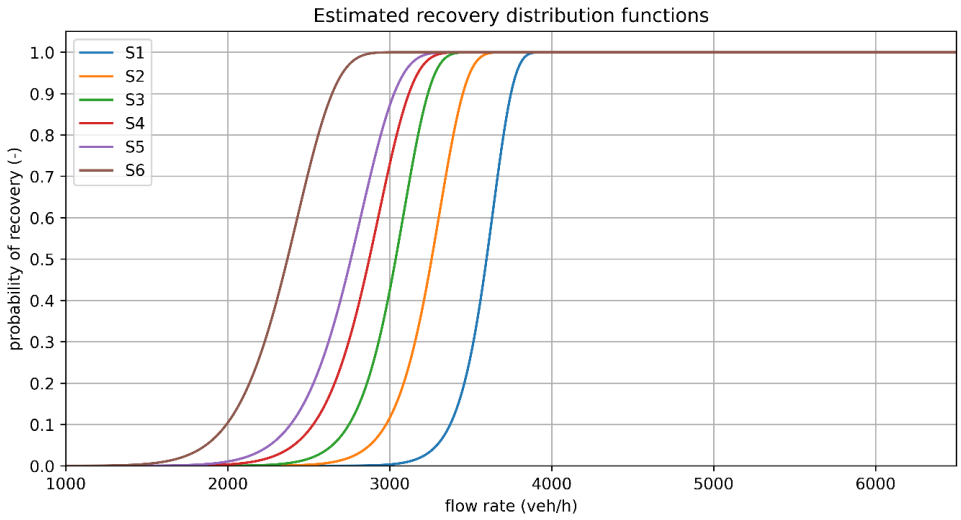


Figure 61 - Estimated breakdown recovery distribution functions (traffic flow composition scenarios)

4.3.1.2 Physical and operational scenarios

In the next scenarios, the future physical and operational characteristics of heavy vehicles are applied, simulated, and put into the calculation framework. The results of the Product Limit Method and the capacity drop computation can be observed in Table 23. First, the base scenario without adaptations to the heavy vehicles is compared to the changed physical characteristics where the size of the heavy vehicles has been changed from 15 meters to 10 meters: scenario F1 – 10% (P). Next, the maximum car following acceleration of the heavy vehicles has been increased from 0.4 to 0.65 m/s^2 as well as the maximum car following acceleration of passenger cars, which increased from 1.25 to 1.5 m/s^2. This scenario is called F2 – 10% (O).

For both scenarios, the number of traffic breakdowns decreased, while the number of identified recoveries increased. The shape parameter A is equal in the base scenario compared to the physical scenario, while the B parameter increases slightly for the PQF. This results in a higher estimated mean capacity of 4101 veh/h compared to 3939 veh/h, while remaining an almost equal standard deviation. The operational scenario shows a lower breakdown capacity mean, combined with a decreased standard deviation. Contrary, the queue discharge flow mean increases while the standard deviation remains equal compared to the base scenario.

Table 23 - Estimated parameters physical and operational scenarios simulated data of capacity distributions

FREEWAY		UNCENSORED VALUES	A	B (VEH/H)	Σ (VEH/H)	μ (VEH/H)	CAPACITY DROP
F0 – BASE – 10%	PQF	32	8.36	4174	529	3939	24.0
	QDF	36	18.90	3096	192	3009	
F1 – 10% (P)	PQF	29	8.49	4343	544	4101	27.5
	QDF	40	19.88	3071	188	2989	
F2 – 10% (O)	PQF	24	11.72	3979	378	3809	9.5
	QDF	40	22.88	3548	184	3465	
F5% (P + O)	PQF	23	16.01	4270	307	4131	10.0
	QDF	21	19.93	3824	225	3722	
F10% (P + O)	PQF	24	13.71	3999	330	3850	7.8
	QDF	32	20.68	3655	208	3561	
F15% (P + O)	PQF	22	8.88	4162	502	3940	13.4
	QDF	34	19.92	3522	208	3429	
F20% (P + O)	PQF	20	7.12	3999	580	3744	14.9
	QDF	34	26.40	3277	148	3210	
F40% (P + O)	PQF	20	9.85	2946	325	2801	0.8
	QDF	32	25.41	2857	134	2797	

The estimated distribution functions are displayed in Figure 62 for the breakdown and in Figure 63 for the recovery. As can be observed, the breakdown distributions for each scenario higher are compared to the recovery distributions. This resulted in a slightly higher capacity drop for the physical scenario and a substantially lower capacity drop for the operational scenario. This is mainly due to the increase in the estimated recovery distribution for the operational scenario.

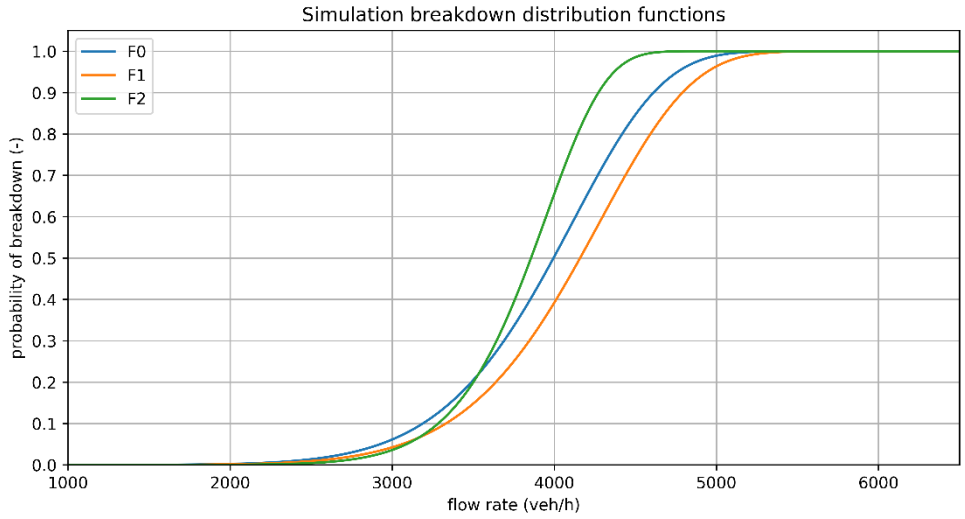


Figure 62 – Estimated breakdown distribution functions (physical and operational scenarios)

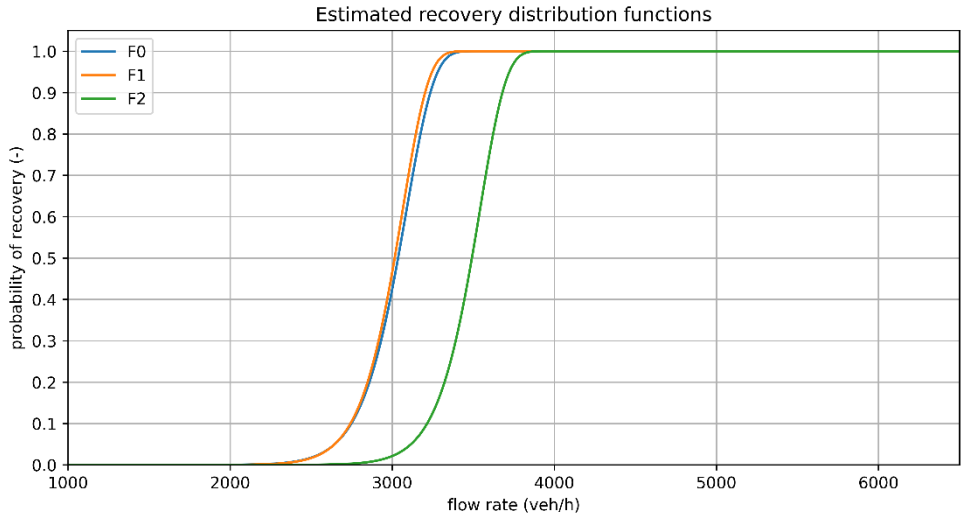


Figure 63 - Estimated recovery distribution functions (physical and operational scenarios)

Next, both characteristics were applied to a similar set of heavy vehicle share scenarios as in 4.3.1.1, increasing 5 per cent stepwise. In Table 23, it can be observed that the number of uncensored measurements keeps constant for different scenarios. Regarding the mean of the estimated Weibull distribution for the PQF, a decreasing trend can be observed, except again for the 15% HV scenario. However, it does not surpass the mean of 5%, which is different compared to the traffic flow composition scenarios. In general, the capacity values of these scenarios do tend to be very close to those values in those scenarios. The standard deviation on the other hand is quite similar for the scenarios, around 320 veh/h, except for the 15% and 20% scenario, where this value is above 500 veh/h. The standard deviation shows a similar pattern for the different vehicle compositions but is considerably lower in the changed combined physical and operational scenarios compared to the regular scenarios. Furthermore, a decreasing trend in the mean QDF was observed while the standard deviation also decreases.

When looking at the breakdown distribution and recovery distribution visualization in Figure 64 respectively Figure 65, a few remarkable observations can be made when comparing the results to the traffic flow composition scenario analysis. First, the decreased standard deviation becomes clear when comparing Figure 64 to Figure 60. The difference of the 5 – 20% heavy vehicle scenarios are combined considerably lower, being about 800 veh/h at the 0.2 probability to about 300 veh/h at the 0.8 probability. This is about 1000 veh/h for both probabilities in the traffic flow composition scenarios.

Furthermore, when comparing the recover distribution functions in Figure 65 to Figure 63, a different pattern is observed. In the traffic flow composition scenarios, a clear decreasing trend was observed for increased heavy vehicle share. This pattern is almost opposite in the recovery distribution functions in Figure 65. Also, the standard deviation becomes smaller as the heavy vehicle share increases in Figure 65 while contrary the standard deviation increases in Figure 63.

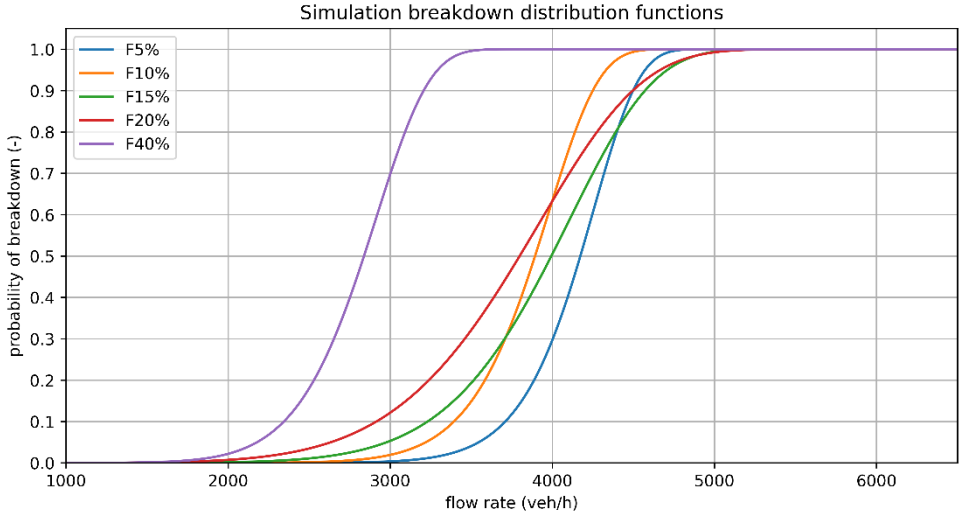


Figure 64 - Estimated breakdown distribution functions (combined physical and operational scenarios)

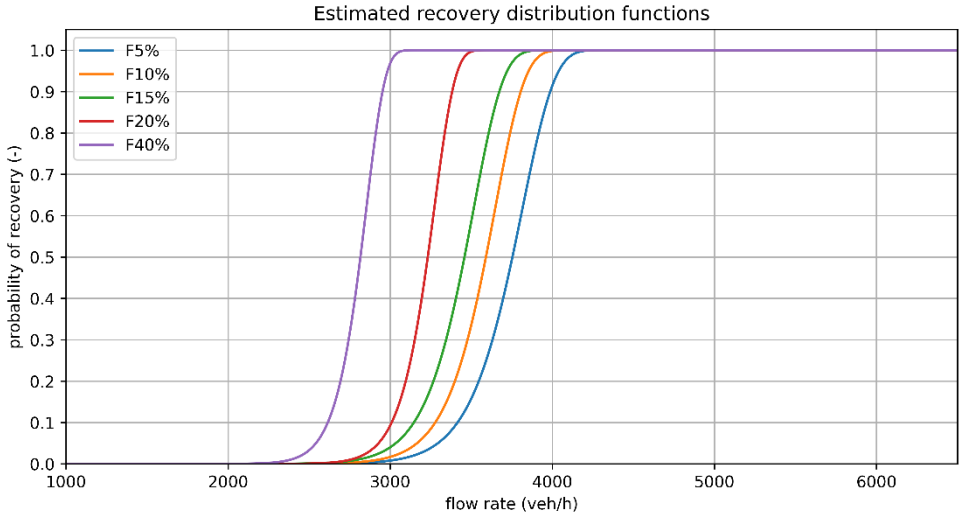


Figure 65 - Estimated recovery distribution functions (combined physical and operational scenarios)

4.3.2 Statistical Test

Now the correlation between the capacity drop and the heavy vehicle share is tested similarly to the tests executed for the empirical data in 3.5.3. The intra-day results will be present in 4.3.2.1.

4.3.2.1 Intra-day

The heavy vehicle flow composition scenarios have been plotted against their observer capacity drop, retrieved from the Kaplan Meier Product Limit Method. The six scenarios presented in chapter 4.3.1 have been supplemented with a few extra measurements at the following traffic flow compositions: 2.5%, 7.5%, 12.5%, and 17.5%, 25%, 30% and 35%. These extra measurements are executed to add more data points to investigate whether a trend exists. The result can be observed in Figure 66.

Through these data points, a 2th-degree polynomial has been plotted. This trendline shows an R^2 of 0.4. Between the 0% and 15% heavy vehicle share scenario, an increasing capacity drop is observed reaching from 16.3% to 34%. The 2.5% heavy vehicle and the 10% heavy vehicle scenario both show a slight decay in capacity drop compared to the previous scenario, while the other scenarios are constantly increasing. After the 15% heavy vehicle scenario, the capacity drop measurements vary widely with a down-going trend. The large variance in capacity drop between the 25% and 30% HV scenario is mainly caused by the high difference in PQF value, which are 3562 veh/h and 2771 veh/h respectively. The 25% HV scenario only reaches a probability of breakdown within the confidence interval of 0.45 with a very high standard deviation of 670 veh/h compared to the standard deviation of the 340 veh/h

As only a few measurements have been obtained beyond the 20% heavy vehicle share scenario, the certainty of the plotted trendline diminishes. When comparing the underlying data regarding the PQF mean and standard deviation up to the 15% HV scenario, the extra measurements fall somewhere in between the measurements displayed in Table 23.

Comparing the results with the observed heavy vehicle share – capacity drop pairs, a few observations can be made. First, the range of the observed heavy vehicle share in the empirical data only ranges between 2 – 10%. The observed capacity drop measurements of the empirical data range between 14 – 29.9%. The measurements in the empirical data and simulation data show the same regularity in the 0 – 10% when fitting a straight line through the measurements. However, it should be noted that the simulated data has completely equal circumstances, except for the heavy vehicle share, which is not the case for the empirical data.

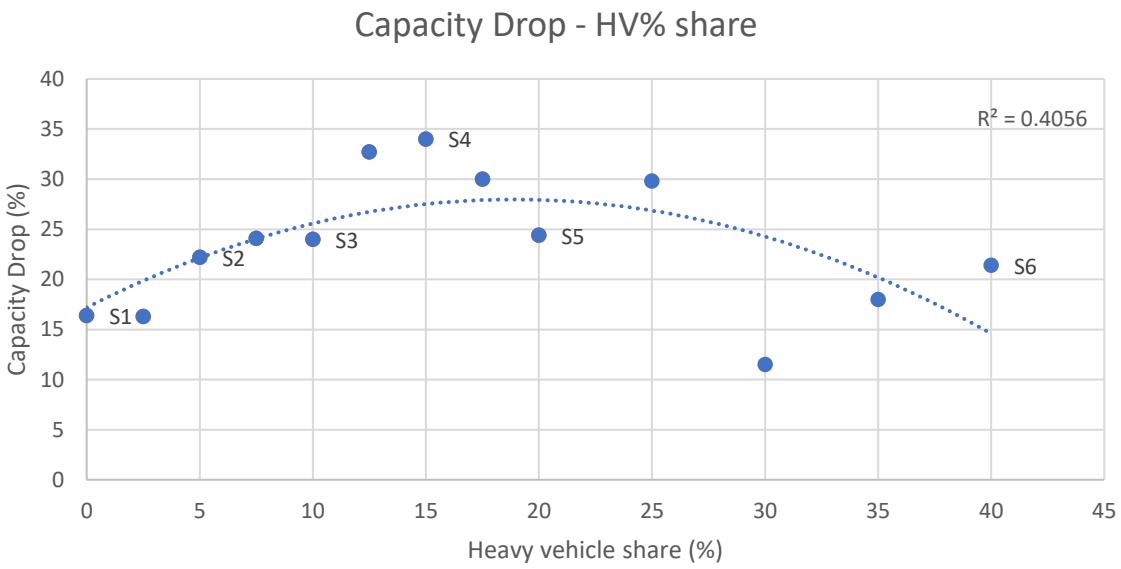


Figure 66 - Capacity drop - HV% share (simulation data)

4.4 Conclusions and Insights

The first step in executing a simulation, is setting up the simulation environment. Although an example scenario of OpenTrafficSim was used and only slightly adjusted to fit the experiments, it was very difficult to get it working due to some specific requirements. As the Calculation Framework was chosen to be re-used, similar data as the empirical data should be gained with the simulation. Locating the loop detectors in OpenTrafficSim led to problems as the loop-detector class enabling macroscopic data collection contained an error with the detection of vehicles. Therefore, unique vehicles were registered multiple occasions sporadically. As the boundaries of the capabilities of the OTS software were tested, it was first thought that the problem was with the correctness of simulation. However, after transposing the data to floating car data, the error was found and a fix was implemented with the help of OTS creator Wouter Schakel.

It was also observed that the OTS software had an very low incentive for heavy vehicles to change lanes. As it was part of the experiment to maintain an equal demand pattern for several heavy vehicle penetration share scenarios, it was required to lower this incentive. It appeared that at high heavy vehicle penetrations, the trucks accumulated at the right lane while the left lane was still in free flow. As the model was calibrated for a 10% heavy vehicle share scenario, the lane change incentive of heavy vehicles (VGAIN) was lower via trial-and-error. The same approach was executed in estimating a demand pattern that would lead to the sequence of free-flow – congestion – free flow such that the capacity drop could be measured in each scenario.

The different heavy vehicle penetration scenarios (0% - 40%) gained insights in how the breakdown and recovery distribution curves change for increasingly more heavy vehicles and subsequently the capacity drop. In assessing the effect of heavy vehicle share on the capacity drop, it was observed that the capacity drop increases from 16.4% up to 34% from 0% HV scenario to the 15% HV scenario. This increase is mainly due to the increase in standard deviation of the estimated breakdown distribution from 198 veh/h to 715 veh/h respectively. This shows that the share of heavy vehicles influences the capacity drop when external factors are kept equal. The breakdown probability becomes more difficult to estimate as the share of heavy vehicles increases, while the recovery probability at a certain flow and heavy vehicle share is quite predictable. This is confirmed by the confidence intervals of the estimated Weibull distributions of both the breakdown and recovery distributions. The first has a certainty up to 0.6 probability, while the latter has a certainty above 0.95 probability.

Next, the reduced truck size scenario and the increased acceleration capability scenario were compared to the base scenario, all at 10% heavy vehicles. This showed that decreased truck size scenario positively impacts the PQF mean, while the increased acceleration scenario positively impacts the QDF mean.

Subsequently, the scenarios were combined with both adaptations. Comparing these results to the heavy vehicle penetration composition scenarios, a few observations were made. The standard deviation of the PQF increased slightly which is useful in estimation the breakdown probability. Second, the standard deviation of the QDF gradually decreases while the opposite was found for the heavy vehicle penetration scenarios. This implies that at low heavy vehicle penetrations, estimating the recovery probability can be done more accurately when the heavy vehicle size and acceleration capabilities are not changed. However, at higher heavy vehicle penetration scenarios, the standard deviation becomes lower for the adapted scenarios. The tipping point is between the 10 – 15% HV scenario.

Finally, the aggregated heavy vehicle share was displayed against the observed capacity drop per scenario, for the traffic flow composition scenarios. Extra measurements were obtained through new simulations for the 2.5%, 7.5%, 12.5%, 17.5%, 25%, 30% and 35% scenario to increase the number of measurements to identify the trend. The plotted 2^{nd}-degree polynomial shows the trend where the capacity drop increases up to around 15%, whereafter a very wide range of measurements was observed. Even in simulation, it is observed that the capacity drop shows chaotic properties. The upward trend in capacity drop up to 15% HV penetration shows clear connection. However, the measured values afterwards are difficult to capture in a trendline due to the high volatility of capacity drop measurements.

5 CONCLUSIONS

In this chapter, main conclusions from the results of the empirical analysis (5.1) and the simulation study (5.2) will be drawn. Subsequently, the main research questions will be answered (5.3).

5.1 Results from Empirical Analysis

In a German study, the shape parameter A in the Weibull distribution typically ranged between 9 and 15 with an average of 13 for 3-lane on German highways (Brilon et al., 2005). Parameters between 14.7 and 20.9 were observed for 2-lane to 4-lane German highways with heavy vehicle penetrations between 3.1 – 14% (Geistefeldt, 2009). A similar shape parameter A was estimated in this study, ranging between 11.5 and 19.4 (Table 14 - Estimated parameters empirical data of capacity distributions) with heavy vehicle penetrations between 2 – 9% at uncensored events. This shows that the method used in this study obtains similar results, despite changing the aggregation technique and temporal resolution of the detector data. It can be argued that the validity of the method has not changed, but changing the aggregation technique and lowering the temporal resolution represents more realistically the raw measured data.

The observed 2-lane roadway mean and standard deviation in Geistefeldt (2009) ranged between 4914 – 5580 veh/h and 359 – 440 veh/h respectively. The mean and standard deviation of the 2-lane roadways in this study varied between 4531 – 5272 veh/h and 256 – 510 veh/h respectively. The means of the observed capacities in this study are slightly lower as well as the standard deviation, except for the outlier of the A4 (510 veh/h). The reason for the higher standard deviation of the A4 is probably the high share of category 2 vehicles. Whether the slight difference in mean and standard deviation is different due to the country or aggregation technique of the data is difficult to say. Further research should investigate with exactly similar methods as in Geistefeldt (2009) if the same lower mean and standard deviation observation are obtained and identify what causes the difference.

The observed capacity drop in this study ranges from 14 – 29.9%. The average capacity drop in Brilon et al. (2005) was about 1180 veh/h. This is a drop rate of about 18.2% for the average mean breakdown flow of 6467 veh/h with a similar computational methodology. The capacity drops observed for their sites varied widely. All attempts of the researchers to identify regularities within this variation failed. In general, a slightly higher aggregated capacity drop value was observed in this study (22.2%), while the drop percentages varied widely as well. The intra-day analysis showed that this phenomenon not only occurs between sites but also at the same site as the pre queue breakdown intensity varies decently. The observed individual drop rates were often not even 'drop rates' as the pre queue flow breakdown intensity was below the queue discharge flow recovery intensity. Although the pre queue flow breakdown intensity varies widely, the queue discharge flow appeared to be more constant as was observed in the empirical data. This implies that the variance in the day-to-day capacity drop is mainly caused by the difference in breakdown intensity rather than the recovery intensity.

When looking at the median of the yearly estimated capacity distribution functions of the A12, A15a, A15b, and A50 a clear downward trend was observed (Figure 50). During the same period, an increase in absolute heavy vehicles was measured (Figure 51). A similar trend of decreasing capacity was observed in the long-term traffic capacity study in Japan (Shiomi et al., 2019). There, a constant passenger car equivalency factor of 1.7 was used to compute an equivalent traffic flow. When comparing the slope of the decreasing capacity to the slope of the increasing heavy vehicle share in this study, the capacity decreases with a larger factor after multiplication with the PCE of 1.7, or the PCE's computed in Table 15. This implies that the identical

phenomenon as in Japan also occur on Dutch highways. One argument could be that during higher heavy vehicle penetrations, each heavy vehicle would take up more space in terms of passenger car equivalents due to the changed driving behaviour of other road users. As it is plausible that heavy vehicle share also increased on Japanese highways over time, this would at least partly explain the decreasing capacity. The theory of changed vehicle performance due to the increase of driver systems, such as ACC, seems a reasonable explanation and should be alarming to roadway authorities at the very least. However, future research should be conducted on the validity of this conjecture.

The C_V of the estimated capacity distribution functions varied between 0.07 and 0.1 and is less than 10% below the C_V for the PCE value of 1. A similar result was observed in Geistefeldt (2009), where it was claimed that the relatively small fraction of the variance of the capacity distributions was caused by the impact of variable heavy vehicle percentages. The PCU values for the PQF estimated in this study vary between 1.5 and 2.0, excluding the A2, A4, and A15c. The latter two are below 1.0, which is unlikely while the A2 has a value of 1.1. This could be because, on a large 5-lane road, the impact of freight traffic could be less compared to 2-lane roads because overtaking is easier. The computed PCUs are comparable with contemporary PCU factors. Geistefeldt (2009) obtained slightly higher values, ranging between 2.2 and 2.6 for 2-lane roads. It should be noted that only passenger cars and heavy vehicles were distinguished in the mixed traffic flow. It is unclear whether the category 2 vehicles in the study of Geistefeldt (2009) were either classified as category 1 or category 3 vehicles, which could possibly explain the large deviations. Furthermore, the observed PCU for the QDF was generally larger than the PCU of the PQF, which ranged between 1.5 and 2.2 when excluding the A15c and A50. The latter was the only on-ramp with a Lane Change Ban, possibly causing the lower value. A higher PCU at QDF compared to during PQF could indicate that heavy vehicles take up more space during recovery compared to breakdown. However, as the C_V values do differ merely minimal, drawing definitive conclusions seems premature.

The intra-location heavy vehicle share – capacity drop correlation (Figure 55) showed a wide variance of observations. The linear trend line drawn through the measurements showed that the correlation between aggregated heavy vehicle share and capacity drop is nihil (R = 0.1). Removing either site would lead to a slightly different trendline. This implies that the trendline is fairly sensitive to specific locations. One explanation could be the wide variety in geometric, traffic, and control conditions for the selected sites. Selecting equal circumstances was complicated due to the limited number of sites suitable. Adding more measurement locations with low heavy vehicle shares would help clarifying the identified trendline.

The intra-day Heavy Vehicle Share – Capacity Drop correlation (Table 16) showed for almost all sites no clear correlation. The Spearman correlation test was accepted for the A1, A15a, and A15c, all showing a negative correlation, but with high variance. The negative correlation implies that the capacity drop reduces as the share of heavy vehicles increases. This is opposite to the intra-location results. Although the trend is significant, the visualization of e.g. the A1 (Figure 56) shows a bad fit of the trendline and the data. It is especially negative capacity drop observations at higher heavy vehicle share percentages that cause the trendline to decrease. Removing all negative capacity drop values would likely lead to no trend at all.

As shown by the PQF – Capacity drop correlation (Table 17), the breakdown flow likewise decreases as the share of heavy vehicles increases while the recovery flow remains rather stable. For all sites, the Spearman and Pearson tests were accepted, showing a positive correlation. This implies that a relation can be yielded between the breakdown flow and capacity drop. However, the trendline does not have an equal slope for different locations. The A4 has a 1:1 breakdown - capacity drop relation, while the A12 has a 1:2 breakdown – capacity drop relation. The other sites are in between. The difference in in slope appears to be not connected with either the share of heavy vehicles or breakdown speed at either of the locations. Local site characteristics are likely the cause for the different slopes, but further research should be conducted to verify this.

5.2 Results from Simulation

For the increasing heavy vehicle share scenarios, it is expected to get decreased capacity and increased standard deviation. However, the scenario with 15% HV (S4) seems to show a tipping point with a larger standard deviation and capacity value compared to the other scenarios. The recovery flow on the other side seems to show a decreasing decline for increased heavy vehicle penetration. These results showed an increased capacity drop when the truck penetration increases up to 15%, where after it declines. The intra-day capacity drop – HV% correlation (Figure 66) shows that even after adding more truck penetration scenarios, the increasing capacity drop trend holds for increased truck penetration. The plotted 2^{th}-degree polynomial visualized the up-going trend in capacity drop, followed by the down-ward trend after 15%. This can be explained by heterogeneity of traffic flow. At very low heavy vehicle shares, the heavy vehicles do mainly all drive at the right line pre-breakdown which complicates merging from the on-ramp, and the drawback of OTS is that the vehicles that are not able to merge wait on the merging lane. Therefore, the pre queue flow becomes very high, subsequently resulting in a high drop of capacity when a vehicle does merge. In the 20% heavy vehicle share scenario, traffic is more homogenised as trucks drive substantially more on both lanes. Due to the more homogenised flow on both lanes, the heavy vehicles characteristics results in a more equally decreasing speed on both lanes, which is advantageous for the merging traffic. The simulated results showed similar capacity drop variance corresponding with the variance of the empirical results in the Heavy Vehicle Share – Capacity Drop measurements (Figure 55).

The adapted scenarios gave insight into how future physical and operational characteristics might impact traffic flow. The decreased physical size of heavy vehicles resulted in a slight increase in capacity and a similar recovery distribution function. As the size of the heavy vehicles was decreased, more vehicles should theoretically fit on the road. This scenario would be beneficial for highway capacity. The increased acceleration capabilities will lead to a slightly decreased capacity, but a substantially higher recovery distribution function. As the acceleration increases, the recovery flow increases as people can accelerate faster up to desired speed out of congestion.

The combined scenarios showed an equal capacity, but with an increased recovery flow. Therefore, the capacity drops observed were lower compared to the base scenarios. The standard deviation of the breakdown flows decreased compared to the base scenarios, which can be declared by slightly equalizing the mixed traffic in terms of physical characteristics. The standard deviation of the recovery flow decreased as the share of heavy vehicles increases in the combined scenarios, while the standard deviation of the recovery flow oppositely increased as the share of heavy vehicles increases in the base scenarios. The cause is likely the increased acceleration capabilities of all vehicles. Percentage-wise, this was increased more for heavy vehicles. It implies that increased acceleration capabilities of all vehicles cause a more wide spread recovery flow probability at low heavy vehicle percentages.

5.3 Research Questions

Three main research questions were established in 1.2. The research questions were answered by executing the activities of the research stages within the Research Framework (Figure 7). The sub-question (Table 1) helped in steering the activities and were answered throughout the corresponding chapters. The main research questions will be answered below.

Research Question 1: Which variables influence the capacity drop rate on both strategic and operational levels?

State-of-the-art literature showed that on an empirical basis the number of lanes, speed in congestion, on-ramp flow, number of lane changes, type of congestion, type of day, and precipitation are of influence on the drop rate. Modelling studies observed that void creation, reaction time, the share of heavy vehicles, inter-driver/vehicle spread, intra-driver variation, spatial lane change location, and acceleration stochasticity are also decisive for the drop rate.

On a strategic level, an increase in heavy vehicles is projected in the future due to growing economic welfare. Measures taken by governments will decide the pace at which commercial traffic is made hybrid or electric. New emerging technologies such as Adaptive Cruise Control of Lane Change Assistant are expected to saturate the market. Changed physical and operational characteristics of heavy vehicles are expected, but as of yet unclear in terms of length, speed, or acceleration capabilities. On an operational level, heavy vehicles have both longitudinal and lateral effects on traffic flow, but the lack of data symbolizes the need for more microscopic data.

Available technological advancements, either in-vehicle or from an infrastructure perspective, are already influencing traffic flow. It is conjectured that cars equipped with ACC deteriorate the road capacity. A fully-autonomous vehicle fleet is expected to optimize traffic flow, but the mixed-traffic in the coming decades will likely have a negative impact. However, several infrastructural measures were proven to improve the effect of heavy vehicles on traffic flow. Truck restrictions, on-ramp meters, and variable speed limits showed better flow-lane utilization, outflow, or even prevention of traffic breakdown.

Research Question 2: How and to what extend does freight traffic impact the capacity drop on highways currently?

The Kaplan-Meier Product Limit Method was used to estimate the capacity and recovery distributions of the sites for extended time periods to overcome the stochastic characteristics of traffic flow. The observed capacity drop ranged between 14 – 29.9%, while the share of heavy vehicles at breakdown ranged between 2 – 9%. The empirical results seem to show a connection between the heavy vehicle share and the capacity drop, but this could not be statistically proven by both the Pearson and Spearman correlation test both intra-location and intra-day. Therefore, it can be concluded the chaotic properties of the breakdown flow seem to superimpose the effects the share of heavy vehicles has on the capacity drop from an empirical perspective.

The variance of the observed capacity drops can mainly be attributed to the high variance of the breakdown flow. Empirical results also show that the pre queue flow and the capacity drop are greatly dependent. This implies that when an breakdown flow has been observed, it is possible to predict the capacity drop in practice. Roadway operators could elaborate on this method to forecast the extent to which traffic flow should be reduced via for example speed reduction by matrix signs to increase the probability of traffic recovery.

Empirical data revealed that the number of heavy vehicles increased on the A12, A15 and A50 by about 10 to 14 veh/h per year during breakdown. Moreover, a decreasing capacity was observed at those sites of about 28 to 90 veh/h per year. The capacity decreased more than the increase in heavy vehicles, even after correction with the estimated PCE values. This raises concern about the underlying cause of the depreciation of roadway capacity. It is conjectured that new innovations such as ACC are responsible. However, this has not been verified. As the cause of this phenomenon has not been discovered yet, researchers should investigate this phenomenon because it could have a large impact on the Dutch roadway capacity.

Research Question 3: How do different vehicle heterogeneity compositions and, physical- and operational characteristics of freight traffic impact the capacity drop?

The simulation study aimed to fill the gap of the empirical analysis by equalizing external factors. The results showed that the estimated capacity drop tended to increase with an increasing amount of heavy vehicle share up to 15%. This reinforces the empirical evidence that a connection between heavy vehicle share and capacity drop exists. As the heavy vehicle share during breakdown only reaches up to 9% on roads currently, it is possible that the effect becomes more influential as the share of heavy vehicles grows further on the Dutch highways. Current empirical data does not show significant correlations, insights from the simulations show that this could be the case in the future as the percentage of heavy vehicle share increases continuously.

Future scenarios, with increased acceleration capabilities of vehicles and decreased truck sizes showed a decrease in the capacity drop. However, it also showed an increased standard deviation of the recovery distributions for higher heavy vehicles shares compared to the base scenarios. This shows that the estimation of the recovery flow shall become more difficult in the future.

For every 5% increase in heavy vehicle share, the capacity drop was estimated up to 40%. After the increase of capacity drop up to 15%, highly volatile but declining capacity drop measurements were observed. Although such scenarios were not observed empirically in this study, it is known that certain locations at the A15 close to the port of Rotterdam do have very high heavy vehicle shares. The simulation results show that the capacity drop under high heavy vehicle percentages have very chaotic properties. This is in accompanied by high standard deviations of both the breakdown and recovery distributions, which in general complicate traffic flow estimations. These simulations are useful in investigating how a future increase of traffic at truck-dominant high way sections would impact the breakdown and recovery distribution, and subsequent capacity drop.

6 DISCUSSION

In the discussion, several steps of the method will be discussed as well as the results obtained. First, it is important to put the results in perspective regarding the research methodology (6.1). The methodology of the Flow Type Classification Model, Data Collection, Simulation Scenarios, and Statistical Tests will be discussed. Subsequently, recommendations for future research will be given (6.2).

6.1 Reflection of the Research Methodology

6.1.1 Flow Type Classification Model Reflection

The Flow Type Classification model has a slightly different accuracy compared to similar studies with different temporal resolutions. Due to the 3-minute smoothing technique applied to the data, the breakdown and recovery events approach more real values compared to a 5-minute time window. However, this also leads to higher variability in the data, which is less flattened out. Especially at speeds around 60 – 70 km/h, the speed tends to fluctuate decently, leading to problems with incorrectly detecting events, such as breakdowns. An extra requirement (Eq. 14) was added to prevent small fluctuations to be classified as uncensored falsely by looking at the speed difference over a longer period. However, measurements were still classified as uncensored falsely after manual revise. It is not possible to give an estimation of the number of falsely classified measurements, since each dataset is too large to perform such a check. The set of criteria was thoroughly tested at several sites for several days almost without any false classifications. By selecting a larger time-window, e.g. 5-minute temporal resolution, the number of falsely classified measurements reduces while the correctly classified uncensored flow values will decrease for the breakdown event due to flattening of the peaks.

It is fair to say that the current flow type detection method has a one-size-fits-all approach in detecting breakdowns and recoveries. As can be observed in Appendix C, which contains an overview of classified measurements at each site, different breakdown and recovery patterns can be observed by the highlighted Gaussian kernel density estimation. For example, the A1 has the maximum observed right before the breakdown at 80 km/h, while this closer to 70 km/h for the A12 and 90 km/h for the A58. The speed threshold for the detection of a breakdown and speed threshold for recovery were now set equal for each site. Future research should focus on designing a model that has a less one-size-fits-all approach and could for example use machine-learning techniques to optimize the breakdown and recovery threshold. Besides, such a technique could be useful in removing falsely classified events, such as shockwaves as either breakdown or recovery events.

As discussed in 6.1.1, the bottleneck detector is not equidistant from the bottleneck at each site. To overcome the risk of comparing intensities located differently relative to the bottleneck, maximum values in a short time window (4 minutes before the classified uncensored event) were classified as the uncensored event.

Due to the set of breakdown and recovery criteria, multiple breakdown or recoveries might have been classified at a single day. This increases the number of classified measurements. However, the model does not recognize dependence between individual uncensored breakdown and recoveries. This implies that a breakdown could be classified without a subsequent recovery. Therefore, the intra-day statistical tests were applied only on breakdown-recovery couples by removing the duplicate classifications on the same day. Executing both tests next to each other guarantees the credibility of the outcome.

6.1.2 Data collection reflection

In comparable studies (Brilon et al., 2005; Geistefeldt, 2009; Regler, 2004), a different approach to reduce the volatility in flow and speed measurements were applied. In those studies, the 5-minute aggregation technique was applied, while in this study the 3-minute smoothening was used. The difference with the lower temporal resolution is the close match with the raw data. The smoothening technique enables more data points to be selected. Getting the maximum flow does not depend on whether this flow is close to the aggregation resolution anymore. A detailed explanation can be found in Appendix A. However, some difficulties arise in the Flow Type Classification Model at the expense of choosing the lower temporal resolution, such as the recognition of stop-and-go-waves, as discussed in 6.1.1.

During the data collection and preparation, decisions have been made to simplify the data processing in the Flow Type Classification Model. The downstream detector does only have to measure speed to rule out spillback from other detectors. Therefore, any type of loop detector would satisfy. However, it was decided for the ease of analysis to only use detectors that measure different vehicle categories. This prevents errors by combining different indexed datasets. This requirement substantially limited the site location options. Because of this, some locations do have the bottleneck detector right at the ramp, while others do have this detector slightly more downstream. Besides, as only a few locations fulfilled the requirements, especially of having multiple detectors around bottlenecks able to measure category 3 vehicles. It can be argued that these specific detectors were located around these bottlenecks for a reason, e.g. to monitor the flow of heavy vehicles near these ramps. Future research could implement any detector as the downstream detector, to increase the number of site locations. Initially, the idea was to compare sites with a different number of lanes with regards to the capacity drop. However, it turned out that mainly 2-lane sites were suited. The selection of the sites was done manually by looking up well-known bottlenecks. Including any type of downstream detector would likely lead to multiple suited 3-lane and 4-lane sites.

Furthermore, less extensive removal of measurement data compared to similar studies was decided. The removal of precipitation days was not executed, while the precipitation data is available by TNO. It was decided not to remove days with precipitation to put the effort and time in the correct operation of the Flow Type Classification Model. Furthermore, the weekend days and holidays have not been removed for a similar reason. The consequence of not excluding the measurements under irregular circumstances is that measurements at the lower range are not removed. A lower capacity was observed on weekend days (Calvert et al., 2016) due to less experienced drivers within the driver population. Also, at days with precipitation (Calvert & Snelder, 2013) as people tend to increase following distance due to increased braking distance and reduced road vision.

6.1.3 Simulation Model reflection

The first point of discussion is the error within OpenTrafficSim, which was causing the loop detector data to be incorrect. Although at first the results seemed somewhat plausible, they did not follow the fundamental relationship of traffic flow theory. After several revisions of the process from the OpenTrafficSim output data to the outcome of the computational framework, it was established that the error was caused outside of the processing of the Flow Type Classification Model. The errors appeared to be within the loop detector recognition of vehicles, which recognized identical vehicles several times in a very short time window. One of the developers of OpenTrafficSim, was consulted and a workaround was established. An algorithm was added to remove multiple measured vehicles from the output data and the results finally matched the visualized simulation.

In the example model implemented, the flow of traffic consists only of two vehicle classes: passenger cars and heavy vehicles. However, the category 2 vehicle class is not implemented while the empirical data of the A4 shows that category 2 vehicles do have a substantially different impact on traffic compared to passenger cars or heavy vehicles. A more realistic simulation environment would also include category 2 vehicles, which would also allow for a direct comparison between empirical and simulation data. However, this implies that the model should be recalibrated, which was out of the scope of this research.

During the investigation of the simulation in the visualized environment, it appeared that the lane utilization during congestion during high heavy vehicle penetration was unrealistic. Due to the high value of the VGAIN parameter of heavy vehicles, the heavy vehicles kept on driving on the right lane. The value helps determining by which speed difference a lane change should be initiated. During a breakdown at the bottleneck, the right lane was fully congested while the passenger cars were still at free flow at the left lane. After lowering the VGAIN parameter of heavy vehicles, both lanes were better utilized during congestion. However, it should be noted that the model has not been recalibrated with the parameter change and the parameter was changed rather instinctively. The performance of the model regarding lane flow distribution for different levels of density is not as clear during congested conditions as it depends on the stochastic input (Schakel et al., 2012). One run did seem to perform better in creating a realistic breakdown compared to the other.

The future scenarios of adapted physical and operational characteristics are merely an investigation on how a change of either characteristic would change the traffic flow and subsequent capacity drop. No profound study has investigated how future characteristics will change over time based on the expected technological advancements. Future studies should focus on mapping the exact change of characteristics such that future simulation studies can use those estimations. Currently, the adapted scenarios are rather straight forward as only the length and following acceleration capability have been changed. Therefore, the results are of limited use and should only be compared to the base scenarios simulated and do not aim to predict future breakdown or recovery values.

The configuration correlation was kept at 0.4 in this study. Changing this to a higher value could represent more H-H cases. This would be an easy way to model assistant driving systems, such as (C)ACC. Future research could increase the configuration correlation and investigate the impact.

6.1.4 Statistical Tests reflection

Two methods to investigate the correlation between the capacity drop and heavy vehicle share were used in this study: intra-location and intra-day measurements. The first method uses all classified measurements to estimate the probability distribution functions. The capacity drop was computed by subtracting the median of the breakdown curve from the median of the recovery curve. The second method only uses breakdown-recovery dependent events by removing all other classified uncensored events. This implies that only individual related events were used to compute the capacity drop. Therefore, on both a large temporal aggregated scale and a small temporal individual scale the capacity drop – heavy vehicle share correlation has been tested. For both variables, the relative value (percentage) has been computed for the correlation tests rather than the absolute value in veh/h. This is a more intuitive representation of the road circumstances as the capacity drop and the heavy vehicle share are different per site.

6.2 Recommendations for Future Research

Improvement of the Flow Type Classification Model is suggested such that the model can use any downstream detector to rule out spillback. This would increase the sites suitable for selection and could subsequently add 3 or 4-lanes sites to the selection. Increasing the site selection would enable the comparison of sites with more similar geometric, traffic, and control characteristics. Besides, the current breakdown and recovery threshold constants should be replaced with site-specific thresholds with the help of, for example, machine learning techniques.

The exclusion of external factors could be improved by removing precipitation days, weekends, and holidays. Combining these changes with a large set of locations would allow to test correlations described in 2.1 on Dutch highways. Factors such as speed in congestion, on-ramp flow, number of lane changes, and type of congestion could also be gathered and processed in the Flow Type Classification Model. The magnitude of the factors could be compared. Increasing the number of variables gathered during traffic breakdowns would possibly expose relations which declare the chaotic properties of the capacity drop which are difficult to expose when only considering a single variable.

Another interesting observation was the decreasing capacity combined with the increasing heavy vehicle share on particular highways. This observation could be tested on a large scale by investigating how this conjecture holds by adding more road sections to the analysis. The current conjecture that driver assistance systems, such as Adaptive Cruise Control, cause the drop should be investigated in detail. A microscopic analysis would be insightful as these factors can be studied in detail rather than from loop detectors.

The Passenger Car Equivalency values should be computed yearly for sites A12, A15, and A50. These sites show a deteriorating capacity with increased heavy vehicles, it could be that the impact of heavy vehicles becomes larger in terms of passenger car equivalents from a spatial perspective. Increased heavy vehicles share might cause changed driving behaviour of other road users.

The collection of microscopic data should have priority for researching institutes and roadway authorities. Building a database identical to the macroscopic loop detector data would help to monitor the impact on traffic flow of traffic with mixed driver assistance systems. Besides, the data would be useful in improving simulation models so that studies similar to this study would benefit from an even more realistic flow of traffic.

Investigating how physical and operational characteristics will change in the coming years will help improving future scenarios of simulation models. As of yet, it is hardly researched how characteristics of heavy vehicles will change in the coming years.

REFERENCES

Aghabayk, K., Sarvi, M., & Young, W. (2012). Understanding the Dynamics of Heavy Vehicle Interactions in Car-Following. *Journal of Transportation Engineering, 138*(12), 1468-1475. doi:10.1061/(ASCE)TE.1943-5436.0000463

ANWB. (2018, 27-12-2018). ANWB: 20 procent meer files op Nederlandse wegen. Retrieved from https://www.anwb.nl/verkeer/nieuws/nederland/2018/december/anwb-20-procent-meer-files-op-nederlandse-wegen

Banks, J. H. (1991). Two-Capacity Phenomenon at Freeway Bottlenecks: A Basis for Ramp Metering? *Transportation Research, 1320*, 83-90.

Banks, J. H. (2006). *New Approach to Bottleneck Capacity Analysis: Final Report*. Retrieved from University of California, Berkeley: https://merritt.cdlib.org/d/ark:%2F13030%2Fm5t43v6c/2/producer%2FPRR-2006-13.pdf

Bertini, R., & Malik, S. (2004). Observed Dynamic Traffic Features on Freeway Section with Merges and Diverges. *1867*, 25-35. doi:10.3141/1867-04

Brilon, W., Geistefeldt, J., & Regler, M. (2005). Reliability of freeway traffic flow: A stochastic concept of capacity. *Proceedings of the 16th International Symposium on Transportation and Traffic Theory*, 125-144.

Brilon, W., & Zurlinden, H. (2003). *Overload probabilities and traffic activity as design criteria for road traffic systems.* (Lehrstuhl fuer Verkerhswesen). FIZ - Fachinformationszentrum Karlsruhe,

Broekman, E. G. (2017). *Implementation and Calibration of Traffic Behaviour Models Around Ramps on Dutch Motorways.* (Master of Science). TU Delft, TU Delft Repository.

Calvert, S. C., Schakel, W. J., & van Lint, J. W. C. (2017). Will Automated Vehicles Negatively Impact Traffic Flow? *Journal of Advanced Transportation, 2017*, 3082781. doi:10.1155/2017/3082781

Calvert, S. C., & Snelder, M. (2013, 6-9 Oct. 2013). *Influence of rain on motorway road capacity - A data-driven analysis.* Paper presented at the 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013).

Calvert, S. C., Taale, H., & Hoogendoorn, S. P. (2016). Quantification of motorway capacity variation: influence of day type specific variation and capacity drop. *Journal of Advanced Transportation, 50*(4), 570-588. doi:10.1002/atr.1361

Calvert, S. C., Van Wageningen-Kessels, F. L. M., & Hoogendoorn, S. P. (2018). Capacity drop through reaction times in heterogeneous traffic. *Journal of Traffic and Transportation Engineering (English Edition), 5*(2), 96-104. doi:10.1016/j.jtte.2017.07.008

Cassidy, M. J., & Bertini, R. L. (1999). Some traffic features at freeway bottlenecks. *Transportation Research Part B: Methodological, 33*(1), 25-42. doi:10.1016/s0191-2615(98)00023-x

Cassidy, M. J., & Rudjanakanoknad, J. (2005). Increasing the capacity of an isolated merge by metering its on-ramp. *Transportation Research Part B: Methodological, 39*(10), 896-913. doi:10.1016/j.trb.2004.12.001

Chen, D., & Ahn, S. (2018). Capacity-drop at extended bottlenecks: Merge, diverge, and weave. *Transportation Research Part B: Methodological, 108*, 1-20. doi:10.1016/j.trb.2017.12.006

Chen, D., Ahn, S., Bang, S., & Noyce, D. (2016). Car-Following and Lane-Changing Behavior Involving Heavy Vehicles. *2561*, 89-97. doi:10.3141/2561-11

Cho Hyun, W., & Laval Jorge, A. (2020). Combined Ramp-Metering and Variable Speed Limit System for Capacity Drop Control at Merge Bottlenecks. *Journal of Transportation Engineering, Part A: Systems, 146*(6), 04020033. doi:10.1061/JTEPBS.0000350

Chung, K., Rudjanakanoknad, J., & Cassidy, M. J. (2007). Relation between traffic density and capacity drop at three freeway bottlenecks. *41*(1), 82-95. doi:10.1016/j.trb.2006.02.011

Coifman, B., & Li, L. (2017). A Critical Evaluation of the Next Generation Simulation (NGSIM) Vehicl Tracjectory Dataset. *Transportation Research Part B: Methodological, 105*, 362-377. Retrieved from http://www2.ece.ohio-state.edu/~coifman/documents/#DataSets

Daamen, W. (2010). *Empirical Analysis Reader*(pp. 214).

Daganzo, C. F. (2002). A behavioral theory of multi-lane traffic flow. Part I: Long homogeneous freeway sections. *Transportation Research Part B: Methodological, 36*(2), 131-158. doi:10.1016/s0191-2615(00)00042-4

de Bok, M., Wesseling, B., Kiel, J., Miete, O., & Jan, F. (2018). A sensitivity analysis of freight transport forecasts for The Netherlands. *Rivista Internazionale di Economia dei Transporti / International Journal of Transport Economics, 45*.

Duret, A., Ahn, S., & Buisson, C. (2012). Lane flow distribution on a three-lane freeway: General features and the effects of traffic controls. *Transportation Research Part C: Emerging Technologies, 24*, 157-167. doi:10.1016/j.trc.2012.02.009

Elefteriadou, L., Roess, R. P., & Williams, M. R. (1995). Probabilistic Nature of Breakdown at Freeway Merge Junctions. *Transportation Research, 1484*.

Geistefeldt, J. (2009). Estimation of Passenger Car Equivalents Based on Capacity Variability. *2130*, 1-6. doi:10.3141/2130-01

Geistefeldt, J., & Brilon, W. (2009). A Comparative Assessment of Stochastic Capacity Estimation Methods. In (pp. 583-602).

Hall, F. L., & Agyemang-Duah, K. (1991). Freeway Capacity Drop and the Definition of Capacity. *Transportation Research, 1320*.

Knoop, V., Hoogendoorn, S., Shiomi, Y., & Buisson, C. (2012). Quantifying the Number of Lane Changes in Traffic. *2278*, 31-41. doi:10.3141/2278-04

Knoope, M., & Francke, J. (2019). *Trendprognose wegverkeer 2019-2024 voor RWS*. Ministerie van Infrastructuur en Waterstaat

Lamberts, F. (2019, 23-10-2019). Tesla Semi electric truck production to start 'with limited volumnes' in 2020. *Electreck*. Retrieved from https://electrek.co/2019/10/23/tesla-semi-electric-truck-production-limited-volume-2020/

Laval, J. A. (2009). Effects of geometric design on freeway capacity: Impacts of truck lane restrictions. *43*(6), 720-728. doi:10.1016/j.trb.2009.01.003

Laval, J. A., & Daganzo, C. F. (2006). Lane-changing in traffic streams. *Transportation Research Part B: Methodological, 40*(3), 251-264. doi:10.1016/j.trb.2005.04.003

Leclercq, L., Knoop, V. L., Marczak, F., & Hoogendoorn, S. P. (2016). Capacity drops at merges: New analytical investigations. *62*, 171-181. doi:10.1016/j.trc.2015.06.025

Leclercq, L., Laval, J. A., & Chiabaut, N. (2011). Capacity Drops at Merges: an endogenous model. *Procedia - Social and Behavioral Sciences, 17*, 12-26. doi:10.1016/j.sbspro.2011.04.505

Leclercq, L., Marczak, F., Knoop, V. L., & Hoogendoorn, S. P. (2016). Capacity Drops at Merges. *Transportation Research Record: Journal of the Transportation Research Board, 2560*, 1-9. doi:10.3141/2560-01

Maciejewski, M. (2010). A comparison of microscopic traffic flow simulation systems for an urban area. *Transport Problems : an International Scientific Journal, 5*.

Minderhoud, M. M., Botma, H., & Bovy, P. H. L. (1997). Assessment of roadway capacity estimation methods. Highway capacity issues and analysis. *Transportation Research Record*(1572), 59-67.

Moridpour, S., Mazloumi, E., & Mesbah, M. (2015). Impact of heavy vehicles on surrounding traffic characteristics. *49*(4), 535-552. doi:10.1002/atr.1286

NDW. (2010). *NDW interface beschrijving*. Retrieved from

Oh, S., & Yeo, H. (2012). Estimation of Capacity Drop in Highway Merging Sections. *2286*, 111-121. doi:10.3141/2286-13

Oh, S., & Yeo, H. (2015). Impact of stop-and-go waves and lane changes on discharge rate in recovery flow. *Transportation Research Part B: Methodological, 77*, 88-102. doi:10.1016/j.trb.2015.03.017

Ossen, S., & Hoogendoorn, S. P. (2011). Heterogeneity in car-following behavior: Theory and empirics. *Transportation Research Part C: Emerging Technologies, 19*(2), 182-195. doi:10.1016/j.trc.2010.05.006

Raj, P., Sivagnanasundaram, K., Asaithambi, G., & Ravi Shankar, A. U. (2019). Review of Methods for Estimation of Passenger Car Unit Values of Vehicles. *Journal of Transportation Engineering, Part A: Systems, 145*(6), 04019019. doi:10.1061/jtepbs.0000234

Regler, M. (2004). *Verkehrsablauf und Kapazität auf Autobahnen (Freeway Traffic Flow and Capacity)*. Retrieved from Bochum:
Rijkswaterstaat. (2018). Smart Mobility - Numbers Dutch Fleet 2016.
Rijkswaterstaat. (2020a). *Maximum speed in the Netherlands*. Retrieved from https://www.rijkswaterstaat.nl/wegen/wetten-regels-en-vergunningen/verkeerswetten/maximumsnelheid/index.aspx#:~:text=Op%20onze%20snelwegen%20geldt%20tussen,of%20100%20km%2Fh).
Rijkswaterstaat. (2020b). *Overtaking ban heavy vehicles*
Rotterdam, C. o. (2019). *Moving towards Zero Emission City Logistics (ZECL) in Rotterdam in 2025*.
SAE. (2018). SAE J3016 Levels of Driving Automation. Retrieved from https://www.sae.org/news/press-room/2018/12/sae-international-releases-updated-visual-chart-for-its-%E2%80%9Clevels-of-driving-automation%E2%80%9D-standard-for-self-driving-vehicles
Sarvi, M. (2011). Heavy commercial vehicles-following behavior and interactions with different vehicle classes. n/a-n/a. doi:10.1002/atr.182
Schakel, W. J., Gorter, C. M., De Winter, J. C. F., & Van Arem, B. (2017). Driving Characteristics and Adaptive Cruise Control ? A Naturalistic Driving Study. *IEEE Intelligent Transportation Systems Magazine, 9*(2), 17-24. doi:10.1109/mits.2017.2666582
Schakel, W. J., Knoop, V. L., & Van Arem, B. (2012). Integrated Lane Change Model with Relaxation and Synchronization. *Transportation Research Record: Journal of the Transportation Research Board, 2316*(1), 47-57. doi:10.3141/2316-06
Shiomi, Y., Xing, J., Kai, H., & Katayama, T. (2019). Analysis of the Long-Term Variations in Traffic Capacity at Freeway Bottleneck. *Transportation Research Record: Journal of the Transportation Research Board*, 036119811984212. doi:10.1177/0361198119842124
Shojaat, S., Geistefeldt, J., Parr, S., Escobar, L., & Wolshon, B. (2018). *Defining Freeway Design Capacity Based on Stochastic Observations*.
Srivastava, A., & Geroliminis, N. (2013). Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. *30*, 161-177. doi:10.1016/j.trc.2013.02.006
Treiber, Hennecke, A., & Helbing, D. (2000). Congested traffic states in empirical observations and microscopic simulations *Physical Review E, 62*, 1803-1824.
Treiber, M., & Helbing, D. (2002). An Adaptive Smoothing Method for Traffic State Identification from Incomplete Information. *32*. doi:10.1007/978-3-662-07969-0_33
Tsugawa, S., Jeschke, S., & Shladover, S. E. (2016). A Review of Truck Platooning Projects for Energy Savings. *1*(1), 68-77. doi:10.1109/tiv.2016.2577499
Van Arem, B., Van Driel, C. J. G., & Visser, R. (2006). The Impact of Cooperative Adaptive Cruise Control on Traffic-Flow Characteristics. *IEEE Transactions on Intelligent Transportation Systems, 7*(4), 429-436. doi:10.1109/tits.2006.884615
Van Beinum, A., Broekman, E., Farah, H., Schakel, W., Wegman, F., & Hoogendoorn, S. (2020). Critical Assessment of Microscopic Simulation Models for Simulating Turbulence around Motorway Ramps. *Journal of Transportation Engineering, Part A: Systems, 146*(2), 04019066. doi:10.1061/jtepbs.0000296
Van Beinum, A., Farah, H., Wegman, F., & Hoogendoorn, S. (2018). Driving behaviour at motorway ramps and weaving segments based on empirical trajectory data. *Transportation Research Part C: Emerging Technologies, 92*, 426-441. doi:10.1016/j.trc.2018.05.018
Vander Werf, J., Shladover, S. E., Miller, M. A., & Kourjanskaia, N. (2002). Effects of Adaptive Cruise Control Systems on Highway Traffic Flow Capacity. *Transportation Research Record: Journal of the Transportation Research Board, 1800*(1), 78-84. doi:10.3141/1800-10
Verkeersprestaties motorvoertuigen; kilometers, voertuigsoort, grondgebied. (2019). Retrieved from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/80302ned/line?ts=1586512647302. Retrieved 10-04-2020, from CBS https://opendata.cbs.nl/statline/#/CBS/nl/dataset/80302ned/line?ts=1586512647302
Verkeersprestaties vrachtvoertuigen naar gewicht en grondgebied 2001-2017. (2018). Retrieved from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/80379NED/line?ts=1586512187007&fromstatweb=true. Retrieved 10-4-2020, from CBS https://opendata.cbs.nl/statline/#/CBS/nl/dataset/80379NED/line?ts=1586512187007&fromstatweb=true
Wegenwiki. (2019). Infrastructuurfonds. Retrieved from https://www.wegenwiki.nl/Infrastructuurfonds

Wiedemann, R. (1991). Modeling of RTI-Elements on Multi lane Roads. *Advanced Telematics in Road Transport by the Commission of the European Community, DG XIII*.

Yang, G., Xu, H., Wang, Z., & Tian, Z. (2016). Truck acceleration behavior study and acceleration lane length recommendations for metered on-ramps. *International Journal of Transportation Science and Technology, 5*(2), 93-102. doi:10.1016/j.ijtst.2016.09.006

Yuan, Knoop, V. L., & Hoogendoorn, S. P. (2015). Capacity Drop. *2491*, 72-80. doi:10.3141/2491-08

Yuan, K. (2016). *Capacity Drop on Freeways: Traffic Dynamics, Theory and Modeling.* (PhD doctoral thesis). TU Delft,

Yuan, K., Laval, J., Knoop, V. L., Jiang, R., & Hoogendoorn, S. P. (2019). A geometric Brownian motion car-following model: towards a better understanding of capacity drop. *Transportmetrica B: Transport Dynamics, 7*(1), 915-927. doi:10.1080/21680566.2018.1518169

Zhang, L., & Levinson, D. (2004). Some Properties of Flows at Freeway Bottlenecks. *1883*, 122-131. doi:10.3141/1883-14

Zhang, Y., & Ioannou, P. A. (2015, 15-18 Sept. 2015). *Combined Variable Speed Limit and Lane Change Control for Truck-Dominant Highway Segment.* Paper presented at the 2015 IEEE 18th International Conference on Intelligent Transportation Systems.

Zhou, J., Rilett, L., & Jones, E. (2019). Sensitivity Analysis of Speed Limit, Truck Lane Restrictions, and Data Aggregation Level on the HCM-6 Passenger Car Equivalent Estimation Methodology for Western U.S. Conditions. *Transportation Research Record: Journal of the Transportation Research Board*, 036119811985145. doi:10.1177/0361198119851451

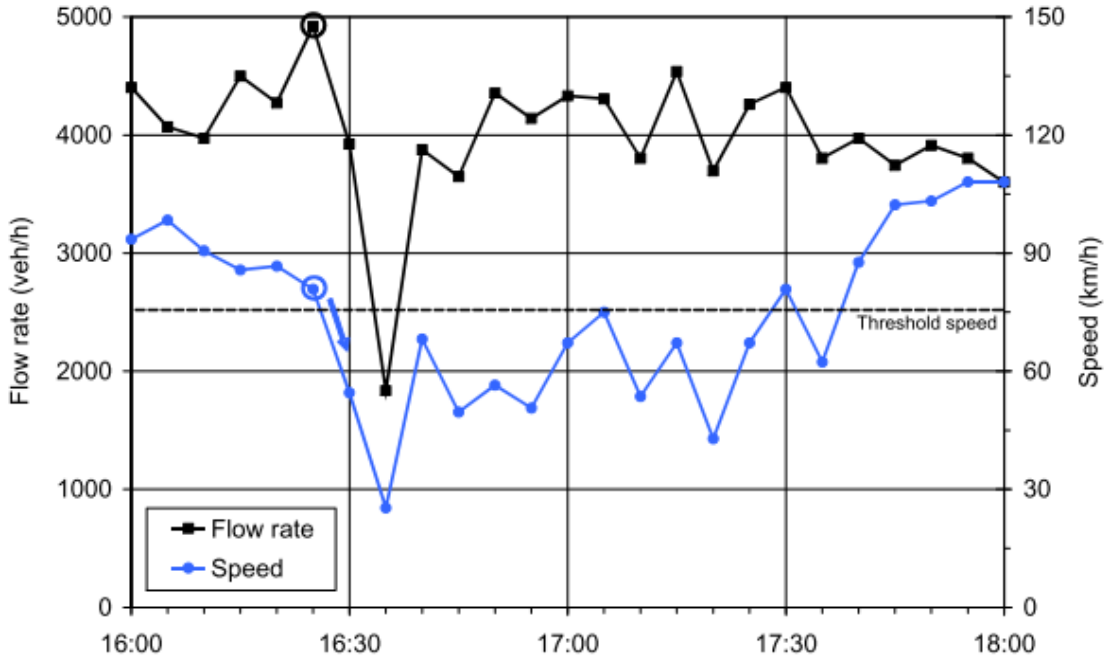
APPENDICES

A Different smoothing methodologies compared
B Explanation of selected locations
C Overview of classified measurements of each site
D Flow, speed and vehicle composition boxplots of sites
E Between-year results
F Technological advancements traffic flow
G Confidence interval of distributions

A Different smoothening methodologies compared

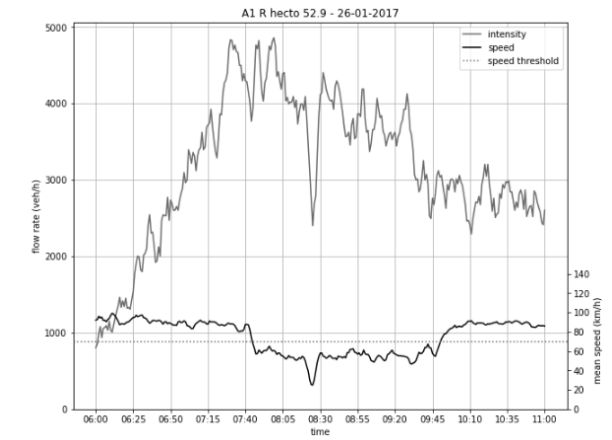
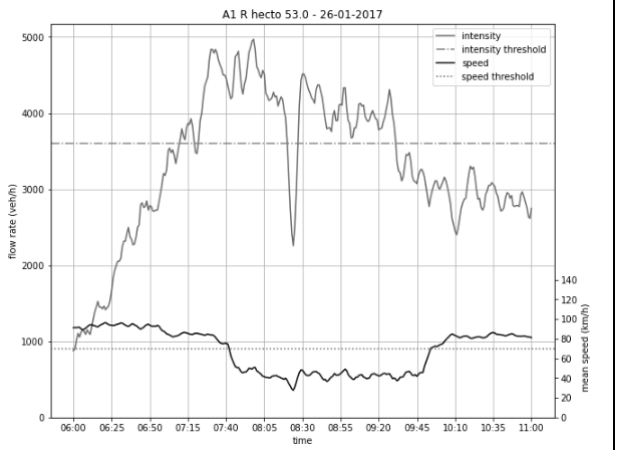
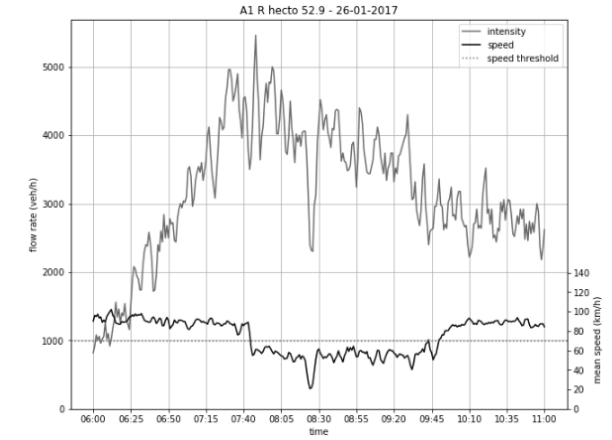
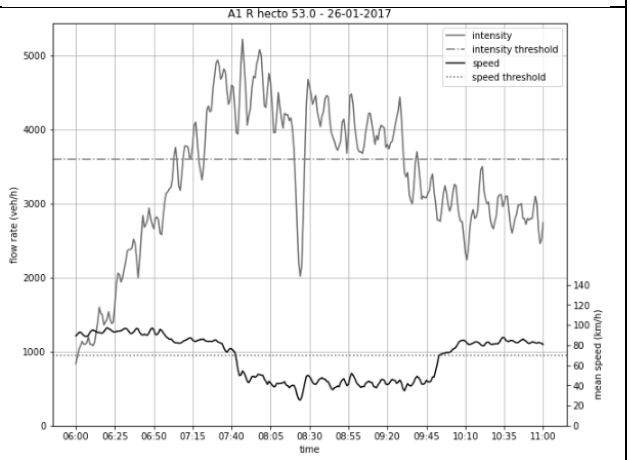
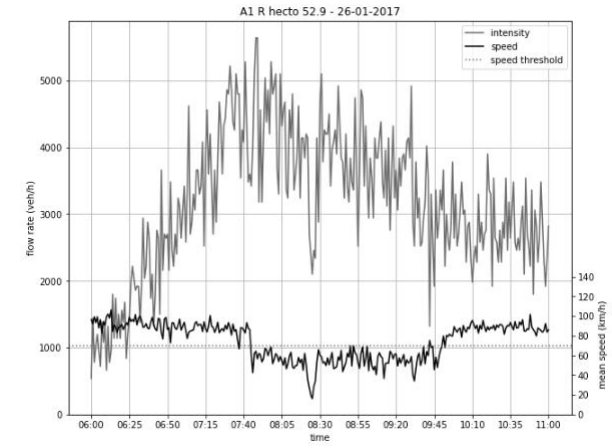
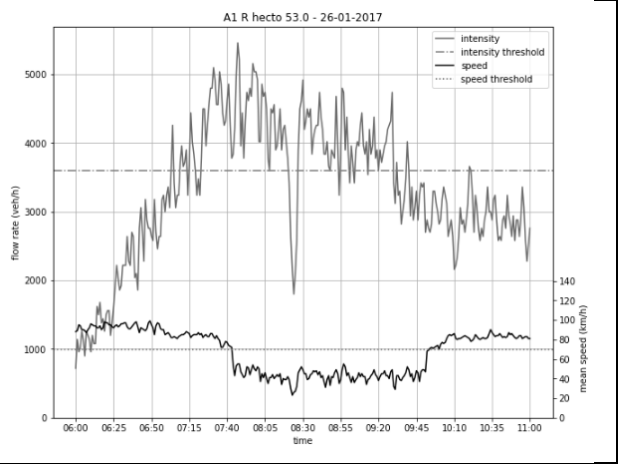
First, the difference between aggregation and smoothening should be clear. Aggregation is the formation of different measurements into a cluster, depending on the time window. This technique, a 5-minute aggregation, was applied to prepare the data for stochastic capacity analysis in Brilon et al. (2005) and Geistefeldt (2009) after experimenting with several time intervals. To visualize this, this technique leads to a total of 25 data points of flow and speed within a 2-hour analysis period, see figure below from Geistefeldt and Brilon (2009).

The main downside of this method is that measurements get aggregated on a 5-minute time scale, while traffic measurements show great variability from minute to minute. Significant speed drops can be observed between two single minutes, which might lead to bad aggregation in case the speed drop happens between the aggregation of two 5-minute time windows. Besides, the measured peaks get flattened with the 5-minute aggregation, which is not preferred when analysing maximum flows that lead to a breakdown.



Therefore a moving-average is a good solution. It allows for minute-accurate breakdown detection which is preferred due to the high variability of traffic from minute to minute. However, an appropriate moving average should be selected with the trade-off of less variability, which eases breakdown detection and disregarded shockwaves, but with flattening peaks or vice versa. The following three options have been investigated, for both the raw and filtered data: no moving average, 3-minute moving average, and 5-minute moving average. Remember that the filtered data has already been kind of smoothened.

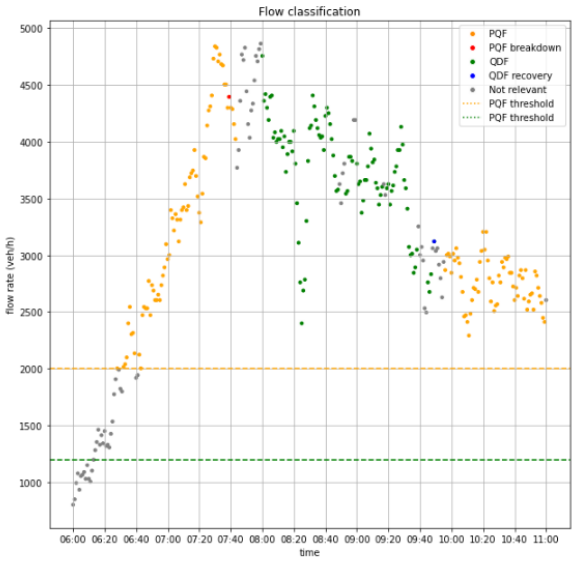
The bottleneck at the A1 near Barneveld will be used as an example to show how different moving averages impact the flow and speed measurements. Remark that the raw data set uses the exact location of the bottleneck selected, which is in this case at hector 52.9 while the filtered data has extrapolated hectometres every 200 meters. Hence hectometre 53.0 is used in the filtered data. The raw and filtered data for different moving average are shown in the following figure below for a single day (26-01-2017):

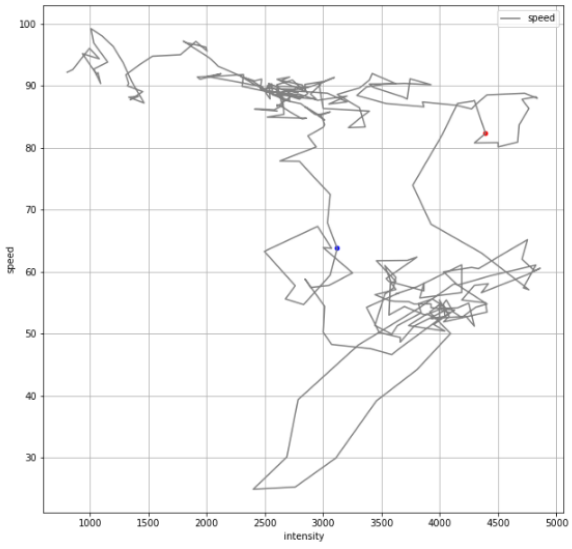
	Raw data	Filtered data
5 MA		
3 MA		
No MA		

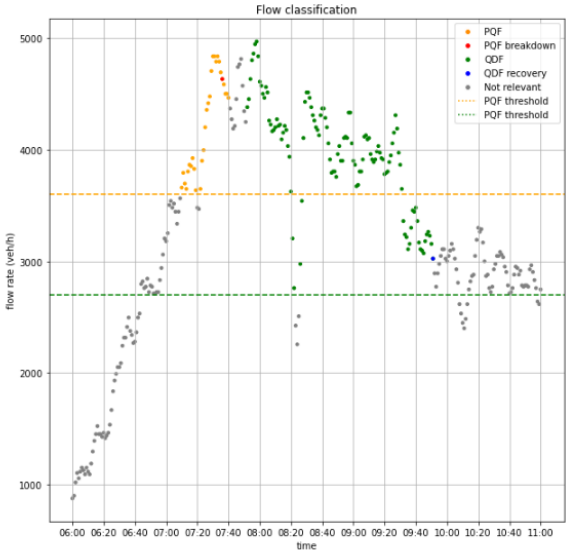
First, when we look at the smoothening of the raw data, it can be observed that non-smoothened data is highly variable. The measured speed does fluctuate around the speed threshold constantly, which complicates the flow type classification. The 3-MA and 5-MA show more stable speed measurements while maintaining the clear speed drop around 7:45. When looking at the maximum observed flow measurements, it can be observed that the 5-MA flattens the peak around 7:45, resulting in a maximum flow of about 4800 veh/h, while the 3-MA and No MA both capture a flow rate above 5200 veh/h.

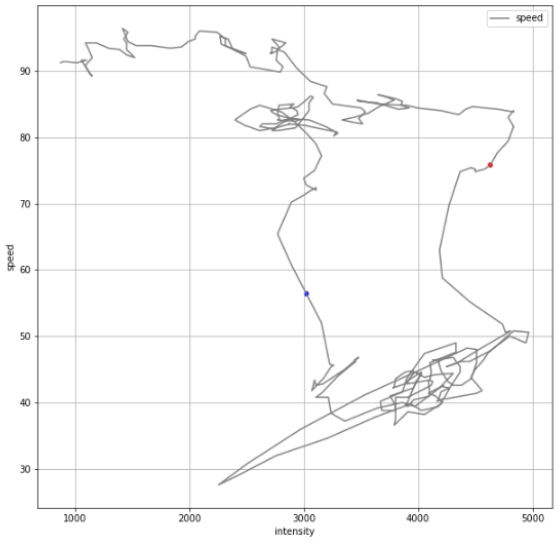
Second, the filtered data seems more stable for all three moving averages. In all cases, the speed drop is clearly below the speed threshold, which eases the breakdown and recovery detection. The 5-minute filtered speed measurements seem most comparable to the 5-minute aggregation technique used in Geistefeldt (2009).

When stop-and-go waves occur, this method seems to best disregard the shockwaves caused. However, again the maximum flow value is flattened and appears to be about 5-10% lower compared to the 3-MA and No MA. Now that the attributes of the smoothening techniques have been investigated, we have to look into how the Flow type classification model handles the different input data.

<table>
<tr><th></th><th>Raw data</th><th>Filtered data</th></tr>
<tr><td>5 MA</td><td>


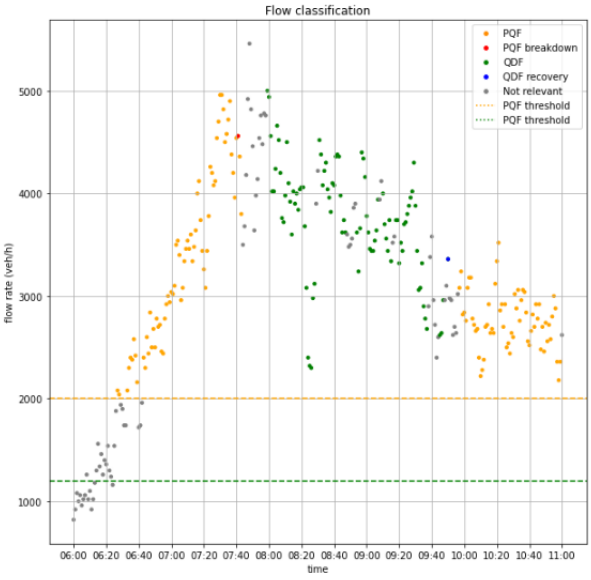


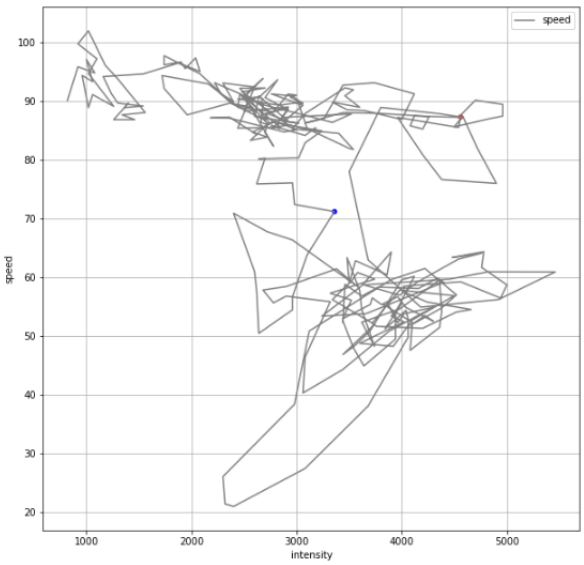
<table>
<tr><th>Time</th><th>Speed</th><th>Speed downstr.</th><th>Flow rate</th><th>Uncen-sored</th><th>Cong. type</th></tr>
<tr><td>07:41</td><td>82.28</td><td>99.73</td><td>4392</td><td>1</td><td>PQF</td></tr>
<tr><td>09:49</td><td>63.78</td><td>89.35</td><td>3120</td><td>1</td><td>QDF</td></tr>
</table>
</td><td>


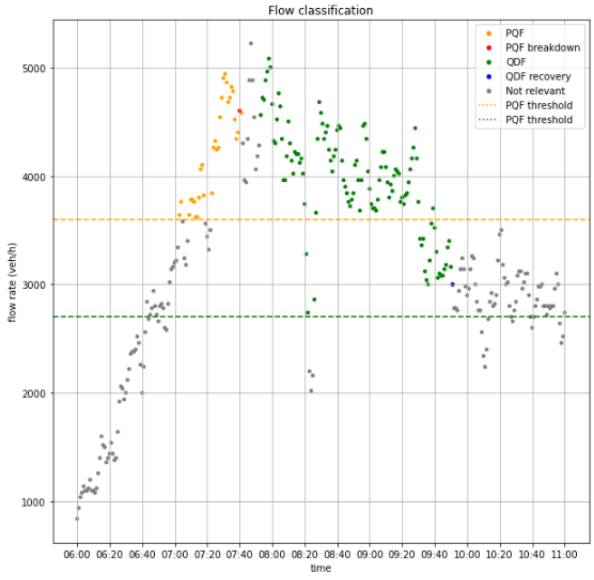


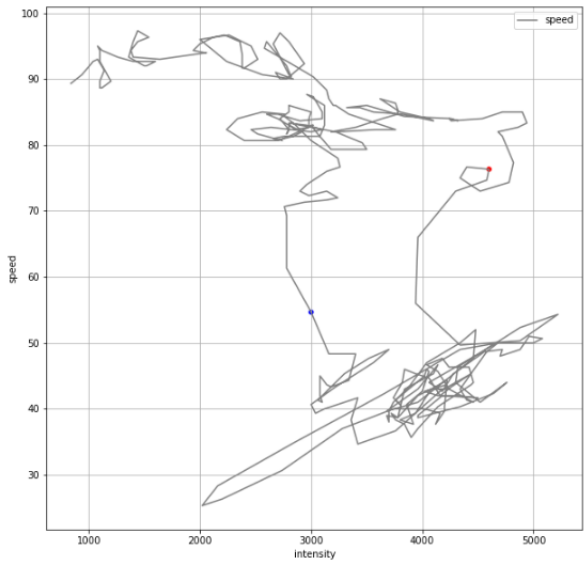
<table>
<tr><th>Time</th><th>Speed</th><th>Speed Downstr.</th><th>Flow rate</th><th>Uncen-sored</th><th>Cong. type</th></tr>
<tr><td>07:36</td><td>75.8</td><td>103.2</td><td>4632</td><td>1</td><td>PQF</td></tr>
<tr><td>09:51</td><td>56.4</td><td>97.0</td><td>3024</td><td>1</td><td>QDF</td></tr>
</table>
</td></tr>
</table>

3 MA



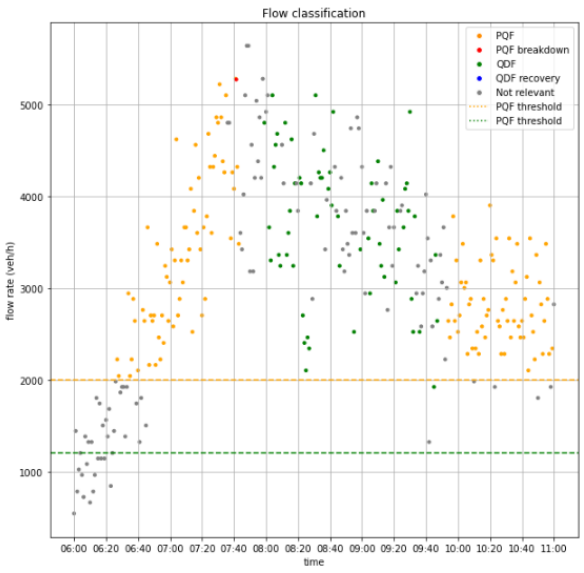


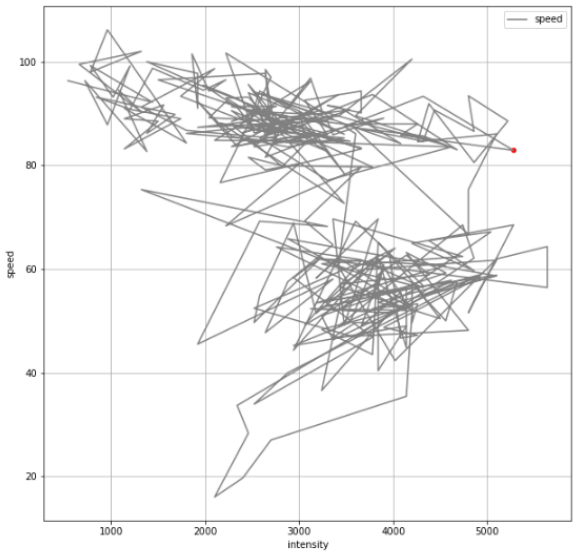


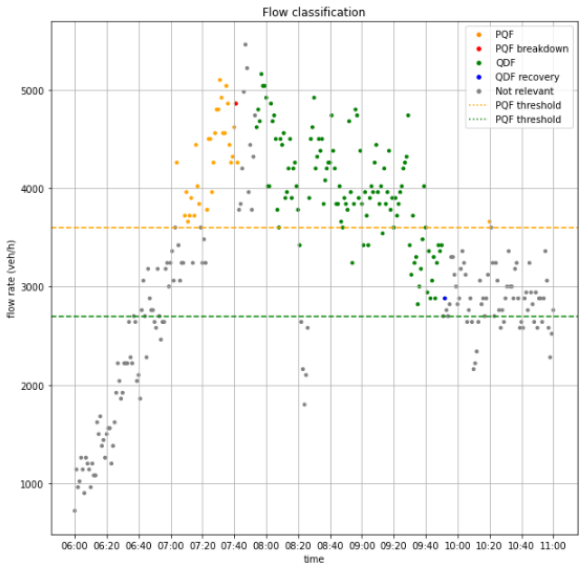


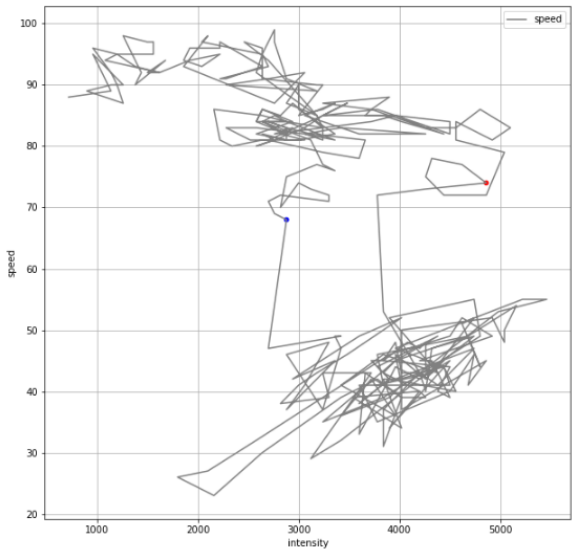
Time	Speed	Speed downstr.	Flow rate	Uncen-sored	Cong. type	Time	Speed	Speed downstr.	Flow rate	Uncen-sored	Cong. type
07:41	87.35	98.25	4560	1	PQF	07:40	76.33	103.67	4600	1	PQF
09:50	71.20	90.03	3360	1	QDF	09:51	54.67	97.0	3000	1	QDF

No MA









Time	Speed	Speed downstr.	Flow rate	Uncen-sored	Cong. type	Time	Speed	Speed downstr.	Flow rate	Uncen-sored	Cong. type
07:41	82.89	93.5	5280	1	PQF	07:41	74.0	100.0	4860	1	PQF
-	-	-	-	-	-	09:52	68.0	97.0	2880	1	QDF

The uncensored classification from the Flow type classification model does an accurate job in obtaining the maximum flow value at similar time stamps for 5 out of the 6 techniques. Now, it becomes clear that the 5 MA filtered data has a flow peak at a different timestamp compared to the other techniques. The short speed recovery before the actual breakdown is not by this technique, while this is the case for the other techniques. The flow value obtained for the PQF, for the 5MA, 3MA, and No MA for the filtered data is respectively 4632, 4560, and 4860. While this is 4392, 4560, and 5280 for the raw data. It becomes clear that the 5 MA value for the raw data is the lowest and thus would underestimate the entire capacity in a larger analysis. On the other hand does the No MA technique not capture the uncensored recovery flow due to the large variability between the measurements from minute to minute.

Concluding, the 3 MA of the raw data seems to be most representative of the actual measurements while allowing the Flow type classification model to detect both breakdowns and recoveries. Therefore, the 3 MA raw data technique is chosen as the technique for modification of the data. Regarding the filtered data, the

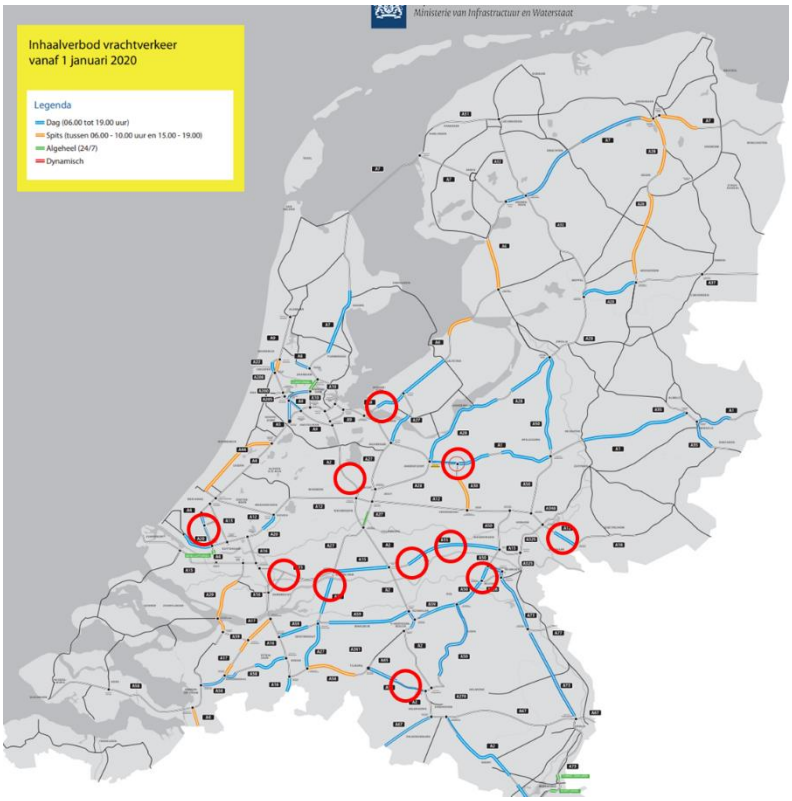
No MA technique seems to be useful in breakdown/recovery detection while remaining as close as possible to the actual measured values. However, the filtered data does not distinguish different vehicle category flows.

B Explanation of selected locations

Free-way	Location	Study period	Geometric			Traffic			Control	
			No. of lanes	Type of bottleneck	Slope (%)	Speed limit	Share of heavy vehicles	No. of traffic breakdowns	Static	Dynamic
A1 R 52.9	Barneveld	2013 - 2020 (E)*	3	Off-ramp	>5	100	2-3%	864	Overtake ban 6:00-19:00	VSL + PHL
A2 R 52.7	Amsterdam – Utrecht	2014 - 2019 (E)*	5	On-ramp	>5	100	3-4%	2307	-	VSL
A4 R 56.6	Delft-Schiedam	2016 - 2019 (E)*	2	Off- and On-ramp	>5	100	4-6%	1490	Overtake ban 6:00-19:00	VSL
A6 R 50.7	Amsterdam – Almere	2013 - 2016 (E)*	2	On-ramp	>5	100	0-3%	432	Overtake ban 6:00-19:00	VSL
A12 R 139.1	Arnhem – Germany	2015 - 2020 (E)*	2	Off-ramp	>5	130	5-8%	4556	Overtake ban 6:00-19:00	VSL
A15a L 80.4	Sliedrecht	2013 - 2018 (E)*	2	Off- and On-ramp	>5	120	6-10%	1225	-	VSL
A15b R 122.2	Gelder-malsen	2014 – 2018 (E)*	2	Off- and On-ramp	<5	130	8-11%	1486	Overtake ban 15:00-19:00	-
A15c R 131.7	Tiel	2014 – 2018 (E)*	2	No emer-gency lane	>5	130	7-11%	381	Overtake ban 15:00 – 19:00	-
A27 R 34.3	Utrecht – Breda	2014 – 2017 (M)*	2	Bridge	>5	100	4-8%	563	Overtake ban 6:00-19:00	VSL
A50 R 139	Ravenstein	2013 – 2017 (E)*	2	On-ramp	>5	130	5-8%	1560	Overtake ban 6:00-19:00	LCB
A58 L 19.6	Moeigestel – Oirschot	2013 – 2018 (M)*	2	On-ramp	>5	120	4-7%	1611	Overtake ban 6:00-19:00	VSL + LCB

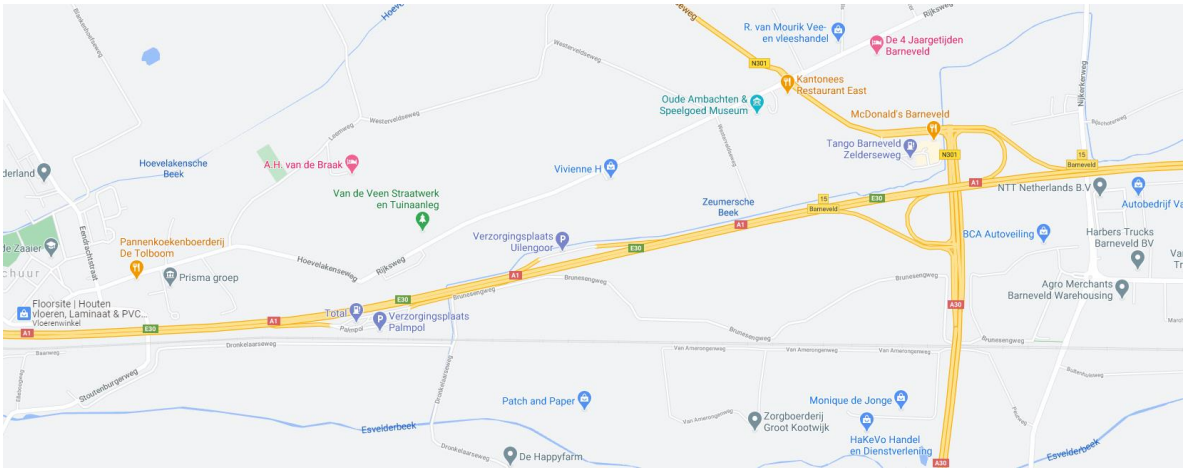
*(E) = Evening (14:00-19:00), (M) = Morning (6:00-11:00)

The spatial location of the roads is visualized on the map below.



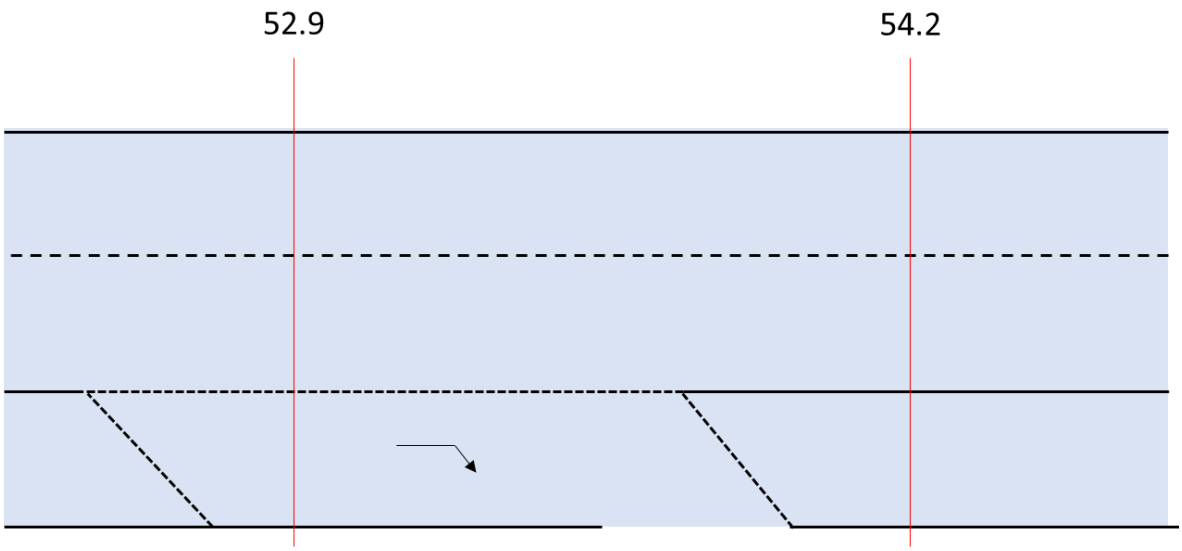
Subsequently, zoomed-in pictures from Google Maps are taken to visualize respective the bottleneck from above, the bottleneck detector, the downstream detector, and the schematic overview.

A1: Barneveld - 14:00-19:00 - 2013-2017

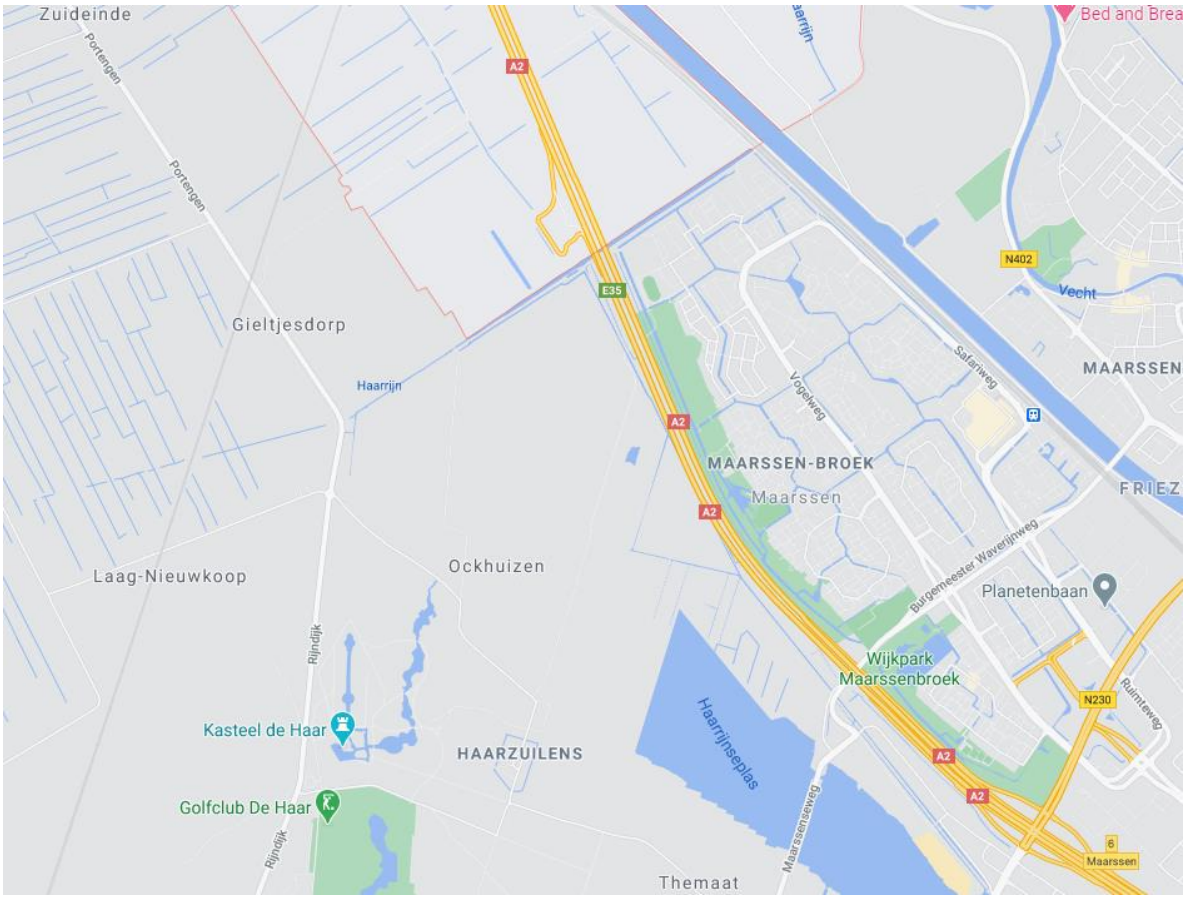


Around hecto 52.9 (2019):

Around bottleneck (2019):

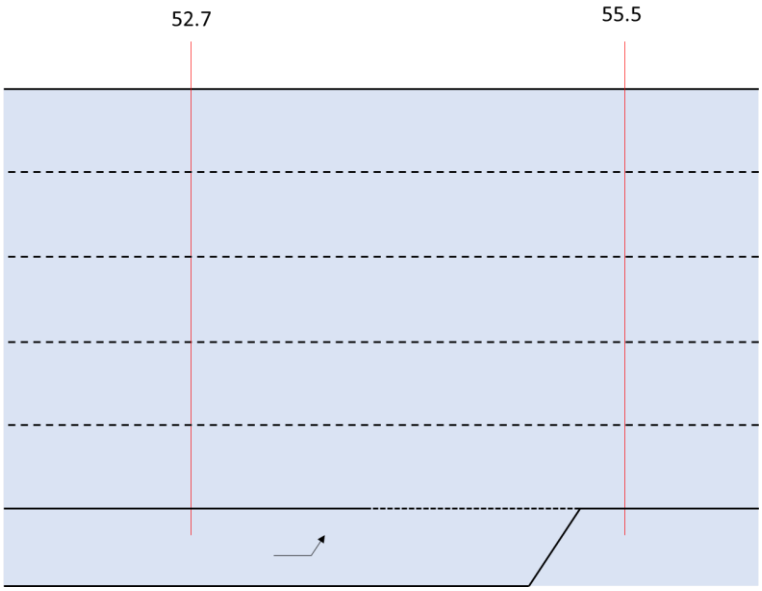


A2: Breukelen – Maarssen – 14:00-19:00 - 2014-2019



Around hecto 52.7 (2018):

Around bottleneck (2018):

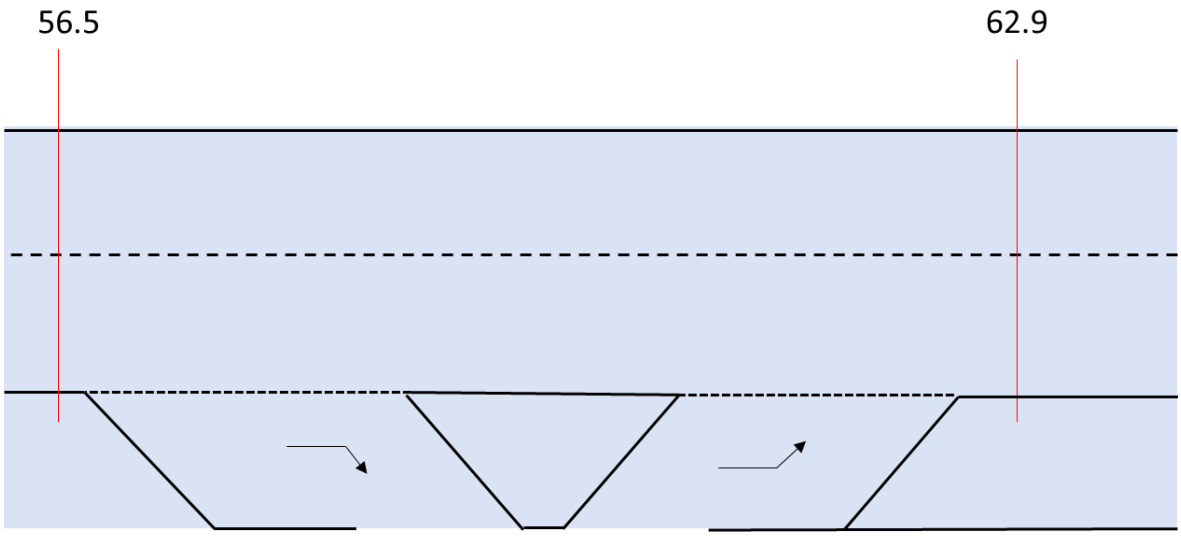


A4: Delft – Schiedam – 14:00-19:00 - 2016-2019

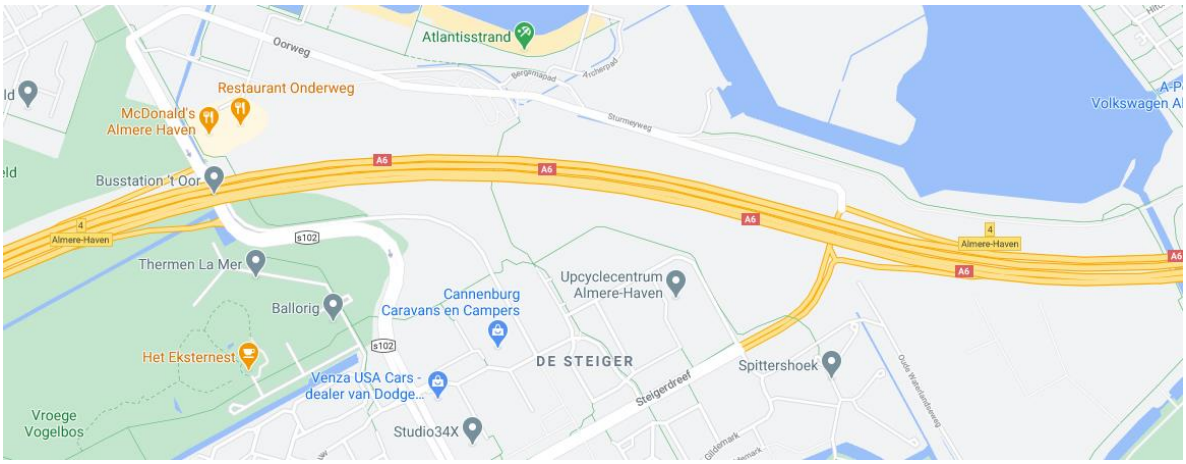
Around hecto 56.5 (2019):

Around bottleneck on-ramp (2019):

Around bottleneck off-ramp (2019):

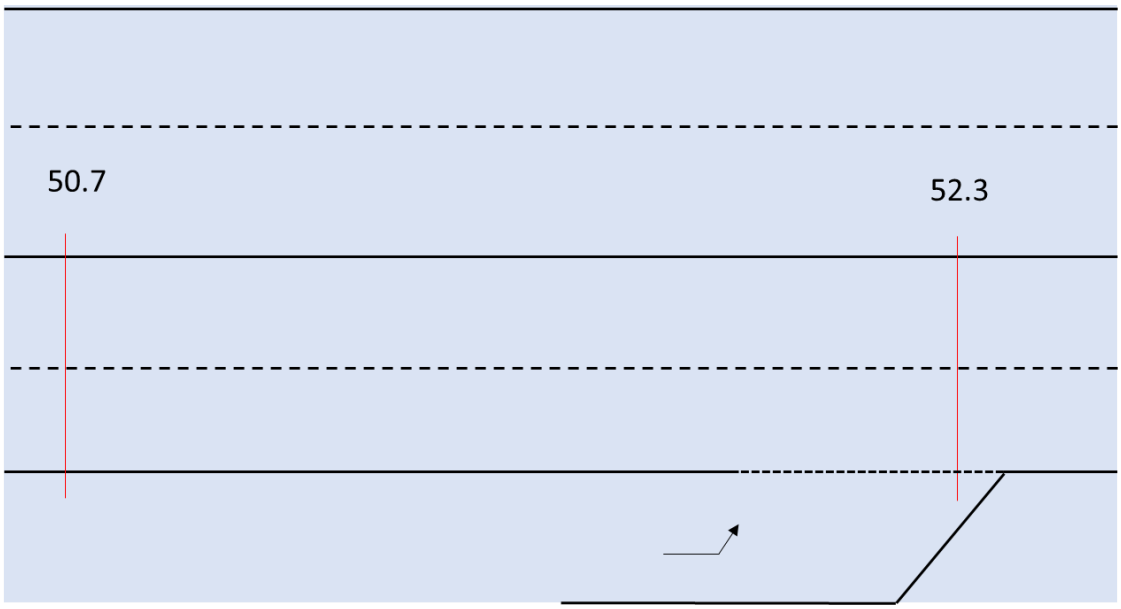


A6: Amsterdam – Almere 14:00-19:00 - 2013-2016

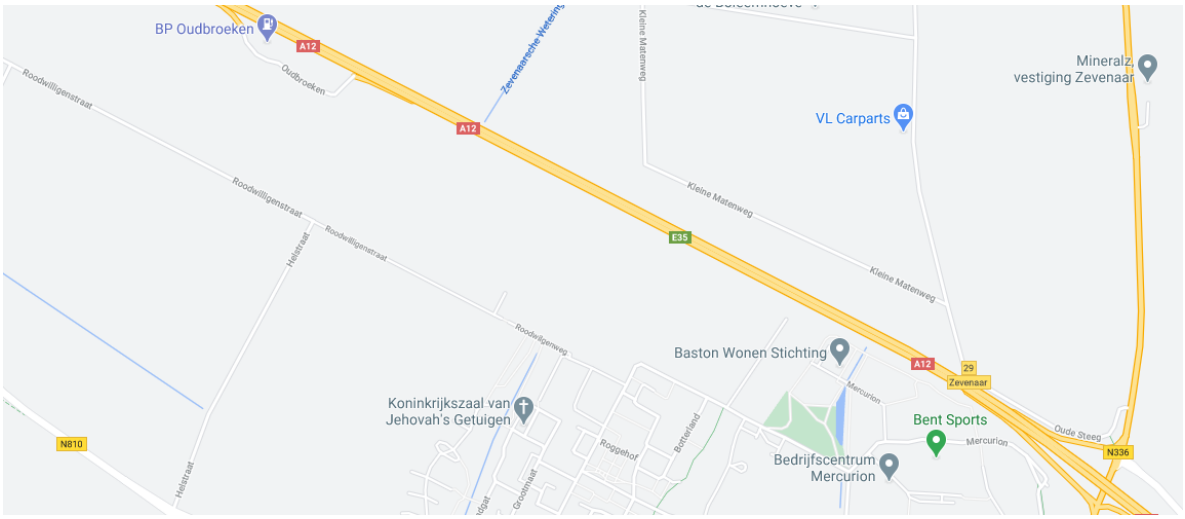


Around hecto 50.7 (2019):

Around bottleneck (2019):

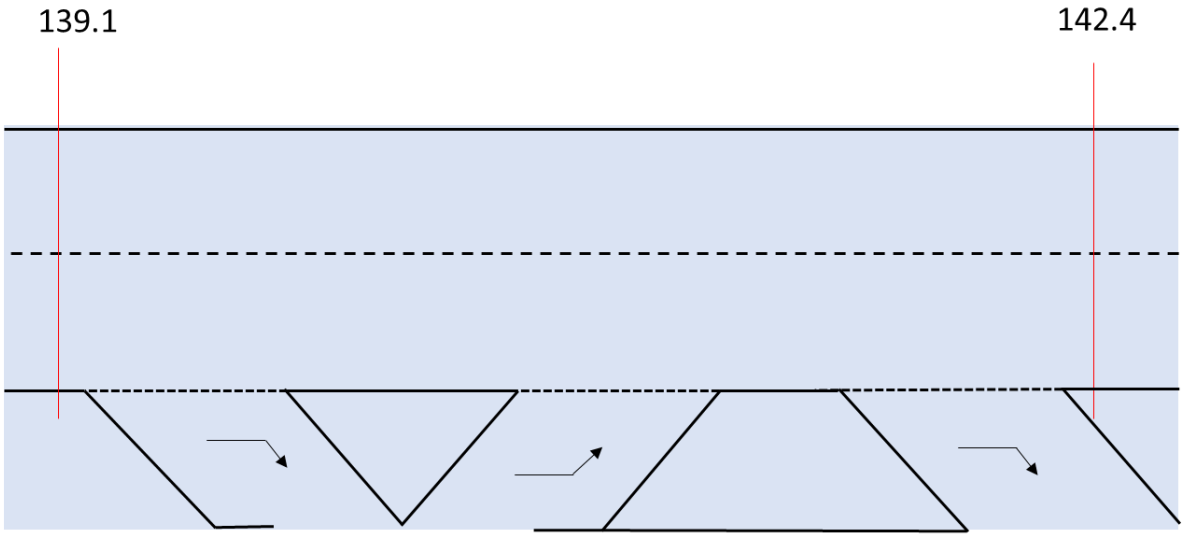


A12: Germany – Arnhem - 14:00-19:00 - 2015-2020

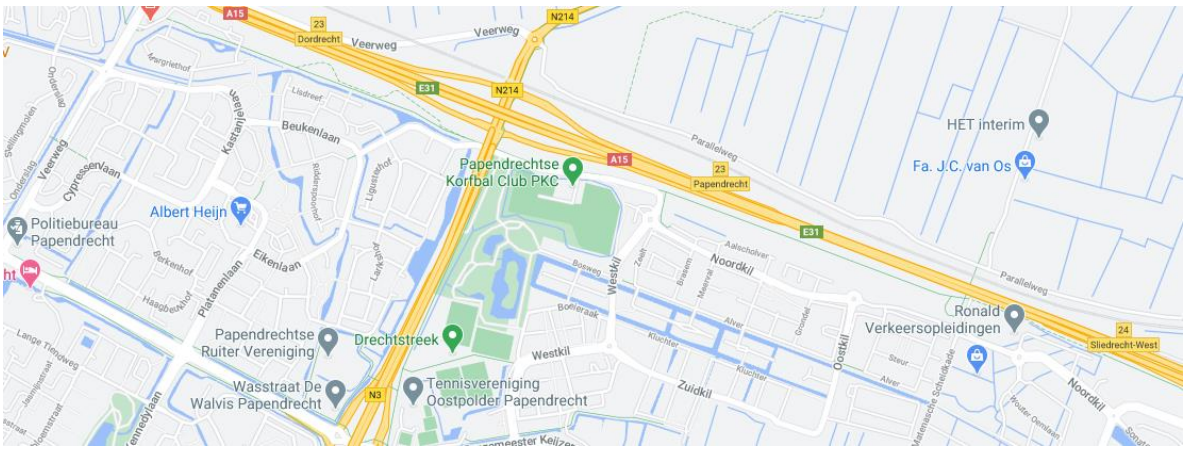


Around hecto 139.1 (2019):

Around bottleneck (2019):

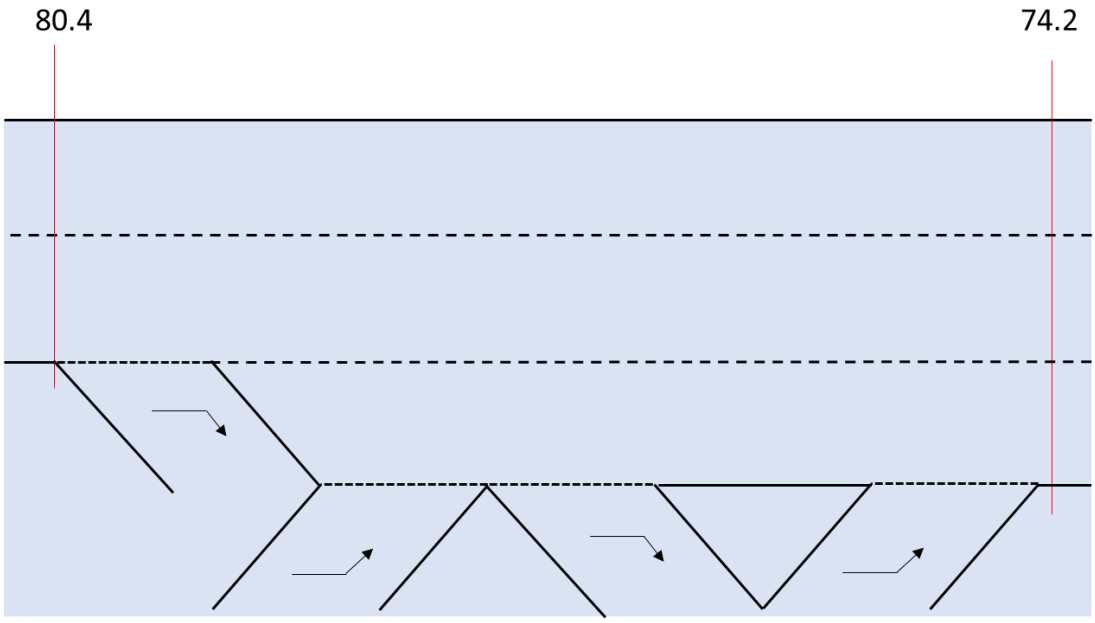


A15a: Sliedrecht – 14:00-19:00 – 2013-2018



Around hecto 80.4 (2014)

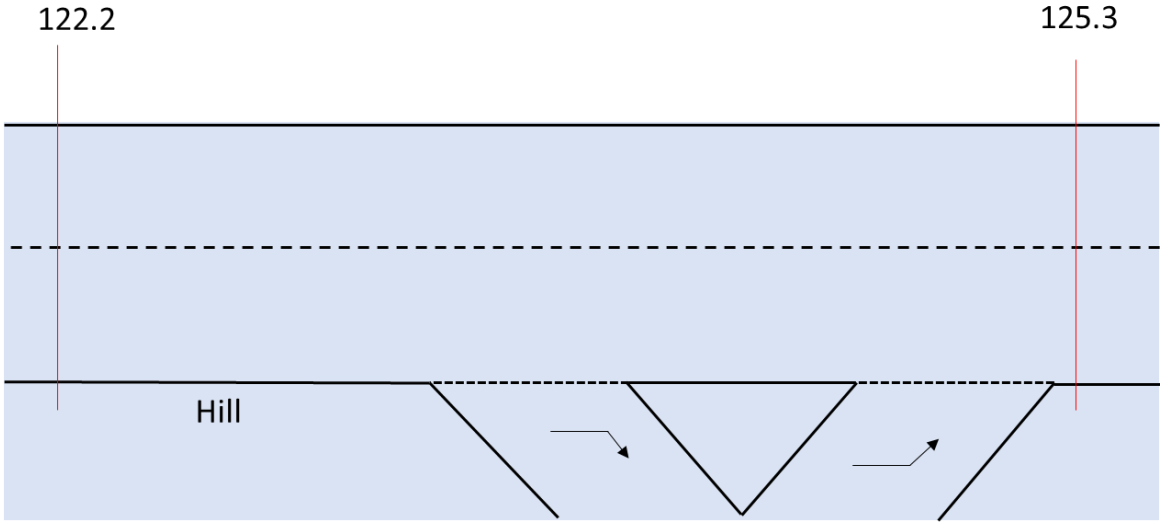
Around bottleneck (2019):



A15b: Geldermalsen – 14:00-19:00 – 2014-2018

Around hector 122.2 (2019):

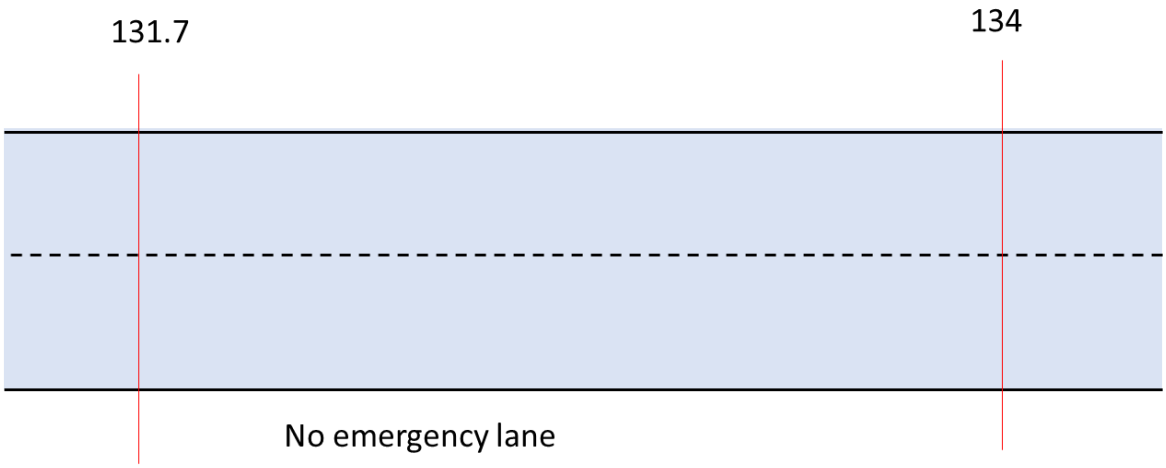
Around bottleneck (2018):



A15c: Tiel – 14:00-19:00 – 2014-2018

Around hecto 131.7 (2019):

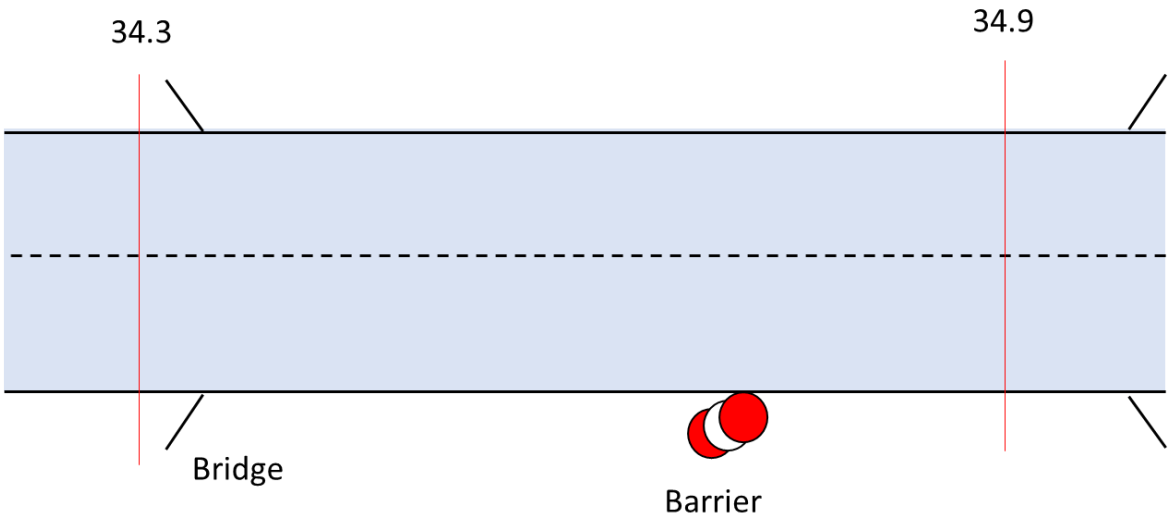
Around bottleneck (2017):



A27: Merwedebrug - 06:00-11:00 - 2014-2017

Around hecto 34.3 (2018):

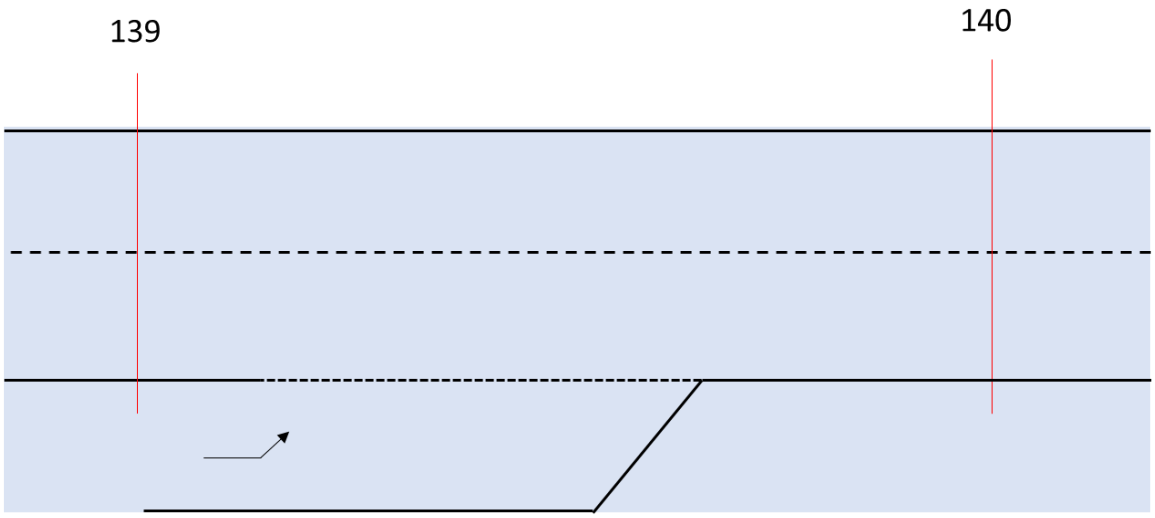
Around bottleneck (2018):



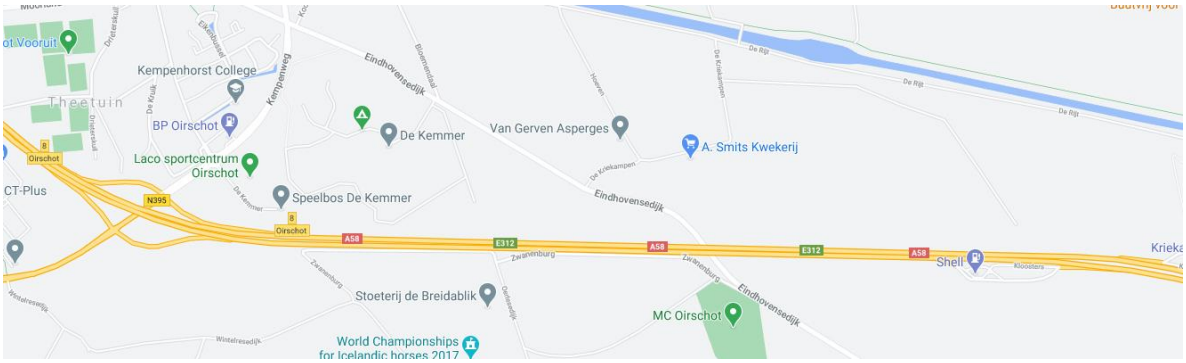
A50: Ravenstein - 14:00-19:00 - 2013-2018

Around hecto 139 (2018):

Around bottleneck (2018):

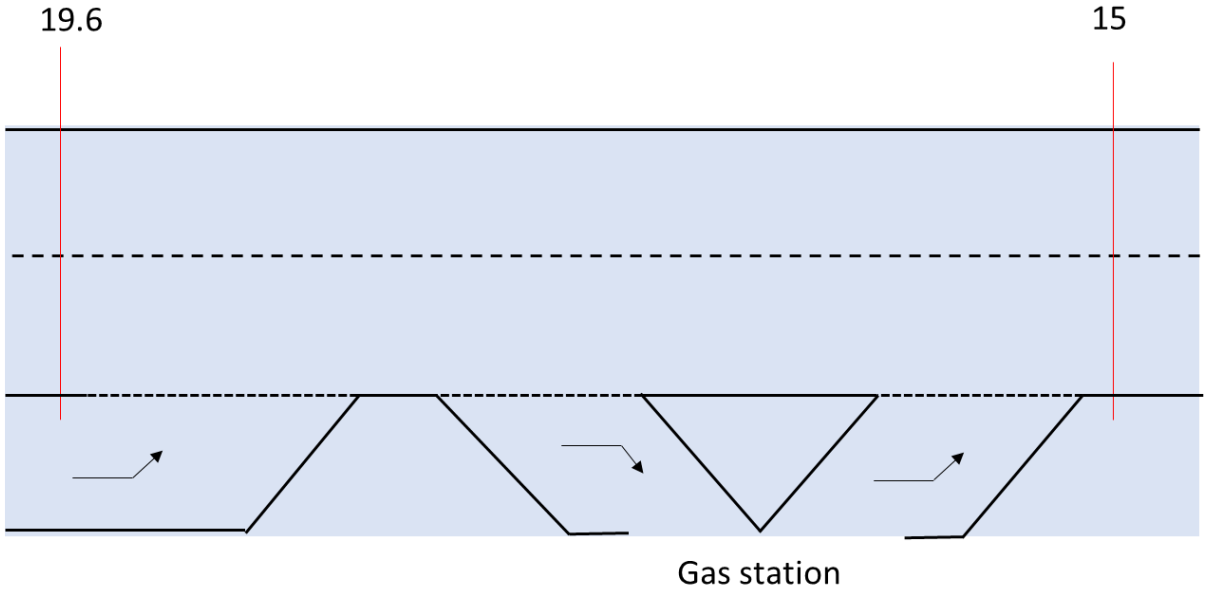


A58: Moergestel – Oirschot - 06:00-11:00 - 2013-2018



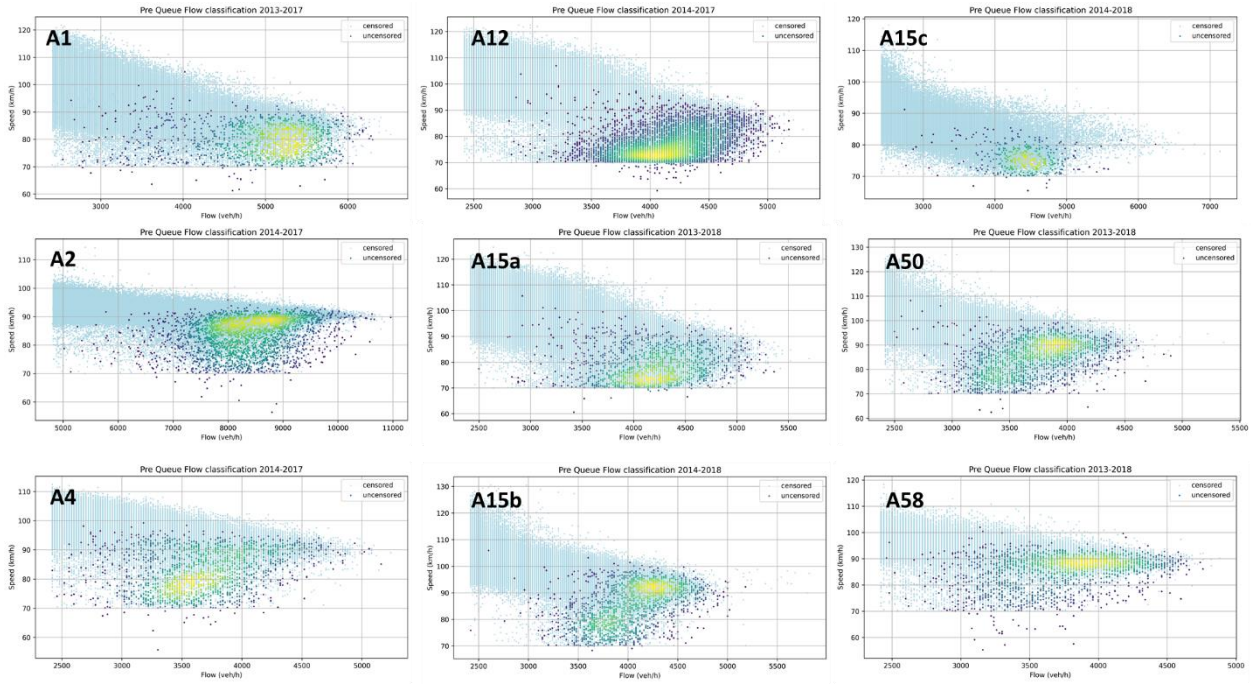
Around hecto 19.6 (2018):

Around bottleneck (2018):

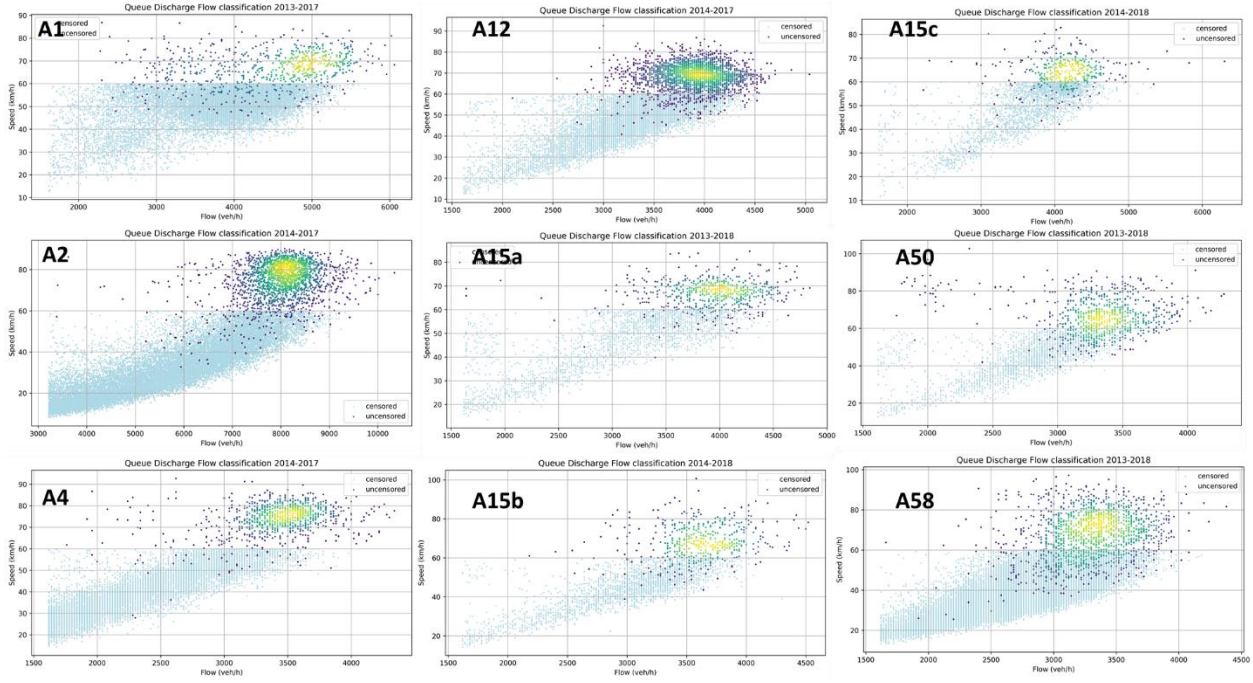

19.6
15
Gas station

C Overview of classified measurements of sites

In the figure below, the PQF measurements for all sites (except A6 and A27) are visualized. The uncensored measurements are coloured light blue, while the censored measurements are coloured according to a Gaussian Density function as explained in 3.4.1.

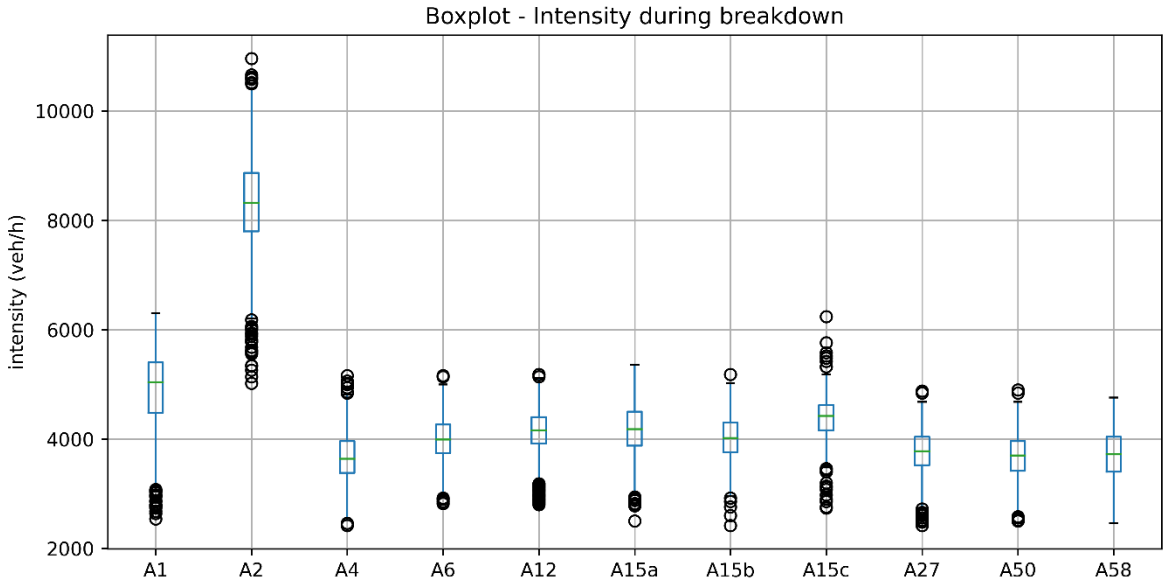


Below, the QDF measurements for all sites (except A6 and A27) are visualized.

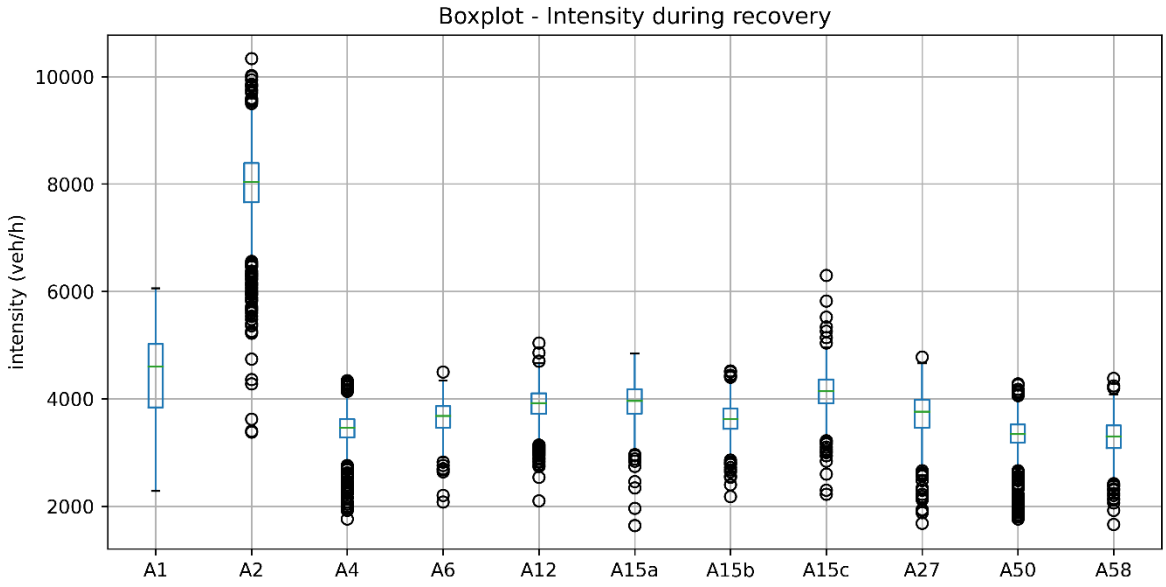


D Flow, speed and vehicle composition boxplots of each site

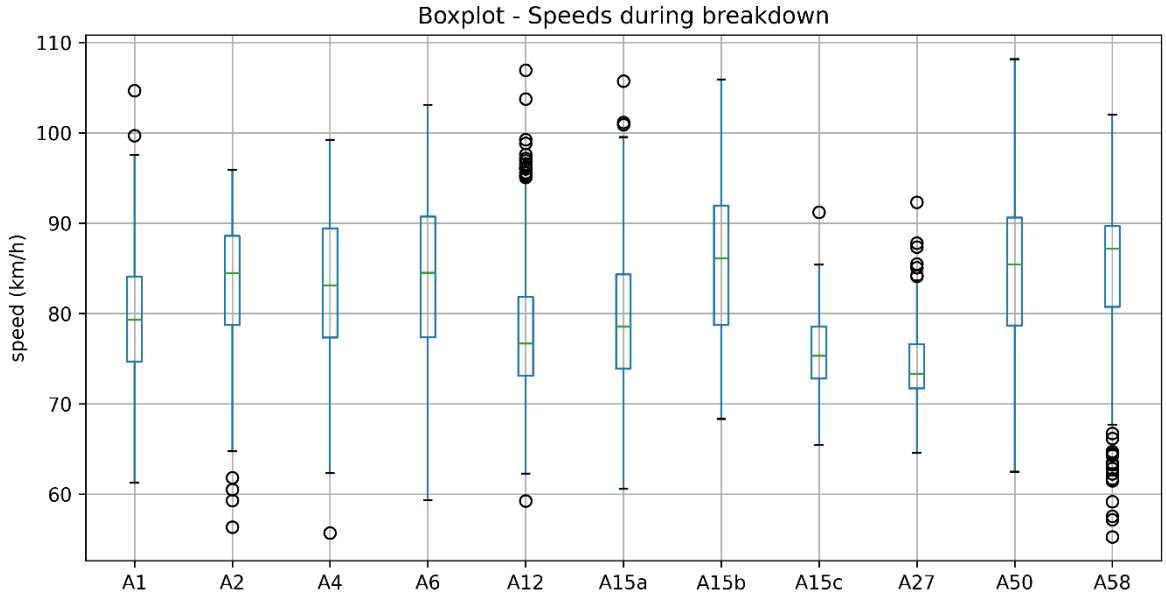
In the figure below, the flow at uncensored classified measurements as breakdowns are visualized in boxplots.



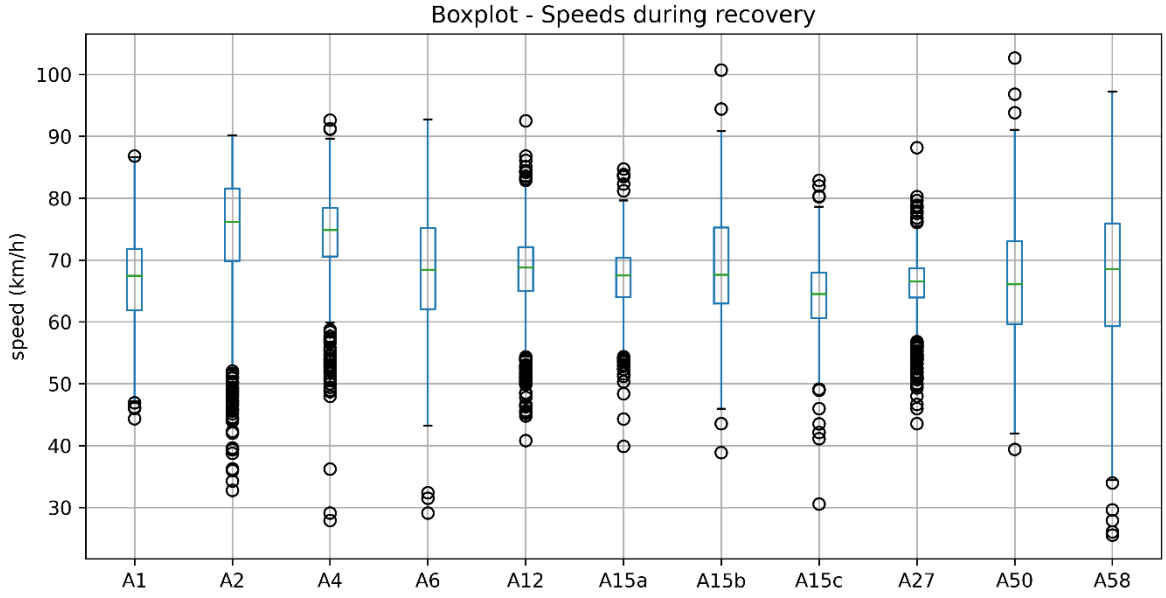
In the figure below, the flow at uncensored classified measurements as recoveries are visualized in boxplots.



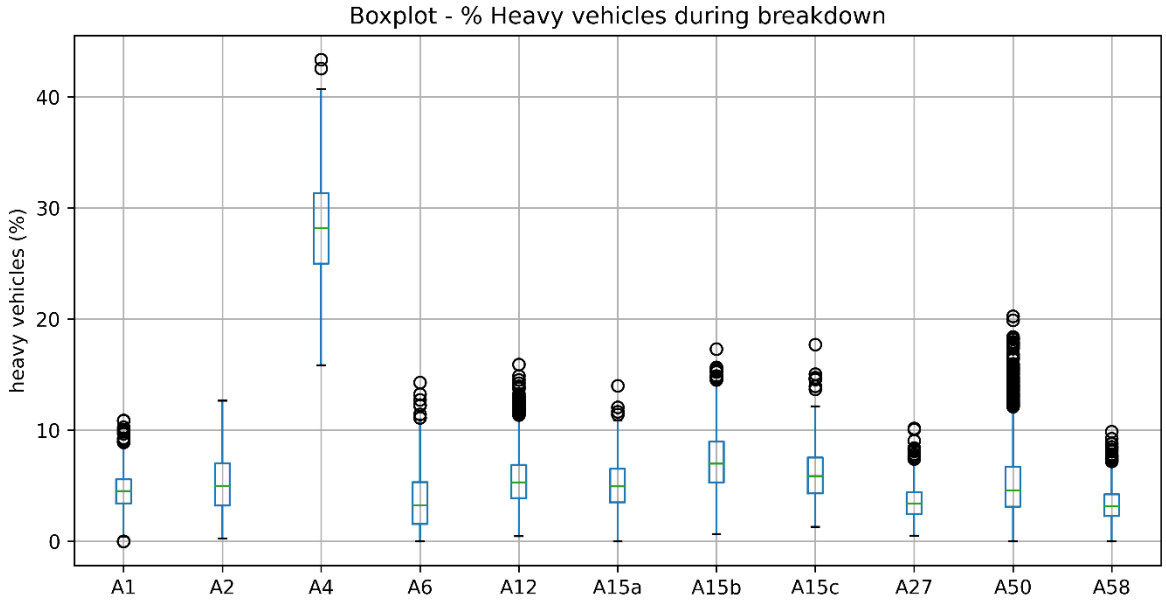
In the figure below, the speed at uncensored classified measurements as breakdowns are visualized in boxplots.



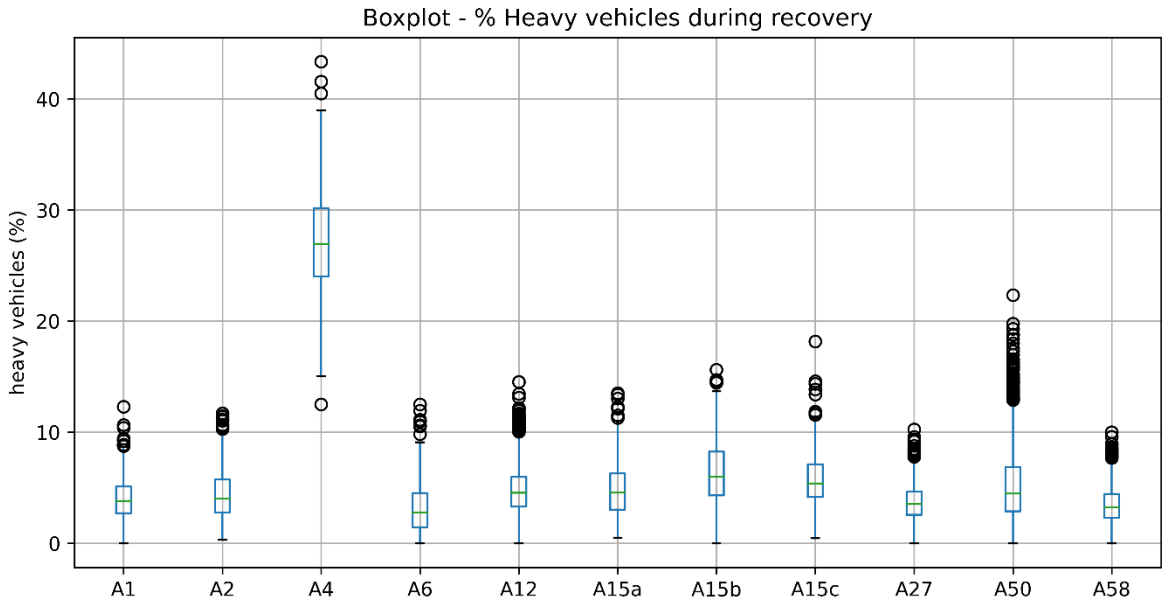
In the figure below, the speed at uncensored classified measurements as recoveries are visualized in boxplots.



In the figure below, the percentage category 2 vehicles at uncensored classified measurements as breakdowns are visualized in boxplots.



In the figure below, the percentage category 2 vehicles at uncensored classified measurements as recoveries are visualized in boxplots.



E Between-year results

For each freeway section per year, the number of breakdowns and discoveries have been displayed. Next, the median of the Weibull fitted distribution is shown for both the PQF and QDF. The capacity drop percentage has been computed based on these medians as explained in 3.4.3.

Table 24 - Computed variables between-year analysis

FREEWAY	VARIABLE	2013	2014	2015	2016	2017	2018	2019	2020
A1	Breakdowns	161	109	164	295	426	300	484	96
	Recoveries	135	79	83	157	297	206	325	50
	PQF median	5688	6420	6200	6042	6093	4116	4052	4199
	PQF std	394	450	379	364	624	336	328	322
	PQF HV	107	122	128	134	129	123	124	141
	QDF median	5732	5323	5186	5315	5465	3461	3440	3493
	QDF std	582	449	361	397	622	273	277	284
	QDF HV	72	95	116	111	110	96	99	127
	Capacity Drop	-0.8	17.1	16.4	12.0	10.3	15.9	15.1	16.8
A2	Breakdowns	54	225	424	655	615	196	117	21
	Recoveries	64	196	393	527	549	143	87	12
	PQF median	11892	10896	10612	10365	10393	10337	10410	11349
	PQF std	963	792	779	761	829	773	775	1014
	PQF HV	236	268	296	300	316	327	370	295
	QDF median	8684	8664	8405	8447	8421	8420	8150	7736
	QDF std	610	612	578	549	544	540	642	838
	QDF HV	232	246	260	271	292	300	327	263
	Capacity Drop	27.0	20.5	20.8	18.5	19.0	18.5	21.7	31.8
A4	Breakdowns				266	500	350	374	
	Recoveries				148	306	251	227	
	PQF median				5172	5093	5080	5136	
	PQF std				469	491	529	556	
	PQF HV				156	174	203	205	
	QDF median				3643	3632	3594	3488	
	QDF std				279	247	263	298	
	QDF HV				147	162	178	179	
	Capacity Drop				29.6	28.7	29.2	32.1	
A6	Breakdowns	49	50	97	205				
	Recoveries	41	42	83	156				
	PQF median	4906	5160	5113	5026				
	PQF std	276	329	346	295				
	PQF HV	69	63	78	86				
	QDF median	3780	3790	3735	3996				
	QDF std	222	212	228	213				
	QDF HV	55	53	59	75				
	Capacity Drop	22.9	26.5	27	20.5				
A12	Breakdowns			931	1105	970	656	752	142
	Recoveries			463	551	508	325	429	84
	PQF median			4910	4872	4829	4721	4678	4707
	PQF std			306	312	314	295	285	278
	PQF HV			255	268	277	293	296	311
	QDF median			4130	4110	4089	4075	4000	4069
	QDF std			279	241	262	297	252	221
	QDF HV			223	244	253	258	250	265
	Capacity Drop			15.9	15.6	15.3	13.7	14.5	13.6

A15A	Breakdowns	234	269	131	258	268	65
	Recoveries	117	118	52	132	122	20
	PQF median	5365	5376	5315	5268	5258	5227
	PQF std	385	413	390	395	410	387
	PQF HV	288	311	310	374	342	365
	QDF median	4208	4195	4209	4090	3974	4089
	QDF std	286	315	306	331	358	274
	QDF HV	251	245	275	327	292	297
	Capacity Drop	21.6	22.0	20.8	22.4	24.4	21.8
A15B	Breakdowns		325	409	368	323	61
	Recoveries		84	119	130	100	18
	PQF median		5019	4915	4730	4711	4667
	PQF std		360	364	305	295	291
	PQF HV		355	356	373	393	406
	QDF median		3881	3861	3807	3794	3731
	QDF std		308	301	248	264	264
	QDF HV		308	322	347	349	359
	Capacity Drop		22.7	21.5	19.5	19.5	20.0
A15C	Breakdowns		105	112	24	81	59
	Recoveries		101	103	35	74	55
	PQF median		6175	6331	6354	6748	5346
	PQF std		550	590	527	694	355
	PQF HV		359	396	333	415	457
	QDF median		4743	4692	4502	4543	4314
	QDF std		375	500	366	448	334
	QDF HV		318	325	277	314	354
	Capacity Drop		23.2	25.9	29.1	32.7	19.3
A27	Breakdowns		97	260	182	24	
	Recoveries		166	406	262	35	
	PQF median		5738	5258	5413	4630	
	PQF std		575	457	469	256	
	PQF HV		249	254	256	230	
	QDF median		3784	3756	3772	3810	
	QDF std		439	397	446	316	
	QDF HV		251	254	267	232	
	Capacity Drop		34.0	28.6	30.3	17.7	
A50	Breakdowns	271	338	347	436	169	
	Recoveries	138	141	165	225	112	
	PQF median	4755	4636	4647	4509	4515	
	PQF std	352	330	347	331	334	
	PQF HV	240	249	247	281	272	
	QDF median	3535	3501	3477	3462	3357	
	QDF std	280	249	280	260	429	
	QDF HV	218	219	235	244	181	
	Capacity Drop	25.7	24.5	25.2	23.2	25.6	
A58	Breakdowns	279	302	273	360	319	78
	Recoveries	237	254	271	338	266	63
	PQF median	4543	4688	4606	4503	4498	4502
	PQF std	325	397	365	338	357	384
	PQF HV	217	213	225	227	247	254
	QDF median	3773	3785	3660	3639	3632	3570
	QDF std	234	256	218	239	232	241
	QDF HV	214	212	233	231	238	253
	Capacity Drop	16.9	19.3	20.5	19.2	19.2	20.7

F Technological Advancements Traffic Flow

Several measures have been taken to improve traffic flow performance. How these new technological innovations impact traffic flow is important for future research areas regarding capacity estimation. New technologies and interventions might lead to an impact on the capacity drop as well and may subsequently be used to reduce or prevent the capacity drop. In the next section, the in-vehicle advancements will be discussed. Thereafter, measures taken from an infrastructure perspective will be elaborated. The summary and discussion is presented in the end.

In-vehicle

The conventional way of driving transitions to an automated way of driving, as technology rapidly evolves. The standard for distinguishing automated driving was created by SAE (2018) and consists of six levels, from no automation to full automation. Conventional cars are of SAE level 0, while these cars at most warn the driver of potential danger but do take over any of the drivers' tasks. SAE level 1 and level 2 support the driver respectively in longitudinal or lateral movements and both combined. One could think of lane centering assist or adaptive cruise control. As more and more manufactured cars do have such features, the market penetration rate is expected to rise almost parabolic the next years (Rijkswaterstaat, 2018). The most recent numbers showed that about 3% of the Dutch fleet has ACC in 2016. This number is forecasted to be 15% in 2020, and 30% in 2025.

The effect on traffic flow, especially capacity, has been simulated and researched in multiple studies. Vander Werf, Shladover, Miller, and Kourjanskaia (2002) researched the increasing proportions of ACC relative to manually driven vehicles in a Monte Carlo simulation. The maximum capacity was achieved at a combination of 60% manual and 40% ACC, as the capacity drops as the ACC proportion increased more. As drivers can set up the following distance in ACC to their preference, the capacity will drop even more when ACC systems are set to keep a larger desired gap compared to manual driving, which is 1.1 seconds. ACC with a mean gap of 1.55 seconds already reduces capacity by an ACC fleet penetration above 20%.

In a long-lasting empirical study, a decreasing tendency of traffic capacity over time was found at multiple bottlenecks (Shiomi et al., 2019). The author thinks ACC could be the cause, as the proportion of hybrid vehicles has increased every year and people likely prefer to increase the following distance with ACC compared to the distance kept when manually driving. Information about the mean gap when accelerating for vehicles with ACC compared to conventional vehicles is also needed to investigate the impact on QDF, and subsequently the capacity drop.

In a naturalistic driving study with 8 participants whose vehicles contained ACC, data was extracted about headway, speed acceleration, and the number of lane changes (Schakel, Gorter, De Winter, & Van Arem, 2017). For accelerations above 0.5 m/s^2, a smaller headway but a larger spacing were found for vehicles with ACC on. Hence, ACC can increase QDF, which would lead to a reduction in capacity drop. Furthermore, the number of lane changes decreased by 36% in congested conditions, which would be also beneficial in reducing the occurrence of traffic breakdowns.

SAE levels 3, 4, and 5 do not require the driver to drive the vehicle and respectively differ in fallback requirement in case of error and road conditions. How these vehicles will impact traffic flow is hard to predict, especially when only a small percentage of the total vehicle fleet will be fully-autonomous in the coming decades (Calvert et al., 2017). The opportunity to optimize traffic flow by vehicle automation seems almost infinite, e.g. with cooperated vehicles that share data extensively. However, the near future will show how a mix of (partially) automated vehicles and conventional vehicles will impact traffic flow.

Infrastructure

Truck specific restrictions

The effects of geometric design on freeway capacity were researched by Laval (2009). A framework for estimating analytical expressions for the capacity reductions caused by a sublet of vehicles forced to slow down at horizontal/vertical curves on multilane freeways was developed. Desired speed in each lane was stochastically modelled and its effect due to moving bottleneck was estimated with the help of Newell's kinematic wave theory. The analytical model showed that for the up-hill section, the increase in capacity due to lane change restriction for heavy vehicles mainly depends on the speed difference between the heavy vehicles and the surrounding vehicles. An increase of 6% capacity could be realised when the heavy vehicles drive about 60 km/h, while the other vehicles drive about 110 km/h. The results show that the improvement is neglectable, but corresponds to the few empirical findings. Notice that the lane-change ratio is kept proportionally low, which is realistic uphill, but does not comply with a bottleneck with on-ramp. Therefore, the author suggests extending the current framework to segments in the vicinity of a freeway entrance/exit and weaving sections. A full explanation of the analytical expression derived could be found in Laval (2009).

The effect of the driving ban for trucks (DBTs) on lane flow distribution (LFD) was examined by Duret, Ahn, and Buisson (2012) on France highways. During summer weekends, the government prohibits trucks to drive on interurban highways to maximize highway capacity for holiday traffic. The researchers compared the traffic flow for days with DBTs and without and report the so-called "U-turn" phenomenon, which characterizes a decrease in shoulder-lane flow in free-flow conditions despite an increase in total flow. When the total flow increases towards capacity, it was observed that how closer to the shoulder lane, the more the throughput on the outer lanes decreases, well below capacity. Consequently, the shoulder lane becomes underutilized, being mainly represented by heavy vehicles. When comparing to DBT circumstances, it was found that the shoulder lane flow varied on weekdays between 675 – 1075 veh/h while during weekend days it was measured to be between 875 – 1125 veh/h. This result suggests that the underutilization of the upper bound flow of the shoulder lane is caused by the number of trucks.

Another measurement to stabilize traffic flow is by implementing truck lane restrictions, in case a large interaction between high truck percentages and large speed differences exists. Trucks overtaking each other may in such cases result in a moving bottleneck, causing surrounding vehicles to get stuck behind the trucks, especially for two-lane roads. Therefore, multiple European countries have introduced lane change restrictions for heavy vehicles (Zhou, Rilett, & Jones, 2019). Moreover, companies place speed delimiters in heavy vehicles to increase safety and reduce fuel usage, which is often below to maximum speed limit, resulting in trucks barely being able to acquire enough speed to overtake another truck. Zhou et al. (2019) simulated several speed limits and truck lane restriction levels for a 4-lane road (e.g., two in each direction) and calibrated the model for Nebraska I-80. As expected, increasing the speed limit of the road, leading to greater speed differences, affected the overall travel time of passenger cars more severely. The results found in these studies show that the PCE values in the HCM-6 do underestimate the impact of heavy vehicles by about 26% for the lowest speed limit. It is even hypothesized that raising speed limits, in this case USA, without adopting mitigation strategies, might lead to traffic flow becoming worse for all vehicles.

Extending on the previous hypothesis, the reverse should also be true. Particularly in the Netherlands, the speed limits recently got lowered during daytime (6:00 – 19:00) on all highways from 130 km/h to 100 km/h (Rijkswaterstaat, 2020a). This new regulation unfortunately was implemented during the lockdown of the Corona-virus, which resulted in way less traffic on the road. A confirmation of this theory therefore cannot be done in the data analysis part but would be interesting when circumstances return to normal.

Ramp-metering

A measurement aimed at directly preventing or reducing the capacity drop is ramp metering. These meters are placed at on-ramps and can subsequently decrease total delay if applied properly. The discharge rate drop can reach up to 15% if left uncontrolled (Cassidy & Bertini, 1999; Cassidy & Rudjanakanoknad, 2005). The

results showed that temporary outflow gains of 10% or more are possible by making use of ramp-metering. However, restrictively metering the on-ramp diminishes the should lane flow below the critical value, which still leads to high outflows in the median lane. Therefore, a relaxing metering rate strategy to allow the shoulder lane flow to recover should be sustained.

Although ramp metering seems to be a good working solution to both prevent breakdown or help recovery after a breakdown, the downside is that the acceleration capabilities of trucks might not be enough depending on the on-ramp (Yang, Xu, Wang, & Tian, 2016). Especially taper-type ramps were, with limited acceleration distance, lead to merging below the mainline speed.

Variable Speed Limit

Another measurement considered to be able to improve traffic conditions is the Variable Speed Limit (VSL) control. Duret et al. (2012) investigated the effect of Variable Speed Limits (VSL) on the utilization of the shoulder lane and the prevention of the capacity drop. With active VSL, the outer lane flow ranged between 775 – 1225 veh/h as opposed to 675 – 1075 veh/h without VSL, indicating that VSL improves the use of the shoulder lane. This effect appeared to be stronger without DBT than with DBT. Subsequently, speed harmonization across lanes can be concluded to improve shoulder lane utilization as passing possibilities diminish due to the lower speed differences. The drop rate for each of the scenarios after congestion sets in has not been computed. Comparing the shoulder lane utilization, due to heavy vehicles, with the drop rate could potentially lead to new insights regarding the capacity drop.

Y. Zhang and Ioannou (2015) investigated whether this measurement could improve truck-dominant highway sections by smoothening and homogenizing the traffic flow. They reported that one of the problems with a VSL scheme is that most lane changes take place in the vicinity of the bottleneck, deteriorating possible improvements provided by VSL. The lane change of a single truck in a queue may already lead to a severe traffic disorder, therefore the researchers combined the VSL control strategy with a Lane Change (LC) strategy. A lane change recommendation was deployed 500m ahead of the bottleneck to give drivers enough space to change lanes. A truck penetration of 30% was simulated and calibrated for freeway data of the I-710 in California. It was found that LC control ensured a high queue discharge rate, while the VSL control dampened the traffic oscillations. The LC control does not reduce the number of lane changes but only forces a more widespread section of them to happen. On the other hand, the VSL was found to be important in homogenizing traffic speed and subsequently reduce lane changes, which is useful in truck-dominant highway sections.

Cho Hyun and Laval Jorge (2020) combined VSL control with a ramp metering system to prevent and recover the capacity drop at merge bottlenecks. The main problem to solve is that VSL prioritizes on-ramp flow as it reduces the speed and consequently flows, on the main road while current Ramp Metering systems tend to prioritize mainline traffic, possibly leading to queue spill back to an upstream arterial road. This combined control is aimed at preventing disruptive lane changes by vehicles both on the on-ramp and/or freeway queues. The results show that VSL control only maintained capacity longer than RM-only control, while the latter recovered the capacity drop more quickly. Combining both control systems did not only prevent capacity drop but also prevented breakdown. The researchers extended their system with on-ramp queue length restrictions and proposed two strategies. Strategy A, which always uses RM and activates VSL when queue spillback is detected, and Strategy S, which is similar but only applies VSL to the shoulder lane. The best performance in terms of reducing delays was found to be the combined VSL and RM model with Strategy S and lead to significant benefits in this simulation experiment.

Summary and Discussion

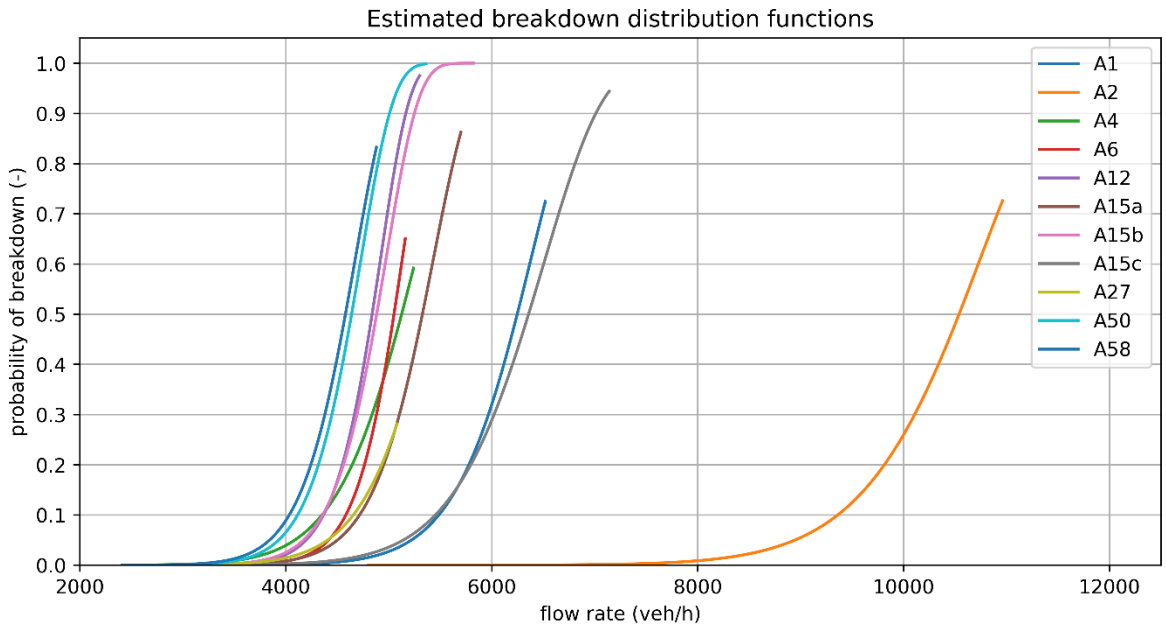
The impact of (technological) advancements on traffic flow can be separated into two categories: in-vehicles and infrastructure. Road authorities can directly influence the traffic flow by adapting infrastructure and subsequently change traffic flow. Technological innovations in-vehicle are the result of car manufacturers trying to win customers while the impact of their technologies on traffic flow is difficult to predict let alone control. One of those technologies that have been implemented on large scale is adaptive cruise control, with an estimated percentage of 15% in 2020 of the fleet (Rijkswaterstaat, 2018). Studies show that after a certain saturation, the effect of ACC-equipped vehicles on capacity will be negative (Vander Werf et al., 2002). Furthermore, since people can set up the desired following distance to their preference, lower capacities over time have already been observed in Japan (Shiomi et al., 2019). In a small study, the QDF was found to be increasing for cars equipped with ACC, and the number of lane changes decreasing (Schakel et al., 2017). For fully autonomous cars, the possibilities to optimize traffic flow are boundless with high saturation rates. However, it is estimated that only a small part of the entire vehicle fleet will be fully-autonomous in the coming decades while mixed traffic might negatively affect traffic flow soon (Calvert et al., 2017).

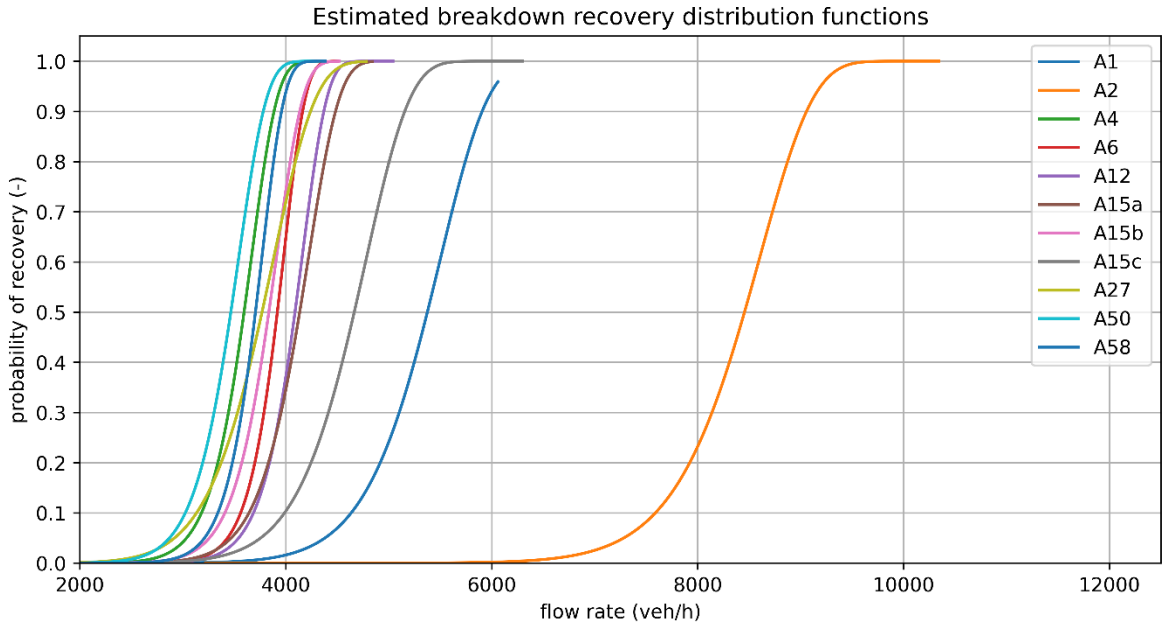
Several infrastructural measures exist to change traffic flow and the impact of heavy vehicles, e.g. truck restrictions, on-ramp meters, and variable speed limits. The effects of driving bans for trucks showed that the number of trucks causes the underutilization of the upper bound flow of the shoulder lane, as the flow increased by about 10% without trucks (Duret et al., 2012). In case passing lane changing trucks lead to moving bottlenecks, lane change restrictions could help to improve stability (Zhou et al., 2019). On-ramp meters have been shown to improve the outflow by 10% when restricting the inflow from on-ramps (Cassidy & Bertini, 1999; Cassidy & Rudjanakanoknad, 2005). However, acceleration limitations of heavy vehicles lead to merging below mainline speed. Especially on-ramps with many heavy vehicles, seem to not benefit from this strategy (Yang et al., 2016). Variable speed limits improve the utilization of the shoulder lane and diminish passing possibilities as speeds get more harmonized (Duret et al., 2012). The possibilities were investigated for truck-dominant highway sections and showed that results drastically improved when combined with a lane change strategy ahead of the bottleneck (Y. Zhang & Ioannou, 2015). Another combination that showed significant benefits in preventing breakdown was the VSL combined with an RM strategy (Cho Hyun & Laval Jorge, 2020).

Multiple infrastructural advancements were investigated and showed insights on how to diminish the capacity drop and improve traffic stability. In-vehicle advancements are less controllable and hard to predict but could be estimated in the short-term. The uncertainty about the electrification of trucks including new emerging technologies such as truck platooning should be thoroughly investigated as different measures from an in-vehicle and infrastructural perspective should be coherent to sophistically improve traffic flow.

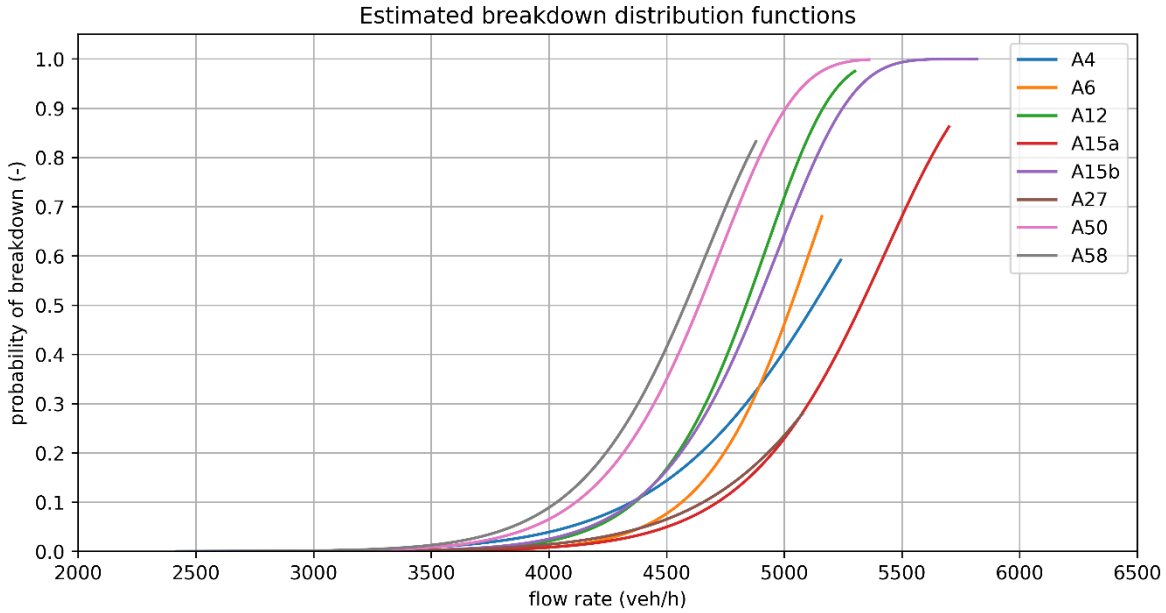
G Confidence intervals distribution functions

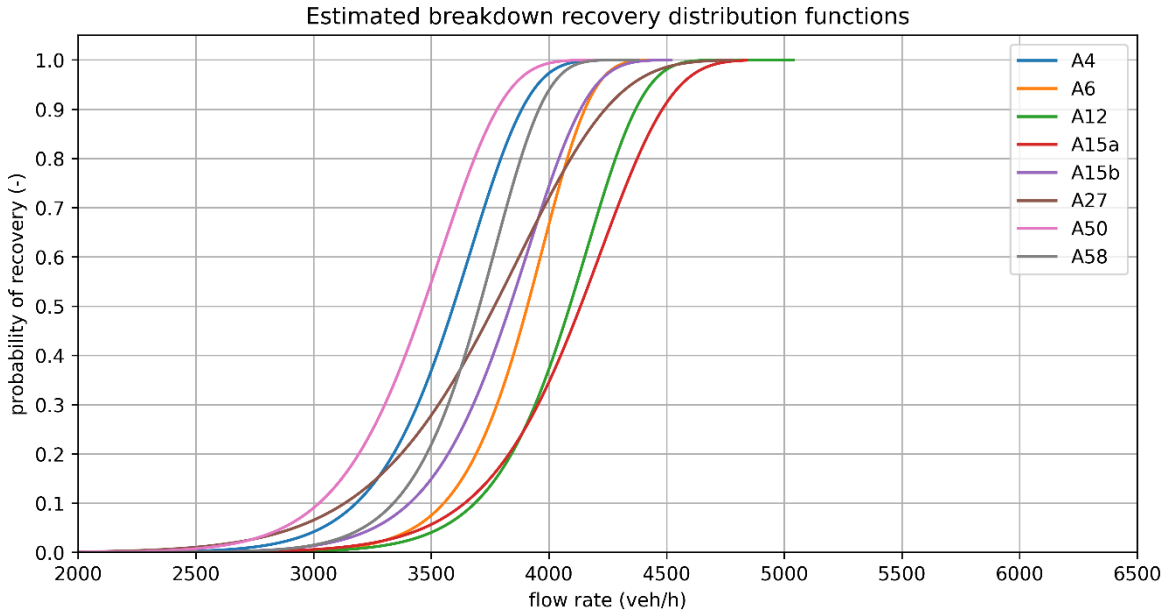
The confidence intervals of the breakdown and recovery distribution functions of the empirical data for all sites can be observed below:





The confidence intervals of the breakdown and recovery distribution functions of the empirical data zoomed-in for the 2-lane sites can be observed below:





The confidence intervals of the breakdown and recovery distribution functions of the simulation data can be observed below: