Het verkeerskundig
laboratorium
voor studenten

ITS EDU LAB

Level of Service estimation with in-vehicle sensor floating car data

Final report

Manasse Hutte

March, 2013

SMART-IN-CAR

Rijkswaterstaat
Ministerie van Infrastructuur en Milieu

TU Delft

Colophon

Published by	ITS Edulab, Delft
Information	Tom Alkim
Telephone	+31 88 798 2468
Author	Manasse Hutte
Graduation Committee	Prof.dr.ir. B. van Arem
	Dr. R.G. Hoogendoorn
	Ir. P.B.L. Wiggenraad
	Ir. L. Visser-Goossens
	Drs. T.P. Alkim
	Dr.ir. R. Happee
Date	March 7th, 2013
Status	Final
Version number	1.00

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

TU Delft

Rijkswaterstaat
Ministerie van Infrastructuur en Milieu

Abstract

In the Smart In-Car project, in-vehicle signals obtained from the CAN (Controller Area Network) bus are collected along with the GPS position from 200 equipped vehicles. The main CAN bus signals in the available data set from the Smart In-Car project are: speedometer, brake usage, steering position, rpm, indicator usage and fuel consumption, but many more CAN bus signals exist, offering possibilities for many applications. In this thesis, the applicability of CAN bus signals for estimating the Level of Service (LOS) on motorways is assessed. The real CAN bus data of single equipped vehicles, obtained from the Smart In-Car project, is used to estimate the LOS, using loop detector data as ground truth. To assess the LOS estimation accuracy at higher penetration rates, simulation data is used. The available real CAN bus signals provide a correct LOS estimation of 37% with single vehicles. The simulation results show correct LOS estimation rates ranging from 52% at 1% penetration to 70% at 100% penetration. LOS estimation using only positional data (GPS) is found to outperform LOS estimation using CAN bus data for most penetration rates, showing that the CAN bus signals do not contribute to better LOS estimation. There are however many other applications which may benefit from CAN bus signals. These will require further research.

Preface

This thesis marks the end of my graduation project, the final stage of the master study Civil Engineering, track Transport & Planning, at the Delft University of Technology. The research leading to this master thesis was conducted at the ITS Edulab, a collaboration between the Transport and Planning group of the Delft University of Technology and the Centre for Transport and Navigation of the Dutch Directorate-General for Public Works and Water Management (Rijkswaterstaat).

This master thesis is part of the Smart In-Car project, which is a cooperation between TASS, Beijer Automotive, NXP Semiconductors, TNO, IBM, Rijkswaterstaat, Cibatax, ANWB, TU/e, Nokia, KPN, Navteq and LaQuSo. This project is part of the subsidiary program Brabant In-car II by the ministry of Infrastructure and Environment, the Province of Noord-Brabant and the Eindhoven Region Collaboration. Brabant In-car II is in turn part of the Beter Benutten (Better Utilization) program, which has the objective to reduce the amount of congestion at specific roads and at bottlenecks where congestion is most severe with 20-30% in three years, by improving the utilization of the existing infrastructure.

Thanks go out to my graduation committee and especially to my daily supervisors.

Manasse Hutte
Delft, March 2013

Executive summary

To improve traffic conditions and safety, a large number of traffic management applications are being used in daily practice. These traffic management applications need information on the prevailing traffic conditions to adequately function. To obtain these, traffic sensors are used. Traffic can be monitored through a large range of different traffic detectors. These detectors are often quite inaccurate and costly. In the Netherlands, induction loops are the most widely used traffic sensors.

The Smart In-Car project offers the possibility to investigate a new type of traffic monitoring. In this project, the values from several in-vehicle sensors, obtained from the CAN (Controller Area Network) bus, and the GPS positions are collected with the uCAN on board unit from a small number of equipped vehicles. In the project, around 200 vehicles were equipped with the uCAN on board unit. In this study, only part of the data is used, containing two months of data from 35 identical vehicles, part of a taxi company fleet, because of the large amount of data and the associated computing time required for the data processing. The main in-vehicle signals in the data set are: speedometer, brake usage, steering position, rpm, indicator usage and fuel consumption.

In the Smart In-Car project, multiple use cases related to traffic state estimation, fleet management, road maintenance, accident registration, traffic management and weather are proposed which could benefit from the in-vehicle sensor floating car data.

In this research, several traffic management related use cases are defined in more detail and one possible application of the CAN bus data is examined. The objective of this research is to determine if Level of Service (LOS), a categorical qualitative description ranging from A to F of the traffic conditions on a road, defined by the density, can be accurately estimated with in-vehicle sensor floating car data.

This is assessed by comparing the accuracy of three LOS estimation methods:

- LOS estimation with in-vehicle sensor floating car data
- LOS estimation with plain floating car data
- LOS estimation with loop detector data

In-vehicle sensor floating car data is defined as the CAN bus data which is collected with the uCAN on board unit. The uCAN on board unit also provides the GPS positions of the equipped vehicles. When LOS estimation with the CAN bus data is considered, the GPS data is only used to determine the location of the vehicles, but not to estimate the LOS. Plain floating car data is defined as exclusively the positional data (provided by the GPS unit of the uCAN on board unit) of the equipped vehicles.

Since the number of equipped vehicles is very limited, the real CAN bus data set is only suitable to assess the LOS estimation with single vehicles. To assess the LOS estimation of in-vehicle sensors with multiple vehicles at various penetration rates, traffic simulation data, containing vehicle trajectories and loop detector data, is used. In the simulation, not only detection point data is used to determine the LOS ground truth, but also the actual density, determined with the vehicle

trajectory data. Using the simulated CAN bus data, the LOS is also estimated with single vehicles, to be able to compare the results with the real CAN bus LOS estimation. Therefore, in total 5 methods are used for LOS estimation:

- LOS estimation with real CAN bus data from single vehicles
- LOS estimation with simulated CAN bus data from single vehicles
- LOS estimation with simulated CAN bus data from multiple vehicles
- LOS estimation with simulated plain floating car data from multiple vehicles
- LOS estimation with real loop detector data

From the real CAN bus data set, the data points of vehicles passing a loop detector location are combined with the respective loop detector data. The loop detector speed and intensity data are used to determine the LOS ground truth at the time and place of the vehicle passage. The combined data is used to identify the CAN bus signals which show the strongest correlation with the LOS. Also, regression analysis is performed to determine functions which are used to estimate the LOS with the single vehicle passages. Of the available CAN bus signals, using only the average values from the speedometer signal yields the most accurate LOS estimation functions. A correct LOS estimation rate of 37% is found.

First, the simulation data is used to assess the LOS estimation accuracy with single vehicle passages, as is done with the real CAN bus data. The determined correct LOS estimation rate is 52%. The difference between the found correct LOS estimation rates has multiple causes and leads to the assumption that the LOS estimation with the simulation data can be used to determine an upper boundary of the real LOS estimation accuracy.

Secondly, the LOS is determined using the simulated CAN bus signals of multiple vehicles, for different penetration rates. The found correct LOS estimation rates range from 52% at 1% penetration to 70% at 100% penetration.

Finally, the LOS at different penetration rates is also determined with the simulated data using just the vehicle positions, to simulate plain floating car data. This is done by dividing the density, based on the equipped vehicles, with the penetration rate to obtain the estimated density and LOS. For this method, the correct LOS estimation rate ranges from 39% at 1% penetration to 100% at 100% penetration.

The results show that in-vehicle sensor floating car data generally provides a less accurate LOS estimation than plain floating car data. Only in case of penetration rates lower than 8%, LOS estimation with CAN bus data may outperform plain floating car data. Also the LOS estimation accuracy with loop detectors, estimated at 89%, is much higher than the LOS estimation accuracy that is reached with CAN bus signals.

Based on the results, it is concluded that it is not possible to accurately estimate the Level of Service with in-vehicle sensor floating car data. This conclusion relates to, and is only valid for, two-lane motorways with a 120 km/h speed limit and the CAN bus signals which were available in the data set, since only that road type and those CAN bus signals where considered in this research.

If CAN bus data becomes available containing signals which can estimate LOS more accurately, such as distance headway measurements with radar, it is recommended to investigate the LOS estimation accuracy with these CAN bus signals. In that case, the research can be improved by using more accurate ground truth data, using other boundaries of the LOS categories, including other vehicle types and drivers, improving the model validity and by including other speed limits and road types. It is also recommended to investigate the CAN bus signal usability for other use cases. The wealth of information offered by the CAN bus signals offer possibilities for many applications.

Management samenvatting

Een groot aantal verkeersmanagementtoepassingen worden gebruikt om de verkeerssituatie en verkeersveiligheid te verbeteren. Deze toepassingen zijn afhankelijk van informatie over de optredende verkeerssituatie om goed te kunnen presteren. Om die informatie te verkrijgen worden verkeersdetectoren gebruikt. Er bestaan veel verschillende soorten detectoren die gebruikt kunnen worden om het verkeer te observeren. Deze detectoren blijken vaak onnauwkeurig en kostenintensief te zijn. In Nederland zijn inductielussen de meestgebruikte verkeerssensor.
Het Smart In-Car project geeft de mogelijkheid om een nieuw soort verkeersmonitoring te onderzoeken. In dit project worden de waarden van verschillende voertuig sensoren, verkregen via de CAN (Controller Area Network) bus, en de GPS posities verzameld met de uCAN on board unit van een klein aantal auto's, die met uCAN zijn uitgerust. In het project zin ongeveer 200 auto's voorzien van een uCAN on board unit. In dit onderzoek is slechts een deel van de data gebruikt, omdat de grote hoeveelheid data en de daarmee gepaard gaande rekentijd voor het verwerken van de data het gebruik van de complete dataset onmogelijk maakte in de beschikbare tijd. De gebruikte data bevatte de CAN bus data van 2 maanden voor 35 identieke auto's, deel van de vloot van een taxibedrijf. The voornaamste CAN bus signalen in deze dataset zijn: snelheidsmeter, remgebruik, stuurpositie, toerental, richtingaanwijzergebruik en brandstofverbruik.
In het Smart In-Car project zijn meerdere use cases gerelateerd aan schatting van de verkeersituatie, vlootmanagement, wegonderhoud, ongevalregistratie, verkeersmanagement en weer voorgesteld, die zouden kunnen profiteren van de sensordata van voertuigen.
In dit onderzoek zijn enkele verkeersmanagement gerelateerde use cases gedetailleerder geformuleerd en is een toepassing van de CAN bus data onderzocht. Het doel van dit onderzoek is om te bepalen of Level of Service (LOS), een kwalitatieve beschrijving van de verkeerssituatie, ingedeeld in categorieën van A tot F, gedefinieerd door de dichtheid van het verkeer, nauwkeurig geschat kan worden met de sensordata van auto's.

Dit is beoordeeld door de nauwkeurigheid van drie LOS schattingsmethoden te vergelijken

- LOS schatting met sensordata van auto's
- LOS schatting met floating car data
- LOS schatting met inductielus data

Sensordata van auto's is gedefinieerd als de CAN bus data die wordt verzameld door de uCAN on board unit. De uCAN on board unit levert ook de GPS positie van het voertuig. Wanneer LOS schatting met sensordata van auto's wordt genoemd, wordt de GPS data alleen gebruikt voor de plaatsbepaling van het voertuig en niet voor de LOS schatting.
Floating car data is gedefinieerd als enkel de positionele data (geleverd door de GPS unit van de uCAN on board unit) van de voertuigen.
Omdat het aantal auto's dat uitgerust is met de uCAN on board unit erg beperkt is, kan de echte CAN bus data alleen gebruikt worden om de LOS te schatten met een enkel voertuig. Om de LOS schatting met sensordata van meerdere auto's te beoordelen voor verschillende penetratiegraden, wordt verkeerssimulatiedata gebruikt, die voertuig trajectoriën en detectiepuntdata bevat.

Bij het gebruiken van de simulatiedata wordt niet alleen de detectiepuntdata gebruikt om de LOS referentiewaarde te bepalen, maar ook de werkelijke dichtheid, bepaald met de voertuigtrajectoriën.
De gesimuleerde CAN bus data wordt ook gebruikt om de LOS te schatten met een enkel voertuig, om de resultaten te kunnen vergelijken met de LOS schatting met de echte CAN bus data.
Het totaal aantal LOS schattingsmethoden dat gebruikt wordt in dit onderzoek komt daarmee op vijf:

- LOS schatting met echte CAN bus data van enkele voertuigen
- LOS schatting met gesimuleerde CAN bus data van enkele voertuigen
- LOS schatting met gesimuleerde CAN bus data van meerdere voertuigen
- LOS schatting met gesimuleerde floating car data van meerdere voertuigen
- LOS schatting met echte inductielusdata

Uit de echte CAN bus dataset worden de datapunten van voertuigen die een inductielus passeren gecombineerd met de data van de betreffende inductielus. De snelheids- en intensiteitsdata van de inductielus worden gebruikt om de LOS referentiewaarde op het moment en op de plaats van de passage van de met uCAN uitgeruste auto te bepalen. Deze gecombineerde data wordt gebruikt om de CAN bus signalen te identificeren die de sterkste correlatie met LOS hebben. Ook wordt regressieanalyse uitgevoerd om functies te bepalen die gebruikt worden om de LOS te schatten met de enkele voertuigpassages. Van de beschikbare CAN bus signalen, blijkt het gebruik van alleen de gemiddelde waarde van de snelheidsmeter de nauwkeurigste LOS schatting te geven. Hiermee wordt in 37% van de gevallen de juiste LOS geschat.
De simulatiedata wordt eerst gebruikt om, net als met de echte CAN bus data is gedaan, de nauwkeurigheid van de LOS schatting te bepalen met de gesimuleerde CAN bus data van passages van enkele voertuigen. In 52% van gevallen wordt de LOS correct geschat. Het verschil tussen beide correct LOS schattingspercentages heeft meerdere oorzaken en leidt tot de veronderstelling dat de simulatiedata gebruikt kan worden om een bovengrens te bepalen van de echte LOS schattingsnauwkeurigheid.
Ten tweede wordt de gesimuleerde CAN bus data gebruikt om de LOS te schatten met meerdere voertuigen voor verschillende penetratiegraden. De resultaten laten correcte LOS schattingpercentages zien tussen de 52% bij 1% penetratie tot 70% bij 100% penetratie.
Ten slotte is de simulatiedata ook gebruikt om de nauwkeurigheid van LOS schatting met de positionele data van meerdere voertuigen voor verschillende penetratiegraden. Dit is gedaan door de dichtheid van de gesimuleerde uCAN uitgeruste voertuigen te delen door de penetratiegraad. Hiermee wordt een geschatte dichtheid verkregen en kan de geschatte LOS worden bepaald. Met deze methode worden correcte LOS schattingspercentages gevonden tussen de 39% bij 1% penetratie en 100% bij 100% penetratie.
De resultaten tonen aan dat met sensordata van auto's in de meeste gevallen LOS minder nauwkeurig geschat kan worden dan met floating car data. Alleen in geval van penetratiegraden lager dan 8% zou LOS schatting met CAN bus data beter kunnen presteren dan met floating car data. Bovendien is het correcte LOS schattingspercentage met inductielusdata, geschat op 89%, veel hoger dan de nauwkeurigheid van de LOS schatting met CAN bus signalen.

Op basis van de resultaten wordt geconcludeerd dat het niet mogelijk is om Level of Service nauwkeurig te schatten met sensordata van auto's. Deze conclusie heeft betrekking op, en is slechts geldig voor, twee-strooks snelwegen met een maximumsnelheid van 120 km/h en de CAN bus signalen die beschikbaar waren in de dataset, omdat alleen dat type weg en die CAN bus signalen beschouwd zijn in dit onderzoek.

Als er CAN bus data beschikbaar komt die signalen bevatten die de LOS nauwkeuriger kunnen schatten, zoals volgafstandmetingen met radar, wordt het aanbevolen om de LOS schattingsnauwkeurigheid te onderzoeken met die CAN bus signalen. In dat geval kan het onderzoek worden verbeterd door nauwkeuriger LOS referentiewaarde data te gebruiken, andere grenzen tussen the LOS categorieën te gebruiken, andere voertuigklassen en bestuurders te gebruiken, het simulatiemodel te kalibreren en wegsoorten en maximumsnelheden te beschouwen. Een andere aanbeveling is, om de bruikbaarheid van de CAN bus signalen voor andere use cases te onderzoeken. De schat aan informatie die de CAN bus signalen leveren, bieden mogelijkheden voor een groot aantal verschillende toepassingen.

Table of contents

Abstract III

Preface V

Executive summary VII

Management samenvatting XI

1 Introduction 1

1.1 Traffic management applications 1
1.2 Traffic sensors 1
1.3 Smart In-Car 1
1.4 CAN bus 3
1.5 Use cases 4
1.6 Level of Service 6
1.7 Research objective 9
1.8 Research relevance 10
1.9 Report outline 10

2 Research approach 13

2.1 Level of Service and driving behaviour 13
2.2 CAN bus data processing 14
2.3 CAN bus data analysis 14
2.4 CAN bus data quality 14
2.5 LOS estimation with real CAN bus data 15
2.6 LOS estimation with simulation 19

3 Level of Service and driving behaviour 23

3.1 Speed 23
3.2 Lane change behaviour 23
3.3 Speed variation 24
3.4 Distance headway 25

4 CAN bus data 27

4.1 CAN bus signal frequencies and reliability 27
4.2 CAN bus data from Smart In-Car 28
4.3 CAN bus visualizations and quality 29

5 Results 31

5.1 LOS estimation with real CAN bus data 31
5.2 LOS estimation with simulation 45

6 Conclusions and recommendations 57
6.1 Conclusions 57
6.2 Discussion 59
6.3 Recommendations 63

References 67

Appendix A: Common traffic management applications 71

Appendix B: Common traffic data collection systems 75

Appendix C: Use cases 81

Appendix D: Traffic management use cases 85

Appendix E: CAN bus data 103

Appendix F: List of used loop detectors 131

Appendix G: Speed limits on Dutch motorways 135

Appendix H: Results: tables and figures 137

Appendix I: Digital appendices 153

1 Introduction

1.1 Traffic management applications

In the Netherlands, multiple traffic management applications, such as ramp metering, travel time prediction and dynamic speed limits are used in daily practice to improve traffic conditions and safety. A more extensive description of various traffic management applications can be found in Appendix A. These traffic management applications need information about the occurring traffic conditions to adequately function. This data can be provided by traffic data collection systems.

1.2 Traffic sensors

Common current methods used to measure traffic include induction loops, license plate recognition cameras, radar, laser and GSM. In the Netherlands, but also worldwide, induction loops are the most widely used traffic sensors [1]. Induction loops however can be unreliable and can be too inaccurate for some applications [2]. In recent studies, it was found that loop detectors in the Netherlands do often not meet the accuracy requirements of the NDW, the National Data Warehouse for Traffic Information, operating the most important traffic information database in the Netherlands (see Appendix B). Traffic intensity and speed are only measured where loops are available. A typical distance between loop detectors on Dutch motorways is 500 meters. For some applications, like incident detection, this resolution of the induction loops is too low to adequately function. Video based monitoring can give more detailed information but also has its limitations and drawbacks, like costs and low detection rate in case of bad weather and at night [3]. A more extensive description of various traffic sensors and their accuracy, found in literature, is provided in Appendix B.

New traffic data collection systems which are able to provide data more frequently and with greater accuracy could open up possibilities for new traffic management applications and improve existing ones.

1.3 Smart In-Car

A new traffic data collection system is tested in the Smart In-Car project. In the Smart In-Car project, data from the CAN bus (described in the next paragraph) of vehicles is gathered along with the GPS position of the vehicle and the acceleration of the vehicle in all three dimensions. Data from the CAN busses are collected and time stamped by the uCAN on-board unit, developed by Beijer Automotive. Depending on the vehicle type, signals like brake usage, windscreen wipers, air-conditioning, speed and steering position can be obtained from the CAN bus. An indicative list of

Table 1: indicative list of CAN bus signals [source: Beijer Automotive].

Accessory	Dimmer	Limiter	Speed10
Airconditioning	DirectionIndicator	LimiterOn	Speed100
AirconditioningTemperature	DoorsLocked	LowBeam	Speed15
Alarm	DoorsUnlocked	Neutral	Speed20
AlarmSet	Drive	NeutralGear	Speed25
AlarmSiren	DriveGear	NoLights	Speed2F
AnyDirectionIndicator	DriverDoor	Odometer	Speed40
AnyDoor	DriverSeatbelt	OdometerHR	Speed5
AnyFogLight	DriverVentilation	OperateFork	Speed50
AnyGear	EngineRun	ParkingAssistant	Speed8
AnyRearDoor	EngineStart	ParkingBrake	StartStop
AnyRemoteKey	FirstGear	ParkingGear	SteeringAngle45
AutoLights	Forward	ParkingLight	SteeringLeft
Brake	FrontDoors	ParkingLightLeft	SteeringPosition
BrakeForce	FrontFogLight	ParkingLightRight	SteeringRight
BrakeGear	FuelConsumption	ParkingRadar	Test100
BrakeHard	FuelLevel	PassengerDoor	Test200
BrakePedalPosition	FuelLid	PassengerSeatbelt	Test300
BrakePosition	FuelLow	Periodic	Test400
BrakeRear	GearShiftLeverSwitch	PTOState	Throttle
BrakeSoft	HazardLights	Radio	TractionControl
Broadcast1	HazardLights_Switch	RDI	Trunk
ButtonMicrophone	Heating	RDI_Active	UnlockAllDoors
ButtonMute	HighBeam	RDI_Pulse	UnlockDoors
ButtonPhone	HighBeamFlash	RDI_Switch	UnlockSlideDoor
ButtonPhoneEndCall	HighBeamLight	RearDoors	VehicleLocked
ButtonVolumeDown	Hood	RearFogLight	VehicleOpen
ButtonVolumeUp	Horn	Remote	VehicleUnlocked
CancelButton	HydroMotor	RemoteAnyKey	VIN
CarUnlocked	Ignition	RemoteClose	Wiper
Clutch	InjectionTime	RemoteLock	WiperActive
ClutchPosition	KeyIn	RemoteOpen	WiperFast
CruiseActive	KeyOut	RemoteTrunk	WiperInactive
CruiseCANButton	KeyOutEngineRun	RemoteTrunkIdle	WiperInterval
CruiseControl	KeyPresent	ResButton	WiperIntervalPosition
CruiseDecreaseButton	KickDown	ReverseGear	WiperIntervalSpeed
CruiseMin	LDI	RightRearDoor	WiperMotor
CruiseOff	LDI_Active	RPM	WiperOn
CruiseOffButton	LDI_Pulse	SeatContact	WiperOnce
CruisePlus	LDI_Switch	SetButton	WiperPosition1
CruiseResButton	LeftRearDoor	SlideDoor	WiperReverse
CruiseSetButton	Lights	SnowSwitch	WiperSlow
DashboardDimmer	LightsAuto	Speed	

the CAN bus signals can be found in Table 1. During the test, the data that has been gathered by the uCAN on-board unit is wirelessly transmitted and stored in a central database server, operated by TASS Technology Solutions.

Around 200 uCAN equipped vehicles participate in the Smart In-Car test. Most of these cars are part of the Cibatax fleet. Cibatax is a taxi company located in Eindhoven. The other equipped vehicles are part of the fleet of the ANWB, the Dutch emergency roadside assistance organisation. The data collection of the Smart In-Car test started in March 2012.

Combined with the GPS position, the CAN bus data may provide insight into many vehicle, driver, road and traffic characteristics. There appear to be many opportunities to use CAN bus data for traffic management. For example, an abrupt steering maneuver may indicate an obstacle on the road, windscreen wiper usage can indicate adverse weather (e.g. rain) and activation of ABS might be an indication for a slippery road surface. This information can be used to warn drivers that are upstream of the incident location and to alarm the road authorities. If the CAN bus data proves to be suitable for providing this sort of information, it can contribute to better traffic management and safer traffic.

1.4 CAN bus

CAN bus was developed in the 1980’s by Robert Bosch GmbH. The constantly growing number of devices in vehicles led to a growing number of necessary connections between these devices. Instead of using separate links between these components, CAN bus interconnects several components at once by connecting them to the same network, thereby reducing the amount of wiring and costs. Additionally, the CAN bus protocol is able to detect errors in the data transmission, for example caused by electromagnetic interference. This is corrected by resending the message.

For over 10 years, CAN bus has been a mandatory protocol for vehicles in the United States and the European Union. Besides the automotive industry, other fields of application of CAN bus include industrial automation, medical equipment and domestic appliances.

The CAN (controller area network) bus is a specialized internal communication network that interconnects components inside a vehicle. It is a vehicle bus standard that allows microcontrollers and devices in the vehicle to communicate with each other. Typically a vehicle has several CAN busses. Each of these connects components, called CAN nodes, which depend on each other. An example is illustrated in Figure 1. In this figure, the traction control unit is connected with the same CAN bus as the engine ignition controller, the fuel injection controller and the throttle controller. These CAN nodes are able to transmit data over the network. Messages that are sent by one of the CAN nodes are sent to all other CAN nodes in the network. Each CAN node uses the message identifier to decide whether it should process the message or not. When multiple CAN nodes are trying to send messages simultaneously, the message identifier is also used to determine the priority of every message. CAN messages can contain up to 8 bytes of information.

More information can be send by sending several messages. For networks up to 40m, the maximum transmission rate is 1Mbit/s. [4][5][6]

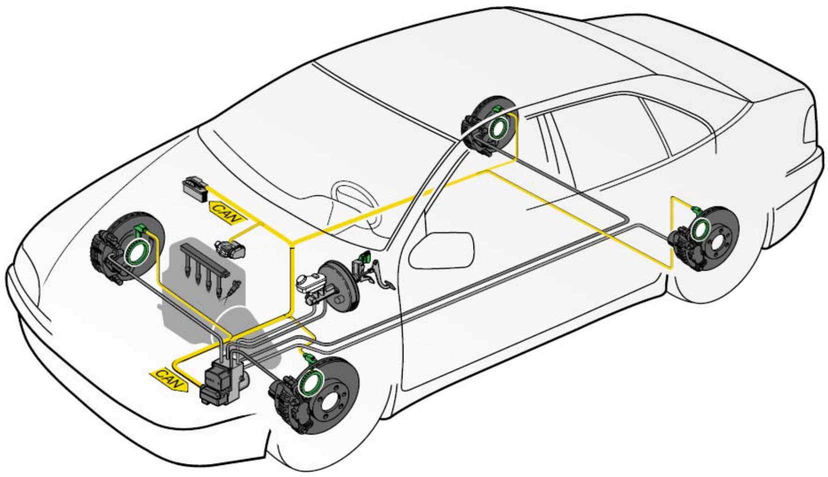

Figure 1: Traction control CAN bus [4]

1.5 Use cases

The data that is provided by the CAN bus and the GPS unit of the uCAN on board unit may have a large range of applications. Several brain storm sessions of the Smart In-Car partners have resulted in a large amount of use cases. Besides traffic state estimation related use cases, also use cases related to fleet management, road maintenance, accident registration, traffic management and weather have been suggested. An extensive list of the use cases is shown in Appendix C.
Recently, four of these use cases were explored by LaQuSo [7], one of the partners in the Smart In-Car project, including traffic jam head and tail detection, rain detection, driver behavior analysis for fleet monitoring purposes and road abnormalities detection (potholes, speed bumps and rough roads). Useful results were found for all use cases, demonstrating a lot of potential for the Smart In-Car technology. In other work [8], a simulation model was used to show a high feasibility of accurate road profile classification using the data provided by the accelerometer in the uCAN on board unit.

Several traffic management related use cases, describing detectable and measurable events and quantities which can be used by traffic management applications, have been defined in more detail. Below, these use cases are listed. Their descriptions, including the measurement and test method, can be found in Appendix D.

- Speed (for Dynamax, traffic jam detection, ramp metering, peak lanes)
- Travel times (for travel information)
- Rain (for Dynamax)
- Wind (for VMS warning)

- Fog (for VMS warning)
- Obstacle detection (lane closure, peak lanes)
- Intensity (for Dynamax, ramp metering, peak lanes)

In this study, a traffic management application of the CAN bus data will be investigated. Below, the use cases are discussed to identify a traffic quantity or event which is suitable to be tested with the available data from the 200 equipped vehicles. Because just 200 vehicles are equipped with a uCAN on board unit, it will be difficult to test the traffic management use cases. Another requirement is that the use case is expected to significantly benefit from the CAN bus signals and does not only need the GPS data from the uCAN.

Some use cases are very useful, but their suitability seems to be trivial. For example, since both the GPS position (and therefore the GPS speed) and CAN bus speed are known, it is trivial that the speed can be measured with the CAN bus data, although the penetration rate will have to be high to give frequent enough measurements for every location and to give a speed measurement which corresponds accurately enough with the average speed of the whole traffic stream. It is decided not to investigate this use case, since the added benefit of the CAN bus data is expected to be small in respect to the GPS data.

The same line of reasoning is followed for the decision to not investigate the applicability of CAN bus data for the measurement of realized travel times. Travel times can be determined by only using the GPS data. The CAN bus data is not expected to contribute to better travel time measurement.

It would be interesting to investigate the rain detection possibilities of CAN bus data. It seems quite trivial that precipitation can be measured when the usage of the windscreen wipers is known. In fact, the reference data, which would be used to confirm the suitability of the CAN bus data for rain detection, is likely to be less accurate than the CAN bus windscreen wiper data. When vehicles in an area are using their windscreen wipers and reference data from weather services claim there is no precipitation at that location, it is much more likely the weather service data is wrong than the CAN bus data, since it is unlikely that many vehicles use their wind screen wipers when it is not raining. Beside traffic management applications, also meteorological services might be interested in this kind of detailed information. A large amount of uCAN equipped vehicles can form a dense network of precipitation detectors. Because the amount of equipped vehicles is small and the possibility of rain detection has already been demonstrated in [7], also this use case is dismissed to examine here.

Like rain, also fog is likely to be suitable to be indicated with CAN bus signals. When the usage of the fog lights of a large amount of vehicles is known, the presence of fog on roads is likely to be able to be determined very accurately. The small amount of uCAN equipped vehicles in combination with the rare occurrence of fog leads to the decision not to investigate this use case.

Similarly to rain and fog detection, when the CAN bus signals would prove to be able to measure wind direction and speed, the CAN bus data may be used to indicate the wind conditions on all roads. Wind measurements are used in traffic management to warn for dangerous wind conditions. The wind is measured locally at locations where strong wind is likely to occur (e.g. bridges). It is questionable what the gain is of detailed knowledge of wind conditions on other locations. For the rain and wind use case, there are CAN bus signals (fog lights and wind screen wipers) which correspond directly to the measurable event. This is however not the case for wind. Some CAN bus signals are expected to show some dependency on the wind speed and direction, like speed, fuel consumption and steering position, but the dependency is expected to be small and therefore, also the accuracy of the wind measurements is likely to be small. This leads, along with the relative small practical relevance of wind detection for traffic management, to the dismissal of the examination of the wind use case.

Also the obstacle detection use case is very interesting. Unfortunately, the current penetration rate of uCAN equipped vehicles and the rare occurrence of obstacles do not allow us to test the obstacle use case in this stage of the Smart In-Car test.

The final use case considered is intensity. This quantity is very relevant for traffic management and the dataset provides enough CAN bus parameters which might be able to show a correlation with the intensity and can thus be used to estimate the intensity. However, the penetration rate of uCAN equipped vehicles is very low and the intensity estimations will have to be made with a single vehicle passage. It is expected that the CAN bus signals of single vehicles will yield intensity estimations with too little accuracy too be practically useful, while for higher penetration rates, the intensity estimations could prove to be accurate enough. Therefore, another traffic condition indicator, Level of Service (LOS), is considered which is expected to be more likely to be able to give reasonable accuracy using CAN bus data from single vehicles. It is decided to investigate the possibility to estimate Level of Service, a categorical qualitative description ranging from A to F of the traffic conditions on a road, with CAN bus data. Since Level of Service is expressed in categories, it is more probable that Level of Service can be estimated with reasonable accuracy than the intensity. In the next paragraph, a description and definition of Level of Service is provided.

1.6 Level of Service

1.6.1 Description

Level of Service is a term used to qualitatively describe the operating conditions and effectiveness of elements of transportation infrastructure. Level of Service qualifications are based on one or several traffic quantities, such as average speed, density and intensity. Level of Service was introduced in the 1965 Highway Capacity Manual. The Highway Capacity Manual defines Level of Service as:

"Level of Service (LOS) is a quality measure describing operational conditions within a traffic stream, generally in terms of such service measures as speed and travel time, freedom to maneuver, traffic interruptions, and comfort and convenience."

There are six levels, ranging from A to F. LOS A is defined as free flow and LOS F reflects breakdowns in vehicular traffic. Table 2 shows the LOS definition from the 2000 Highway Capacity

Manual. Table 3 shows the traffic quantities for every category. These are also graphically presented in Table 4.

Table 2: Level of Service category definitions [9]

Level of Service category	Definition
A	Free-flow operations. Vehicles are almost completely unimpeded in their ability to maneuver within the traffic stream.
B	Reasonably free-flow. Free-flow speeds are maintained. The ability to maneuver within the traffic stream is only slightly restricted, and the general level of physical and psychological comfort provided to drivers is still high.
C	Flow with speeds at or near free-flow speed of the freeway. Freedom to maneuver within the traffic stream is noticeably restricted, and lane changes require more care and vigilance on the part of the driver.
D	Speeds begin to decline slightly with increasing flows and density begins to increase somewhat more quickly. Freedom to maneuver within the traffic stream is noticeably limited, and the driver experiences reduced physical and psychological comfort levels.
E	Operation at capacity. Vehicles are closely spaced. Maneuverability within the traffic stream is extremely limited, and the level of physical and psychological comfort afforded the driver is poor.
F	Breakdowns in vehicular flow.

Table 3: HCM Level of Service criteria for basic freeway sections [10]

1994 HCM Level of Service Criteria for Basic Freeway Sections

	70 mph Free-Flow Speed				65 mph Free-Flow Speed				60 mph Free-Flow Speed			
LOS	Density[a] (pc/mi/ln)	Speed[b] (mph)	Maximum[c] V/C	MSF[d] (pcphpl)	Density[a] (pc/mi/ln)	Speed[b] (mph)	Maximum[c] V/C	MSF[d] (pcphpl)	Density[a] (pc/mi/ln)	Speed[b] (mph)	Maximum[c] V/C	MSF[d] (pcphpl)
A	≤ 10.0	≥ 70.0	0.318/0.304	700	≤ 10.0	≥ 65.0	0.295/0.283	650	≤ 10.0	60.0	0.272/0.261	600
B	≤ 16.0	≥ 70.0	0.509/0.487	1,120	≤ 16.0	≥ 65.0	0.473/0.457	1,040	≤ 16.0	60.0	0.436/0.412	960
C	≤ 24.0	≥ 68.5	0.747/0.715	1,644	≤ 24.0	≥ 64.5	0.704/0.673	1,548	≤ 24.0	60.0	0.655/0.626	1,440
D	≤ 32.0	≥ 63.0	0.916/0.876	2,015	≤ 32.0	≥ 61.0	0.887/0.849	1,952	≤ 32.0	57.0	0.829/0.793	1,824
E	≤ 36.7/39.7	≥ 60.0/58.0	1.000	2,200/2,300	≤ 39.3/43.4	≥ 56.0/53.0	1.000	2,200/2,300	≤ 41.5/46.0	53.0/50.0	1.000	2,200/2,300
F	Variable	Variable	Variable	Variable	Variable	Variable	Variable	Variable	Variable	Variable	Variable	Variable

[a] Density in passenger cars per mile per lane.
[b] Average travel speed in miles per hour.
[c] Maximum volume-to-capacity ratio.
[d] Maximum service flow rate under ideal conditions in passenger cars per hour per lane.

≤ less than or equal to
≥ greater than or equal to

Note: In table entries with split values, the first value is for four-lane freeways, and the second is for six- and eight-lane freeways.

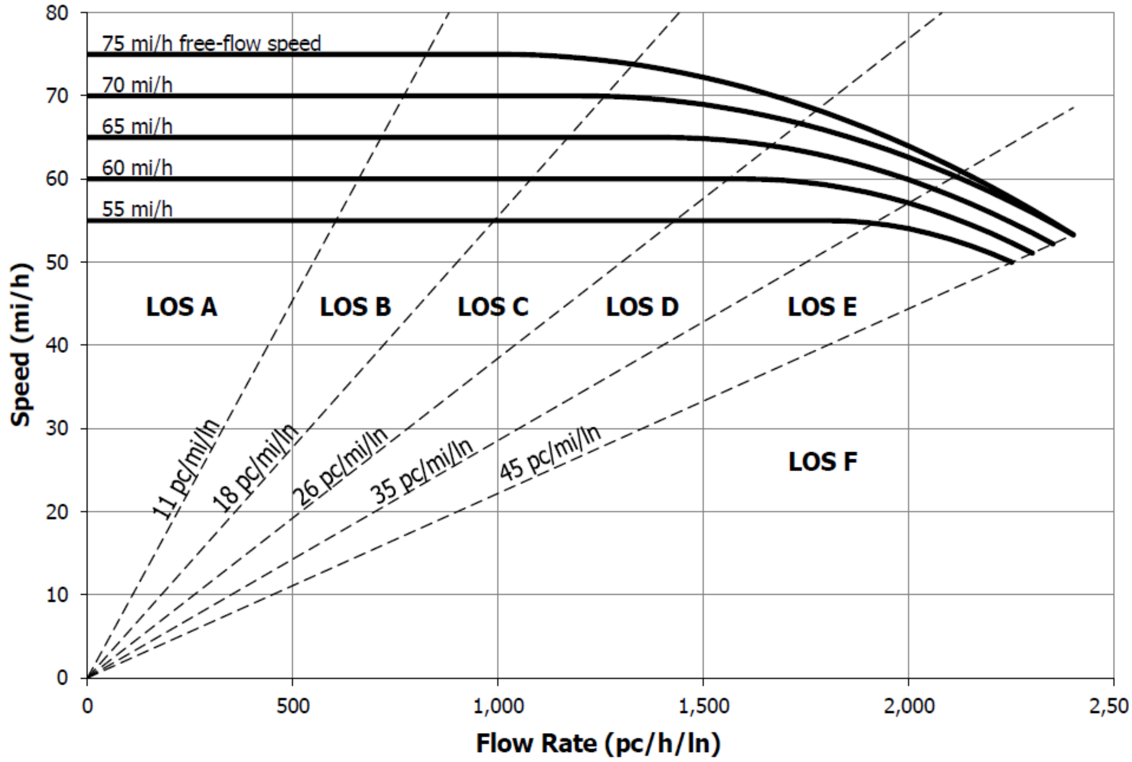

Figure 2: Level of Service graph [11]

The number of levels and the boundaries between the levels are quite arbitrary. After the first definition of the Level of Service, in the Highway Capacity Manual 1965, several suggestions have been made to change the amount of levels, to change the boundaries between the levels or to change the quantities that are used to measure the LOS.

Choocharukul et al. [12] investigated the user perception of highway Level of Service. They concluded that LOS F does not adequately represent the conditions that are perceived by the road users as the ‟breakdown in vehicle flow'' that this LOS is intended to represent. They also showed that there are a number of attributes besides traffic density which determine public perceptions of LOS.

Brilon [13] and Cameron [14] both point out that different traffic qualities exist within LOS F. Cameron proposes to use additional levels beyond LOS F (G, H and I) to properly quantify congestion conditions. Brilon proposes to use a more objective parameter to determine the thresholds between subsequent LOS. He introduces efficiency, as a combination of demand, capacity and quality of flow, which should be able to determine the boundaries between the LOS, especially for LOS D and E.

Level of Service can be used for traffic information, but also for other traffic management applications. Landman et al. [15] have shown that route guidance control based on Level of Service is able to improve network performance, while taking into account policy objectives and route priorities of road authorities.

1.6.2 Definition

In this research, the LOS definition of the 2010 HCM will be used. The metric classification values are shown in Table 4.

Table 4: Level of Service boundaries density values.

LOS	max. density (pc/mile/lane)	max. density (pc/km/lane)
A	11	6,8
B	18	11,2
C	26	16,2
D	35	21,7
E	45	28,0
F		

1.7 Research objective

The objective of this master thesis is: *"to assess the possible applicability of real-time wireless transmitted CAN bus data, combined with GPS position, for estimating Level of Service at motorways."*

1.7.1 Research questions

In this master thesis, the following main research question will be answered:

> Is it possible to accurately estimate the Level of Service with in-vehicle sensor floating car data?

To answer this main question, a set of sub-questions has to be answered:

1. Which main driver behavior characteristics are dependent on Level of Service?
2. Which CAN bus signals are provided by the uCAN equipped vehicles?
3. What is the accuracy of Level of Service estimation with in-vehicle sensor floating car data from single vehicles?
4. What is the accuracy of Level of Service estimation with in-vehicle sensor floating car data from multiple vehicles?
5. Does in-vehicle sensor floating car data provide a better LOS estimation than plain floating car data?
6. Does CAN bus data provide a better LOS estimation than loop detector data?
7. Does the transmission frequency of the CAN bus signals influence the LOS estimation accuracy?

Some of the terms in the sub questions need clarification. In-vehicle sensor floating car data is defined as the CAN bus data which is collected with the uCAN on board unit. The uCAN on board

unit also provides the GPS positions of the equipped vehicles. When LOS estimation with the CAN bus data is considered, the GPS data is only used to determine the location of the vehicles, but not to estimate the LOS.
Plain floating car data is defined as exclusively the positional data (provided by the GPS unit of the uCAN on board unit) of the equipped vehicles.

1.8 Research relevance

This research contributes to the practice of Rijkswaterstaat. It explores the applicability of CAN bus signals for the estimation of Level of Service. It also gives insight in the possibility of density estimation, intensity estimation and speed estimation with CAN bus signals. Especially intensity and speed are important quantities which are used by many traffic management applications used by Rijkswaterstaat. This research is a first step in the investigation of applicability of CAN bus signals for traffic management. Besides Level of Service estimation, several use cases are defined which could be suitable to be measured with CAN bus signals. This new kind of traffic sensor could prove to be more accurate and cost effective than existing traffic measurement systems, thereby not only reducing costs and effectiveness of existing traffic management applications, but also providing opportunities for applications which are currently not used, because the required data to adequately function is not available, like obstacle detection.
Another aspect of this research is that it provides insight in the properties of the signals which can be provided by the CAN bus. This knowledge can be used for the investigation of applicability of CAN bus signals for the proposed traffic management use cases or the other use cases relevant for Rijkswaterstaat, like measuring road surface conditions for road maintenance purposes.
The CAN bus signals may also be used for research on driving behaviour. CAN bus signals, which can be used to measure parts of the driving behaviour, are ideal to measure driving behaviour in real driving conditions. Better understanding of the driving behaviour in various traffic conditions can contribute to better traffic modelling.

1.9 Report outline

1.9.1 Research approach

In chapter 2, the phases are described that are executed to accomplish the research objective. First, literature is reviewed to find driving behavior characteristics which show a dependency on Level of Service. Then code is written to import, process and analyze an initial CAN bus data set. With the findings, a part of the uCAN equipped vehicles is selected, which supply CAN bus signals which are most likely to be able to measure the driving behavior characteristics which were found to depend on LOS.
The CAN bus data of these selected vehicles is combined with data from loop detectors on motorways in the Eindhoven area, which are used to determine the LOS ground truth. With the

combined data, the CAN bus signal correlations with LOS are determined and a LOS estimation function is defined which uses the CAN bus signals of a single uCAN equipped vehicle to estimate the LOS.
Using a traffic simulation model, LOS estimation with CAN bus signals from multiple vehicles is investigated. The model is also used to estimate LOS using only the positional data (simulating GPS). The results of the real CAN bus data LOS estimation, the simulated CAN bus data LOS estimation and the simulated positional data LOS estimation are compared and discussed.

1.9.2 Level of Service

In chapter 3, literature is reviewed to identify the main driving behaviour characteristics which show dependency on LOS. Sub question 1 is answered in this paragraph.

1.9.3 CAN bus data

In chapter 4 the process of importing, processing and analysing the CAN bus data from the Smart In-Car project is described. Visualizations are presented which are used to interpret the data and give a first impression of the available data. Frequency, accuracy and several other properties of the CAN bus data are discussed. Sub question 2 is answered in this paragraph.

1.9.4 Results

Chapter 5 contains the results of the phases which are described in chapter 2.
In paragraph 5.1 the LOS estimation accuracy with real CAN bus data is determined. Based on the findings in the previous section, data from a part of the vehicle fleet is selected, containing those CAN bus signals which are most likely to be suitable to estimate LOS. The data from 35 equipped taxis over a period of two months is collected. Also loop detectors, located near Eindhoven where the taxis operate, are selected to provide data to determine the LOS ground truth. The data from the loop detectors and the uCAN equipped vehicles is combined and used to show the dependencies between LOS and several CAN bus parameters. The effect of various radii (the distance from the loop detector that defines the area in which data points are collected) and signal periods (the CAN bus signal transmission periods) is discussed and the LOS is estimated using single and multiple CAN bus parameters. It is found that using a single CAN bus signals, the average speed, yields the most accurate LOS estimation. LOS estimation using the real CAN bus speed from single vehicle passages is correct in just 37% of the cases. Sub question 3 and 7 are answered in this paragraph.
In paragraph 5.2 the LOS estimation accuracy with simulated CAN bus data is determined. This is done with single vehicle passages, to allow a comparison with the results of the LOS estimation with real CAN bus data, and with multiple vehicle passages, for several penetration rates.
This is done using several LOS ground truths, based on detection point measurements and the actual density. Also, the LOS is estimated using only the vehicle positions, simulating plain floating

car data, for example from GPS units. This paragraph is concluded with a comparison between both LOS estimation methods. For high penetration rates, the correct LOS estimation rate using the simulated CAN bus data reaches 70%, while the correct LOS estimation rate using only the vehicle positions, simulating GPS data, reaches 100%. Sub question 4, 5 and 6 are answered in this paragraph.

1.9.5 Conclusions and recommendations

In chapter 6 the answers to the main research question and the sub questions are presented and discussed, resulting in several recommendations.

2 Research approach

This chapter describes the phases that are executed to accomplish the research objective, which was formulated in the previous chapter as: *"to assess the possible applicability of real-time wireless transmitted CAN bus data, combined with GPS position, for estimating Level of Service at motorways."* The methods that are used to answer the research questions are described. Also, the data which are used in the analysis is discussed. The research approach is schematically shown in Figure 3. The various elements and their relations will be discussed in this chapter.

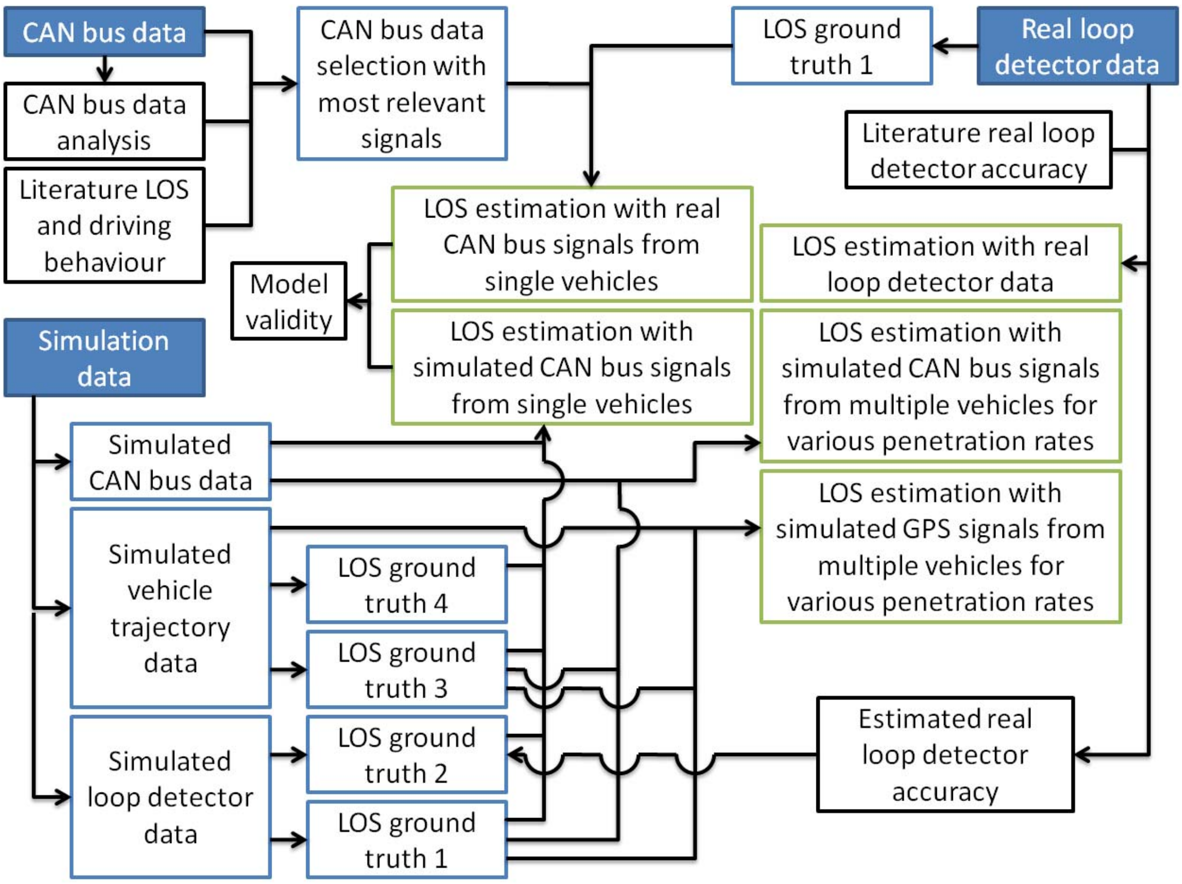

Figure 3: Schematic research approach

2.1 Level of Service and driving behaviour

In this phase, literature is reviewed to examine if there are certain driving behaviour characteristics which depend on the Level of Service. Since the in-vehicle sensors are able to capture part of the driving behaviour, this knowledge will be used to select the vehicles that produce sensor data from those sensors which are expected to have the largest probability of showing a correlation between the LOS and the values of the sensor signals. Not all data from all vehicles is used, since processing all data would require too much calculation time (several months).

2.2 CAN bus data processing

After selecting the vehicles which will be used for the analysis, the CAN bus data from those vehicles will have to be obtained and processed. Since the data is distributed in text files in json format (JavaScript Object Notation, a text-based open standard, see Figure 11 for an example), the data will first be imported to Matlab and converted to a matrix format. The goal of this phase is to write Matlab code which is able to import and convert the raw data files to Matlab matrix files, which can be easily and quickly accessed, queried, edited and processed.
The final Matlab matrices will have a single row per vehicle per time instant, showing, in separate columns: the time, the date, the vehicle id, the GPS position and all CAN bus data available in that time instant.

2.3 CAN bus data analysis

When the data set has been imported and converted, it will be explored. The goal of this phase is to obtain knowledge about the available information and its characteristics in the data set.
For starters, the locations of the vehicles will be looked into. The results will be used to determine which and how many loop detectors will be used to obtain the ground truth data, described in section 2.5.1 Also the available sensors and their frequencies will be investigated. Insight into the CAN bus data will be obtained as well by visualizing the data in various ways. For example by plotting several CAN bus data values versus each other or by plotting the CAN bus data versus the position, given by the GPS position.

2.4 CAN bus data quality

Before using the CAN bus data for LOS estimation, we need to gain insight in the accuracy and reliability of these CAN bus signals. This will be done by reviewing literature on the CAN bus and by comparing the values from the several CAN bus signals with each other. For example, the accuracy of the CAN bus speedometer signal can be compared with the speed, calculated from the GPS positions. The visualizations, described in the previous paragraph, can also give insight in the accuracy and reliability of the CAN bus signals.

2.5 LOS estimation with real CAN bus data

In this paragraph, the steps that will be followed to accomplish a LOS estimation with real CAN bus signals are described.

2.5.1 Determining the LOS ground truth

To investigate whether the CAN bus data can accurately estimate the Level of Service, reference data is needed to determine the LOS ground truth at the position and time of the uCAN equipped vehicle presence. The LOS ground truth will be determined by using the speed and intensity data of the loop detectors in the road. This data is aggregated per minute. Using this information, the LOS will be determined for every minute, according to the definition in section 1.6.2. Not all roads are equipped with loop detectors, so the research will have to be restricted to those roads that are equipped with loop detectors. The map in Figure 4 shows the locations of the loop detectors in the Eindhoven area. All loop detectors are located on motorways, so no loop detector data of secondary roads is available.

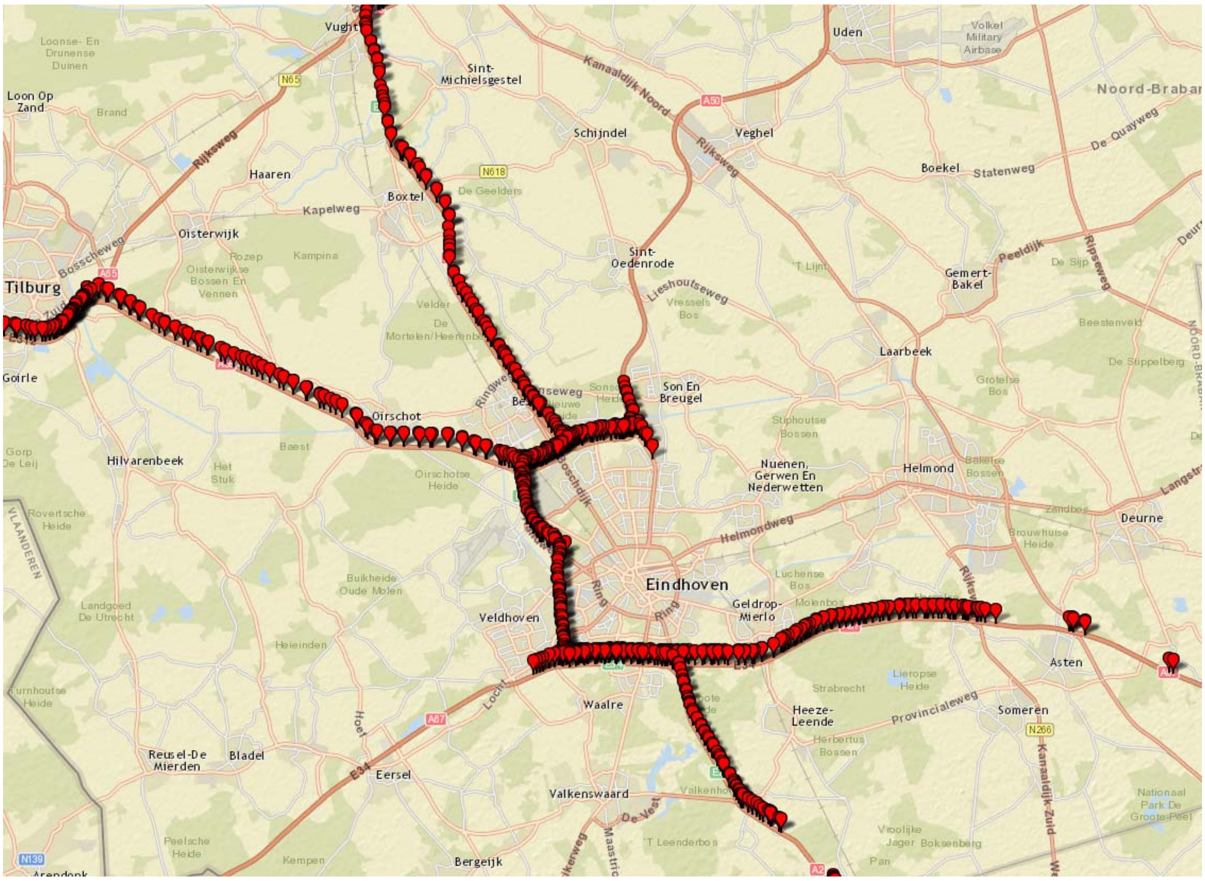

Figure 4: Detection loops in the Eindhoven area. The red icons indicate the locations of the loop detectors.

Only 35 uCAN equipped cars are included in this study. These vehicles are taxis, which are expected to primarily make short trips in the city or to the airport. Therefore, the number of uCAN equipped vehicle passages in the available dataset on the motorways near Eindhoven is expected to be very limited. It will be found that on average the detection points on the motorways in the Eindhoven area only passed once per day by one of the 35 selected vehicles. Because of this, only road sections that have a similar cross section will be included in this research. The most common cross section in the Eindhoven area is a 2*2-lane motorway, with a shoulder lane (see Figure 5). The data from these road sections will be combined to get a larger amount of similar measurements. Road sections with on-ramps, off-ramps and weaving sections will not be part of this research, to ensure all measurement are made in similar road conditions.

Figure 5: 2*2-lane motorway

2.5.2 Acquiring the relevant CAN bus data

To obtain the CAN bus data from the uCAN equipped vehicles that drive on road sections like in Figure 5, the data will have to be filtered. Since the LOS is determined from the data of a single point on the road and the CAN bus data is collected all over the road section, a part of the road section has to be selected where the CAN bus data is collected. This will be done for various road section lengths, in the ensuing defined by the radius: the distance from the loop detector that defines the area in which CAN bus data points are collected. With the findings resulting from the following steps an optimal road section length will be determined which will be used in the rest of the research. The optimal radius is expected to be around 200m, because it is assumed to be a reasonable concession between minimizing the length to assure a good comparison with the loop detector data and maximizing the length to obtain enough CAN bus data points from the passing vehicles. At 130 km/h, a vehicle will provide a CAN bus data point every 36 meters, since it will be found that the CAN bus signals are transmitted around once per minute. With a road section length of 400m, at least 10 data points will be collected for each passing vehicle. Initially, radius values of 50m, 100m, 200m and 400m will be considered. Additional radius values will be added if this proves to be necessary to acquire the optimal radius.

When the data is filtered by the GPS position, there will be some data points, coming from vehicles on roads that over- or underpass the motorway or from vehicles on parallel roads, which will contaminate the data. So instead of just filtering by GPS position, virtual checkpoints will be used to assure that we only get the right data points. At the start and end of each road section a virtual checkpoint is created. When a vehicle passes the checkpoint at the start of the road section, all data between that moment and the moment when the vehicle passes the checkpoint at the end of the road section will be saved in a separate file and used for the further filtering.

In Figure 6, the CAN bus data filtering is visualized. From the CAN bus data points between both checkpoints, the ones which are within the radius of one of the loop detectors are stored. Notice that the area defined by the radius is large and also includes roads other than the motorway. This does not cause problems, since the checkpoints will make sure that only the data from the motorway and from vehicles driving in the right direction will get selected. The size of the checkpoint area has to be large enough, because the positional data is only given once per second. The path through the checkpoint area has to be longer than the distance a fast vehicle covers in a second (50 meters at 180 km/h), therefore as checkpoint radius of 30m is used.

The data points of every vehicle passage will be processed, to get a single line of data from every vehicle passage which can be combined with the single line of loop detector data. This line will contain the average value and the standard deviation of the CAN bus parameters, as well as the minimum and maximum value.

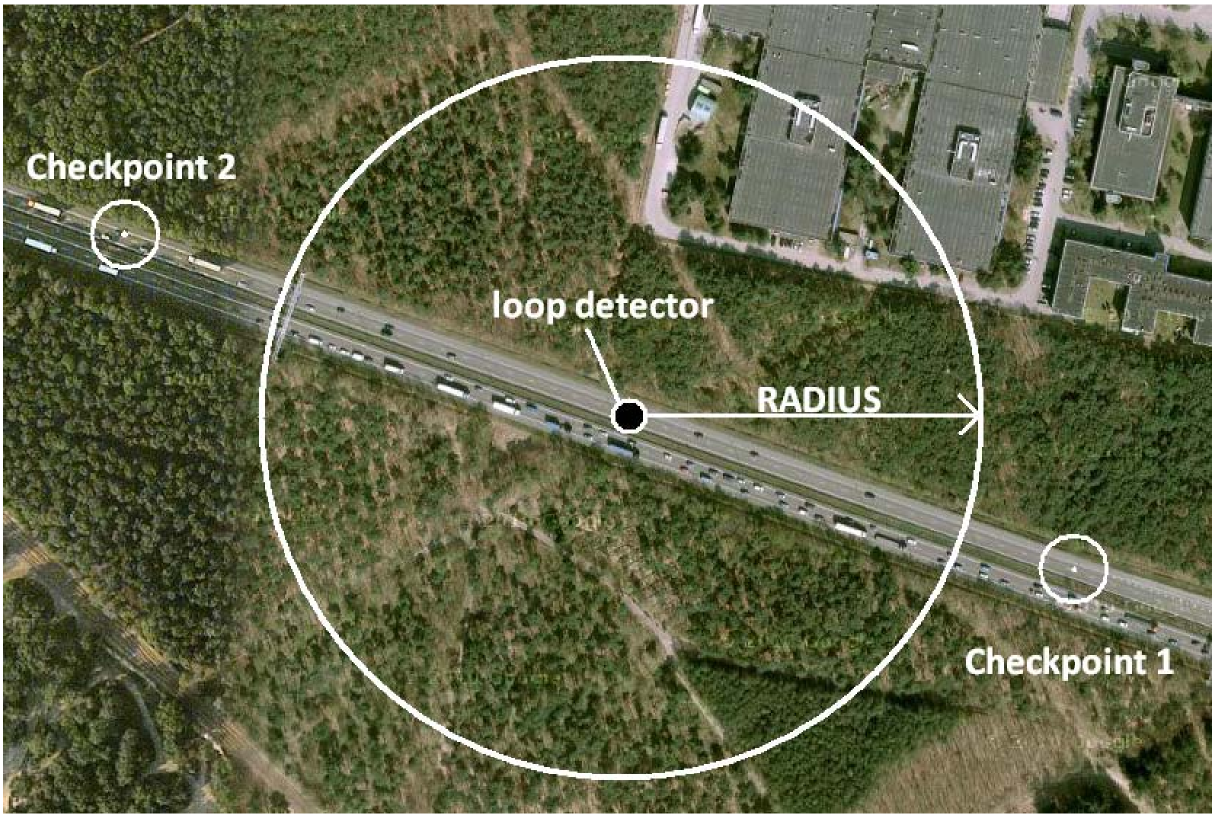

Figure 6: CAN bus data filtering. uCAN equipped vehicles are checked for subsequently passing checkpoint 1 and checkpoint 2 with their GPS data. When they do, it is certain they are driving on the motorway and in the correct direction. From the CAN bus data points between both checkpoints, the ones which are within the radius of one of the loop detectors (just one in the figure) are stored and can be used to combine with the loop detector data.

2.5.3 Combining the loop detector data with the CAN bus data

The single CAN bus data line will be combined with the loop detector data by matching the CAN bus timestamp and the loop detector timestamp. This is visualized in Figure 7. After combining the data, we get a table including all loop detector data, the CAN bus parameters and the LOS. The contents of this table are listed below.

- Date
- Time
- VehicleID
- Road
- direction
- longitudinal position
- 1-minute aggregated intensity from loop detector
- 1-minute aggregated speed from loop detector
- LOS
- For every CAN bus parameter: the average value during the passage
- For every CAN bus parameter: the standard deviation during the passage
- For every CAN bus parameter: the minimum value during the passage
- For every CAN bus parameter: the maximum value during the passage

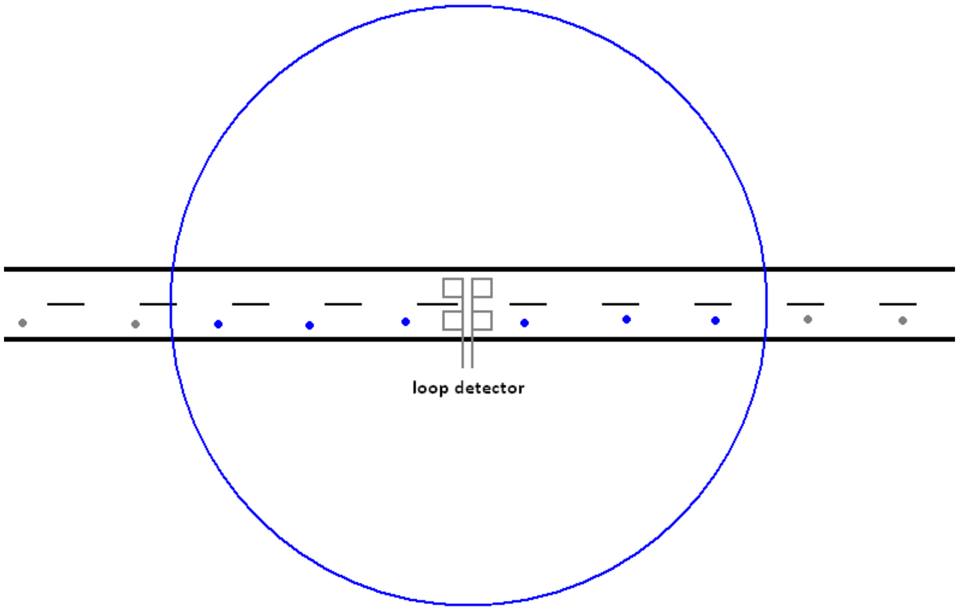

Figure 7: Combining the CAN bus data with the loop detector data. The dots indicate the GPS coordinates of the data points of a uCAN equipped vehicle. The blue circle, determined by the loop detector position and the radius, indicates the area in which the CAN bus data points are collected.

2.5.4 Correlation and regression analysis

With the combined data from the CAN bus and the loop detectors, first the relationship between the individual CAN bus parameters will be visualized by plotting the measured CAN bus parameter versus the measured Level of Service. Subsequently, the correlations between the CAN bus parameters and the LOS will be determined. Regression analysis will be used to determine the combination of CAN bus parameters which give the best LOS estimation. Since the LOS is in an ordinal scale, multinomial logistic regression will be used. With the results of the regression analysis, LOS estimation functions will be constructed of the form:

$$P(LOS=y) = e^{(\beta y * Xi)} / \Sigma_k\, e^{(\beta k * Xi)}$$

where: P(LOS=y) is the chance that the Level of Service is y

βy is the regression coefficient for LOS=y

Xi is the value of the CAN bus parameter

βk are the regression coefficients for the 5 other LOS (LOS≠y)

This will be done by using the *mnrfit* function in Matlab. This function uses Maximum Likelyhood Estimation to determine β's which yield the best fit.

With this resulting function, the LOS can be estimated with a single vehicle passage, using the *mnrval* function in Matlab. Using the β's, determined with *mnrfit*, and the values of X_i this function calculates for every LOS the probability that this LOS occurs. The LOS with the highest probability will be selected as the estimated LOS. This will be done with the available dataset. The results of these estimations will be shown and discussed.

2.6 LOS estimation with simulation

Since the number of vehicles which are equipped with the uCAN on board unit is very limited, it is not possible to investigate the LOS estimation accuracy at higher penetration rates. Therefore, data from traffic simulation software is used to simulate LOS estimation at higher penetration rates.

For the simulation study simulation software package OpenTraffic [16] will be used. This is an open source simulation package developed in Java. For the simulation set up the settings used in [17] are largely adopted. In the simulation, a 2 lane road with a length of x = 6000 and a simulation time t = 10800 with time steps of 1s is used. Starting at x = 2500 vehicles are able to merge on the main carriageway. 9 loop detectors were placed, namely at x = 500, x = 1000, x = 1500, x = 2000, x = 3500, x = 4000, x = 4500, x = 5000 and x = 5500 (before and after the on-ramp). During the simulation, the intensity on both the motorway and onramp is gradually increased. The demand eventually exceeds the capacity of the road section downstream of the on-ramp. This is done to ensure that congestion will emerge and spill back from the onramp to the 2-lane section, so every LOS will occur at this 2-lane section during the simulation. Only one vehicle class is applied (passenger cars). The applied car following model is a modification of the Intelligent Driver Model [18]. For this model, the parameter settings described in Kesting et al. [19] are adopted. As a lane changing model, the Lane changing Model with Relaxation and Synchronization model (LMRS) described in Schakel et al. [16] will be used.

The simulation software is not able to produce many CAN bus parameters. The floating car parameters that are produced by the simulation software are: speed, acceleration, longitudinal position and lateral position (lane). The detector loops from the simulation provide 1-minute aggregated speed and intensity data. Although this traffic simulation software is not able to provide a vast amount of simulated CAN bus parameters, the speed is expected to be an important indicator for LOS estimation. Therefore, we expect the simulation can be used to indicate a lower boundary for the real LOS estimation performance. In case in reality more in vehicle signals are suitable to be used as LOS indicator, the LOS estimation accuracy will be higher than the found LOS estimation accuracy with the simulation.

2.6.1 LOS ground truths

The LOS estimation with the real CAN bus data, as described in section 2.5, uses LOS ground truth based on the loop detector data. This is not the most accurate way to determine the LOS for three reasons:

- First of all, the density, which the LOS is based on, is calculated with the detector speed and intensity data, but this is not the real density at a road section.
- Secondly, a vehicle passage is matched with a LOS ground truth based on the 1-minute aggregated loop detector data, but the traffic conditions can change drastically during a minute.
- Finally, the loop detectors can be quite inaccurate, as was found in section 1.2, leading to inaccurate LOS estimation.

Therefore, the simulation data will be used to determine the LOS ground truth in multiple ways. This is possible, because besides the detection point data, the simulation data also provides all vehicle trajectories.

In the remaining part of this section, four different LOS ground truth determination methods are described which will be used in this part of the research.

- The first LOS ground truth is determined in the exact same way as the LOS ground truth for the real CAN bus data. Using the 1-minute aggregated speed and intensity data from the simulated detection points, the 1-minute aggregated density is calculated with $k=q/u$. Using this density, LOS ground truth 1 is determined.
- Although LOS ground truth 1 is determined analogously to the LOS ground truth for the real CAN bus data, there is a difference. The data from the simulated detection point is perfect, while the real loop detectors can be quite inaccurate. Therefore, LOS ground truth 2 will also be determined with the same simulated detection point data, but with an added error. This LOS ground truth and the corresponding LOS estimation should provide the best simulation of the LOS estimation with real CAN bus data. Also, by comparing the LOS estimation results for both LOS ground truths, the effect of the loop detector accuracy on the LOS estimation with real CAN bus data can be assessed.
- The third LOS ground truth is based on the average density during a minute on the road section (200m) near the detection point. This ground truth can be used for LOS estimation with single and multiple vehicles.

- LOS ground truth 4 resolves all three issues which were mentioned at the beginning of this paragraph. It is also based on the average density on the road section (200m) near the detection point, but only for the 5 seconds around the moment of the passage. This LOS ground truth can only be used for LOS estimation with single vehicles.

2.6.2 LOS estimation with CAN bus signals from single vehicles

Using the simulation data, first the same steps will be taken as with the real CAN bus data. Using the four LOS ground truths, the same LOS versus CAN bus parameter figures will be made, the LOS versus CAN bus parameter correlations will be calculated and LOS estimation functions will be determined with multinomial logistic regression. The results can be compared with the results of the real CAN bus data. Using these comparisons, the model validity can be discussed.

2.6.3 LOS estimation with CAN bus signals from multiple vehicles

Subsequently, the LOS will be estimated using multiple vehicle passages. The average values of the CAN bus parameters of all vehicles which are considered to be uCAN equipped and pass a detection point in the same minute are taken. These values are used to estimate the LOS estimation functions with multinomial logistic regression. This is done for different penetration rates. The LOS estimation accuracy for every penetration rate will be determined.

2.6.4 LOS estimation with plain floating car data from multiple vehicles

The LOS will also be estimated by only using vehicle positions, simulating vehicles which do not provide CAN bus data, but only positional data. This LOS estimation method uses the following formula to estimate the density:

$$k_{estimated} = k_{measured} / P$$

where: $k_{estimated}$ is the estimated density

$k_{measured}$ is the density of the equipped vehicles

P is the penetration rate

Also for this method, the LOS estimation accuracy will be determined at various penetration rates.

2.6.5 Comparison between LOS estimation methods

Finally, the results of the LOS estimation accuracy for the various methods will be compared.

3 Level of Service and driving behaviour

In this chapter, the literature about driving behavior in different traffic conditions is reviewed which make it possible to assume that LOS can be estimated with the CAN bus data. Several relationships between traffic conditions and driving behaviour are found. Since part of the driving behaviour can be measured with the CAN bus signals, it is probable that also the LOS can be estimated with these signals.

3.1 Speed

First of all, the fundamental diagram shows a clear relation between the speed and density. Although there are many different models for the diagram, they all have in common that as the density increases, the average speed will decrease, as is shown in Figure 8.

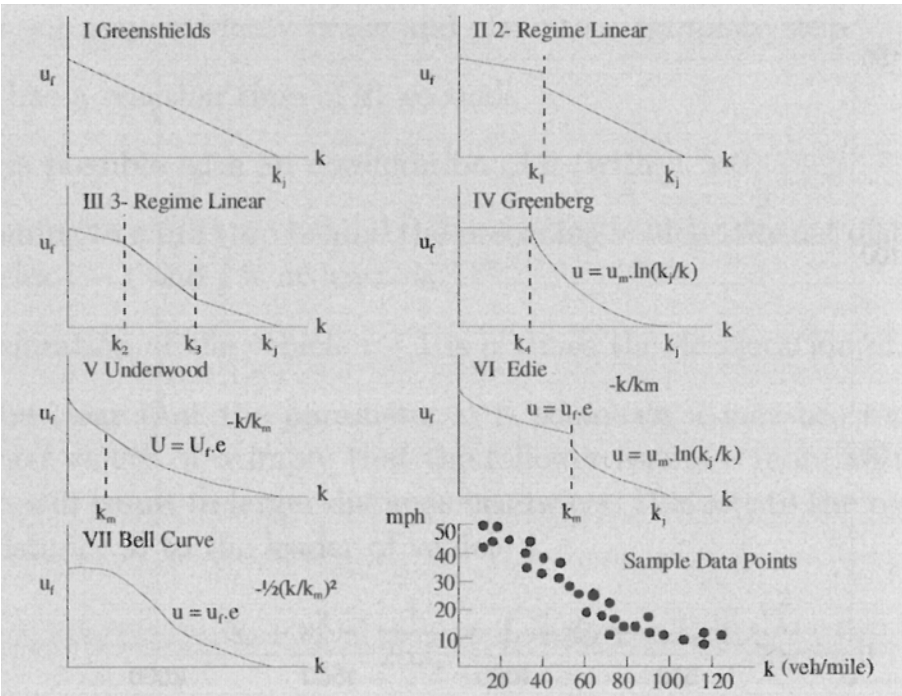

Figure 8: Fundamental diagram models [20]

3.2 Lane change behaviour

Tang Chang-Fu et al. [21] investigated lane change rate versus density. They found that the amount of lane changes on a 2-lane freeway increases when the density increases, up to 17.5 veh/km/lane (0.035 veh/m in Figure 9). If the density gets larger, the amount of lane changes decreases again. Figure 9 shows how the amount of lane changes is affected by the density.

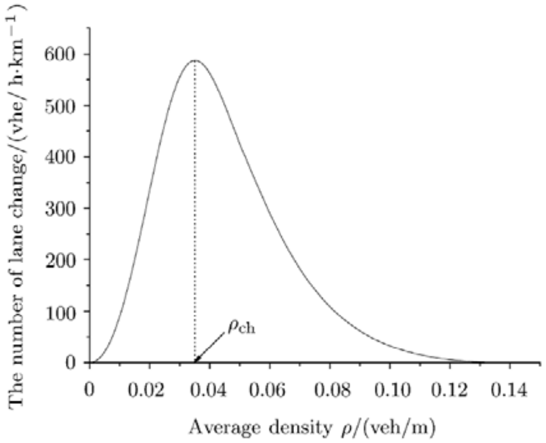

Figure 9: Density versus amount of lane changes [21]

Knoop et al. [22] investigated the dependency of lane change rate on the density of the origin and target lane. Figure 10 shows the dependency they found of the lane change rate between the second and third lane of a 3-lane motorway.

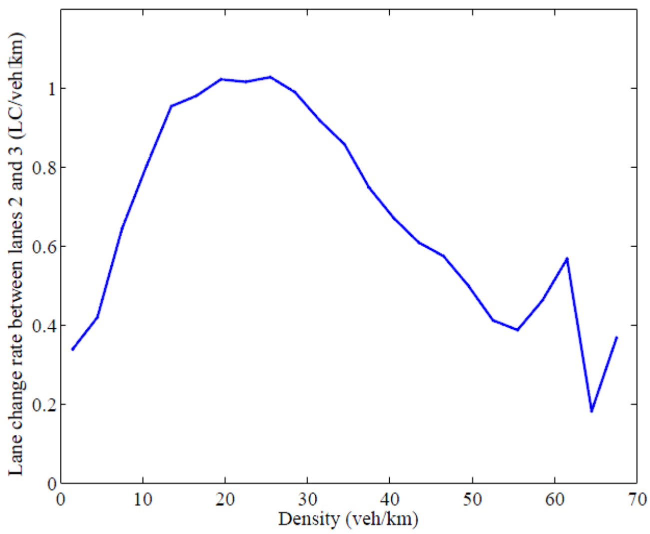

Figure 10: Lane change rate density dependency [22]

Although the GPS data from the uCAN on board unit does not prove to be accurate enough to detect lane changes, the changes in the steering position and usage of direction indicators might show relationships with the density.

3.3 Speed variation

Ko et al. [23] investigated the relation between speed variation with the Level of Service. They used the acceleration noise, defined as the root mean square deviation of the vehicle acceleration, as a measure of the speed variation. It was found that in general the speed variation increases as the density increases. Also, it was observed that the speed variation is more sensitive to traffic microscopic traffic conditions than speed for LOS A, B and C.

3.4 Distance headway

Also the following distance of vehicles depend on the LOS. This is trivial, since the definition of the following distance is the reciprocal of the density:

$$s = 1 / k$$

where: s is the average distance headway (km)

k is the density (veh/km)

4 CAN bus data

In this chapter, the CAN bus data processing method is described and the processed CAN bus data is explored. Also, the accuracy and reliability of the CAN bus signals is investigated, by reviewing literature and by using the processed CAN bus data. The findings in this paragraph will be used in combination with the finding of the previous paragraph, to select those uCAN equipped vehicles which are most likely to be able to accurately estimate LOS.

4.1 CAN bus signal frequencies and reliability

In literature, some information is found on the CAN bus signal frequencies and reliability.
The devices that are connected to the CAN bus differ per vehicle. In general, modern vehicles will have more CAN bus nodes than older vehicles. Every signal is send over the CAN bus network with its own frequency. In Table 5 the period of a number of CAN bus signals is shown.

Table 5: CAN bus signal periods [24]

signal number	signal	period (ms)
1	Traction Battery Voltage	100.0
2	Traction Battery Current	100.0
3	Traction Battery Temp, Average	1000.0
4	Auxiliary Battery Voltage	100.0
5	Traction Battery Temp, Max	1000.0
6	Auxiliary Battery Current	100.0
7	Accelerator Position	5.0
8	Brake Pressure, Master Cylinder	5.0
9	Brake Pressure, Line	5.0
10	Transaxle Lubrication Pressure	100.0
11	Transaction Clutch Line Pressure	5.0
12	Vehicle Speed	100.0

Tindell et al. [24] have investigated the message response times of the CAN bus. They analyzed the worst-case response time of a message. Table 6 shows the worst case response times for a few signals of a CAN bus network (the signal number corresponds to the number in Table 5).
The frequencies of the CAN bus signals are high, ranging from 1 Hz to 200 Hz. The worst case response times of CAN bus signals are in the order of milliseconds, indicating very high reliability. Both properties are expected to be more than sufficient for LOS estimation.

Table 6: CAN bus worst case response times [24]

signal number	period (ms)	worst case response time (ms)
8,9	5.0	2.128
7	5.0	2.632
11	5.0	3.720
15,16,17,19,20,22,26,27	10.0	9.040
18	100.0	10.128
1,2,4,6	100.0	18.944
12	100.0	19.448
10	100.0	19.552
3,5,13	1000.0	20.608

4.2 CAN bus data from Smart In-Car

Ideally, the available data from the uCAN equipped vehicles would be made available in a matrix format, containing all parameter values in single row. This however, is not the case. In Figure 11, a couple of lines from the raw data files are shown. Instead of containing all parameters, most lines only contain one CAN bus parameter. They are logged individually in so called json format. Since we want to be able to perform analyses on the data, the raw data has to be converted to a proper format which allows these analyses. Therefore the raw data is imported in Matlab and converted to a matrix format.


```
{ "timestamp" : 1335874566255000 , "gpsdata" : { "lat" : 51.46685527777779 , "lon" : 5.473162500000002} ,
    "uvm" : { "id" : 119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874566511100 , "candata" : { "Speed" : 22.579200744628906} , "uvm" : { "id" :
    119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874566656500 , "candata" : { "RPM" : 1797.0} , "uvm" : { "id" : 119611000024168 ,
    "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874567031300 , "candata" : { "Speed" : 22.924800872802734} , "uvm" : { "id" :
    119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874567335000 , "gpsdata" : { "lat" : 51.466869444444455 , "lon" : 5.473101944444446} ,
    "uvm" : { "id" : 119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874567551200 , "candata" : { "Speed" : 23.443201065063477} , "uvm" : { "id" :
    119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874567666100 , "candata" : { "RPM" : 1822.0} , "uvm" : { "id" : 119611000024168 ,
    "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874568071100 , "candata" : { "Speed" : 23.50080108642578} , "uvm" : { "id" :
    119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874568332000 , "gpsdata" : { "lat" : 51.46687861111112 , "lon" : 5.473010833333335} ,
    "uvm" : { "id" : 119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874568590900 , "candata" : { "Speed" : 23.96160125732422} , "uvm" : { "id" :
    119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874568667700 , "candata" : { "RPM" : 1848.0} , "uvm" : { "id" : 119611000024168 ,
    "licenseplate" : "16-XN-FT"}}
```


Figure 11: Example of the raw CAN bus data

A complete description of the conversion methods can be found in Appendix E. A few rows and column from the resulting Matlab matrix is shown in Table 7. These are the converted lines from Figure 11.

Table 7: 1 second aggregated data in Matlab

year	month	day	hour	minute	second	epoch timestamp	uvm	lon	lat	rpm	speed
2012	5	1	12	16	6,4742	1335874566474200	119611000024168	5,4732	51,4669	1797	22,5792
2012	5	1	12	16	7,3959	1335874567395900	119611000024168	5,4731	51,4669	1822	23,1840
2012	5	1	12	16	8,4154	1335874568415420	119611000024168	5,4730	51,4669	1848	23,7312

4.3 CAN bus visualizations and quality

In section 1.2 of Appendix E, many visualizations of the CAN bus data are presented. These are, together with the analysis in section 1.3 of Appendix E, used to assess the reliability and accuracy of the CAN bus signals. The most important findings are presented here.

Although the uCAN equipped vehicles provide a vast amount of data, the penetration rate is limited. The frequency of the CAN bus parameters differs per vehicle and parameter, but are often around once per second. Also the available CAN bus parameters differ per vehicle.

Except for a few clear outliers which mostly occur when the vehicle engine is started or turned off, which can be easily filtered, the CAN bus data shows realistic values. When comparing some of the CAN bus parameters, they also show expected relationships. Also, when a single variable is measured in multiple ways, the results match. Not all parameters have been thoroughly examined. The accuracy of these parameters can't be guaranteed.

The GPS position in areas with builtup close to the road and below trees can show large errors from the actual position. The GPS position on motorways is quite accurate. The unit of the fuel consumption is unknown.

When using the data for analyses, there are some limitations that will have to be kept in mind. The data is only gathered by taxi drivers. It is quite probable that their driving behaviour differs from other road users. The taxis are mainly present in the city centre of Eindhoven. Their usage of the motorways is relatively small. Also, the dataset only contains data from a limited amount of vehicle types.

5 Results

5.1 LOS estimation with real CAN bus data

The definition of Level of Service, the determined LOS dependent driving behaviour factors and the findings from the CAN bus data exploration in the previous paragraphs will be used in this paragraph to estimate the LOS with CAN bus data. The goal of this chapter is to determine the accuracy of LOS estimation with CAN bus signals and to assess the effect of different signal transmission frequencies and different data collection radii.

5.1.1 CAN bus data vehicles

Based on the findings in chapter 3, regarding the driving behaviour dependency on density, CAN bus signals can be identified which are likely to be able to measure these behavioural factors. Speed, speed variation, the amount of lane changes and the distance headway were shown to depend on density. Speed can be measured with the speedometer. The speed variation can also be measured with the speedometer, but the brake signal and the RPM signal may also be indicators for speed variation. The CAN bus signals which are most likely to be able to indicate lane changes are the indicator signals and the steering position, since these values will change during a lane change. Besides the available signals (Table 20, Appendix E), there could be other CAN bus signals which are not available in the data set, like cruise control and brake force related CAN bus signals, which are related with Level of Service.

It is decided that the CAN bus data of the 35 Mercedes E-class (212) 2009 taxis will be used, because many of these CAN bus signals are provided by these vehicles. Another reason to select only one vehicle type is to ensure the signals from different vehicles will have the same properties. As was shown in Table 20 in Appendix E, the CAN bus data of these vehicles include the following signals:

- RPM
- Speed
- Brake
- SteeringPosition
- RDI_Switch (right direction indicator)
- LDI_Switch (left direction indicator)
- FuelConsumption
- WiperMotor

The CAN bus data from these vehicles which was collected between August 1st 2012 and September 30th 2012 is used. This period is chosen because during this period these vehicles were known to collect the desired CAN bus signals.

5.1.2 Loop detectors

From the available loop detectors, shown in Figure 4, all loop detectors are selected which are located on road sections with a 2*2-lane cross section, as described in section 2.5.1, and located at least 200m from on-ramps, off-ramps and weaving sections. In total, around 140 loop detector locations are included in the research. A list of these loop detectors is included in Table 24 in Appendix F. The locations of the loop detectors are shown in Figure 12. The speed limit on all these motorway segments is 120 km/h. This hasn't changed after September 1st 2012, when the speed limit on most Dutch motorways was increased to 130 km/h, see Figure 69 in Appendix G. This enables us to combine the data from the different loop detector locations, since the conditions on all road segments are to large extent the same. A drawback is that the data does not give the opportunity to assess the effect of different speed limits. When the speed limit is lower and thus closer to the speed near capacity conditions (LOS E), it is probable that the differences in driving behaviour (e.g. speed) in the LOS's will be smaller, making estimation of the LOS based on CAN bus signals less accurate.

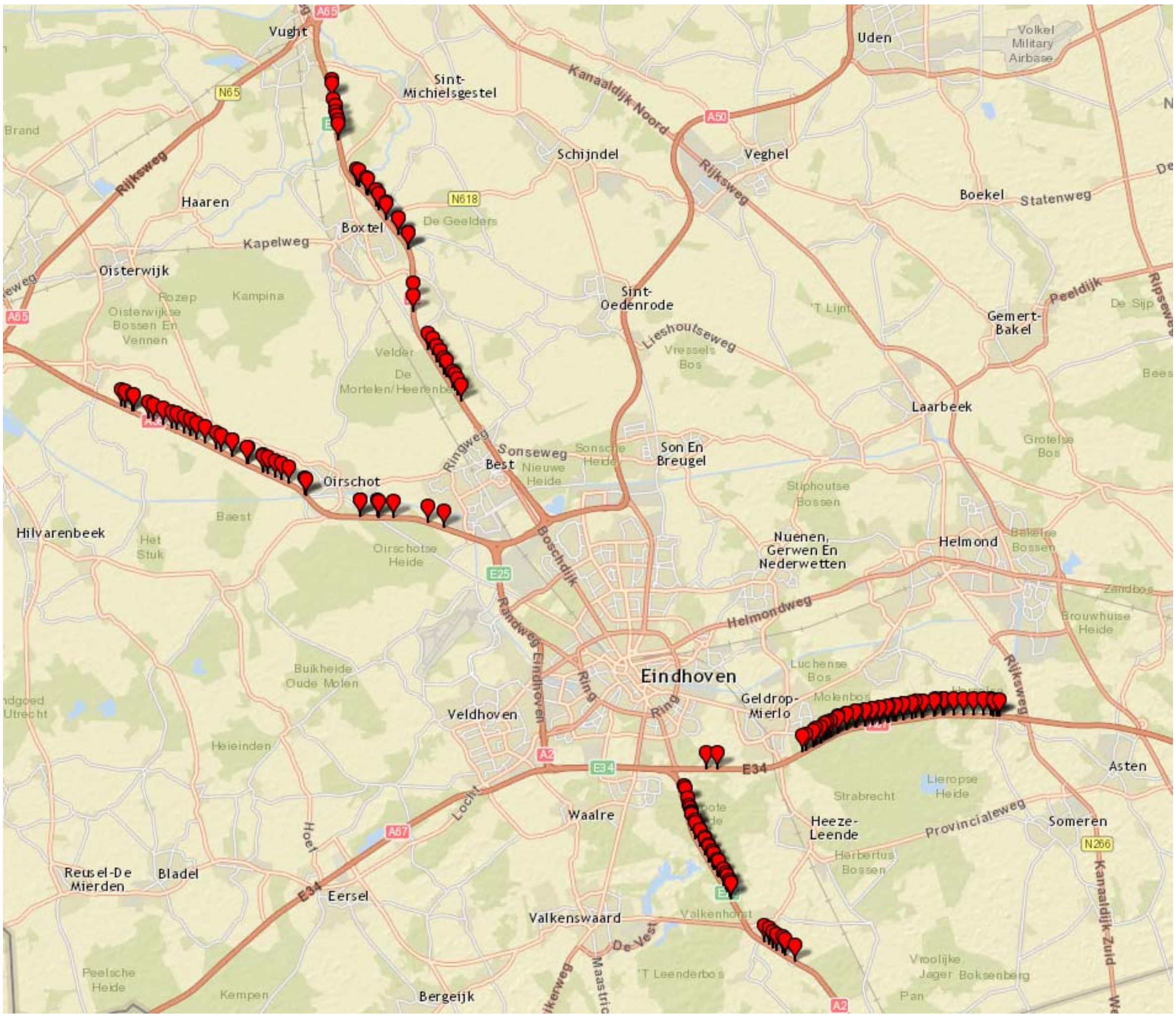

Figure 12: Locations of the used loop detectors. The red icons indicate the locations of the used loop detectors.

5.1.3 LOS ground truth

As was shown in section 1.2 in Appendix B, the loop detectors used in the Netherlands can be quite unreliable. Many loop detectors where found which do not meet the NDW requirements [25][26]. Unfortunately, the real accuracy of the loop detectors, which are used for determining the LOS ground truth, is unknown. However, the effect of the inaccuracy of the loop detectors on the LOS ground truth determination can be assessed by adding a measurement error to the loop detector measurements.

The MAPE of the loop detectors is set at 5%, the limit of the NDW requirement (see section 1.7 in Appendix B). The assumption is made that the error is normally distributed. The following simulated erroneous loop detector values are generated:

$$v_{error} = v_{measured} * (1+normrand(\mu,\sigma))$$

$$i_{error} = i_{measured} * (1+normrand(\mu,\sigma))$$

with: normrand(μ,σ) is a function which randomly selects a point from the normal distribution

$\mu = 0$

$\sigma = 0.0625$

The v_{error} and i_{error} is calculated for all 7184 passages. Using v_{error} and i_{error}, k_{error} is calculated and LOS_{error} is determined. When comparing LOS to LOS_{error}, in 11% of the passages, LOS and LOS_{error} do not match. This means that, if the assumed MAPE of 5% corresponds with the real MAPE of the loop detectors, in 11% of the passages, the LOS ground truth will have been determined incorrectly (see Appendix I). This also means, that the currently used LOS determination using loop detectors, even when complying with the NDW requirements, has an inaccuracy of up to 11%. If the unknown real MAPE would actually be larger, as is found by [25][26], this percentage will be even larger. It is also likely that individual loop detectors will not have an average error of zero. They are more likely to have a negative or positive bias ($\mu \neq 0$), which could also lead to a higher inaccuracy of the LOS ground truth determination.

The effect of loop detector inaccuracy on LOS ground truth determination will be further investigated in section 5.2, using the simulated loop detector data.

5.1.4 Combining the CAN bus data with the loop detector data

In total, during August and September 2012, 7184 times one of the 35 selected uCAN equipped taxis passed one of the selected loop detectors.

Every time a vehicle passed a loop detector, the data points of that vehicle that are near the loop detector are collected. To assess the effect of different signal transmission frequencies and using different detection point radii, this is done with variable values of the signal transmission frequency and the detection point radius. The different signal periods which are used are 1s, 2s, 3s, 4s, 5s and 6s. The detection point radii are 50m, 100m, 200m and 400m. This yields 24 sets of data points for every passage.

Note that in some cases, the period is too long and the radius too small to obtain a data point from a vehicle which passes the loop detector. This can occur when the vehicle travels over 2*radius in

the time window of the period. This means that some period-radius combinations result in missed detections.
For every set of passage data points, the mean, standard deviation, minimum and maximum of every CAN bus parameter is stored. This information is combined with the corresponding loop detector data, including the determined LOS. Table 8 shows the information that is included in the resulting table with passages. Information about the units of the parameters can be found in Table 19 in Appendix E.

Table 8: Combined CAN and loop detector data per passage

column	parameter	description
1	dp_speed	loop detector 1-minute aggregated speed during passage
2	dp_intensity	loop detector 1-minute aggregated intensity during passage
3	dp_density	1-minute aggragated density (intensity/speed) during passage
4	dp_LOS	1-minute aggregated density LOS during passage, determined with density
5	passagenr	unique number of passage
6	Road	road number of motorway
7	direction	side of road (0=left, 1=right)
8	hectometer	the longitudional coordinate of the loop detector position
9	year	moment of passage
10	month	moment of passage
11	day	moment of passage
12	hour	moment of passage
13	minute	moment of passage
14	second	moment of passage
15	epochtime	epoch time of passage
16	uvm	unique vehicle number
17	lon	average GPS longitude during passage
18	lat	average GPS latitude during passage
19	rpm_avg	average of motor RPM during passage
20	rpm_std	standard deviation of motor RPM during passage
21	rpm_max	maximum of motor RPM during passage
22	rpm_min	minumum of motor RPM during passage
23	speed_avg	average vehicle speed during passage
24	speed_std	standard deviation of vehicle speed during passage
25	speed_max	maximum of vehicle speed during passage
26	speed_min	minimum of vehicle speed during passage
27	brake_avg	average brake usage during passage
28	brake_std	standard deviation of brake usage during passage
29	brake_max	maximum of brake usage during passage
30	brake_min	minimum of brake usage during passage
31	steeringposition_avg	average steering position during passage
32	steeringposition_std	standard deviation of steering position during passage
33	steeringposition_min	maximum of steering position during passage
34	steeringposition_max	minimum of steering position during passage
35	steeringofcenter_avg	average steering position of center during passage
36	steeringofcenter_std	standard deviation of steering position of center during passage

column	parameter	description
37	steeringofcenter_min	maximum of steering position of center during passage
38	steeringofcenter_max	minimum of steering position of center during passage
39	RDI_avg	average right direction indicator usage during passage
40	RDI_std	standard deviation of right direction indicator usage during passage
41	RDI_min	maximum of right direction indicator usage during passage
42	RDI_max	minimum of right direction indicator usage during passage
43	LDI_avg	average left direction indicator usage during passage
44	LDI_std	standard deviation of left direction indicator usage during passage
45	LDI_min	maximum of left direction indicator usage during passage
46	LDI_max	minimum of left direction indicator usage during passage
47	hazardlights_avg	average hazard lights usage during passage
48	hazardlights_std	standard deviation of hazard lights usage during passage
49	hazardlights_min	maximum of hazard lights usage during passage
50	hazardlights_max	minimum of hazard lights usage during passage
51	RearFogLight_avg	average rear fog light usage during passage
52	RearFogLight_std	standard deviation of rear fog light usage during passage
53	RearFogLight_min	maximum of rear fog light usage during passage
54	RearFogLight_max	minimum of rear fog light usage during passage
55	FrontFogLight_avg	average front fog light usage during passage
56	FrontFogLight_std	standard deviation of front fog light usage during passage
57	FrontFogLight_min	maximum of front fog light usage during passage
58	FrontFogLight_max	minimum of front fog light usage during passage
59	WiperMotor_avg	average wiper motor usage during passage
60	WiperMotor_std	standard deviation of wiper motor usage during passage
61	WiperMotor_min	maximum of wiper motor usage during passage
62	WiperMotor_max	minimum of wiper motor usage during passage
63	FuelConsumption_avg	average fuel consumption during passage
64	FuelConsumption_std	standard deviation of fuel consumption during passage
65	FuelConsumption_min	maximum of fuel consumption during passage
66	FuelConsumption_max	minimum of fuel consumption during passage

5.1.5 LOS estimation with CAN bus signals

After collecting the CAN bus data and the LOS ground truth and combining them in the previous paragraphs, we can now investigate the LOS estimation possibility with CAN bus signals using the resulting table from section 5.1.4. First of all, the CAN bus signal values of every passage will be plotted against the LOS, to get a first impression of the LOS dependency of the CAN bus signals. Secondly, the obtained CAN bus signal passage data will be used to investigate the correlation between the CAN bus signals and the LOS, to identify the CAN bus signals which have the strongest correlation with LOS. With this information also a combination of a period and radius will be chosen, which shows the strongest correlations with LOS, which will be used in the further analysis. After this step, multinomial logistic regression is used to determine a LOS estimation formula with single and multiple CAN bus parameters. Finally, these LOS estimation formulas are used to estimate the LOS during the passages and the accuracy of these estimations is presented for every CAN bus parameter or CAN bus parameter combination.

5.1.6 LOS versus CAN bus parameter values

In this paragraph, the values of some of the CAN bus parameters will be plotted against the LOS, determined with the measured 1-minute aggregated speed and intensity values from the loop detectors. In all figures in this paragraph, the central red mark inside the blue box indicates the median. The edges of the box represent the 25th and 75th percentiles. The whiskers extend to the most extreme data points not considered outliers and outliers are plotted individually with red crosses.

Figure 13 - Figure 20 show the measured values of several CAN bus parameter at the different LOS. In all figures, the used radius is 200m and the used period is 1s. Figure 13 shows the LOS versus the average speed of the single vehicle passages. For increasing LOS, the average speed decreases. The average speed at LOS A and B is very similar, but the differences between the values at higher LOS are larger. The overlap between subsequent LOS is however substantial.
Figure 14 shows the standard deviation of the measured speed versus the different LOS. On average, the standard deviation of the speed increases for increasing LOS. Also here, there is a lot of overlap of values at different LOS.
Figure 15 and Figure 16 show the LOS versus average fuel consumption and standard deviation of fuel consumption respectively. Also for these CAN bus parameters, a respectively decreasing and increasing trend is visible. But the overlap between all LOS is very large, disallowing differentiation of LOS based on the fuel consumption parameters.
Figure 17 shows the LOS versus the standard deviation of the measured steering position. This parameter is very similar for all LOS, only showing slightly higher values at LOS F. This makes it very unlikely that this parameter will be useful to estimate the LOS.
Figure 18 depicts the brake usage of the passing vehicles for the different LOS. For all LOS, the median value of the brake usage is 0, meaning that the brake is not used at all in the lion's share of the passages. For LOS A-E, also the 75th percentile is 0. In LOS F, the 75th percentile value is larger than 0, meaning that over 25% of the vehicles in LOS F are using their brakes at least once during the passage.
In Figure 19 and Figure 20 show the average RPM and standard deviation of the RPM respectively versus the LOS. These figures are very similar to the speed related Figure 15 and Figure 16, although the scatter within a LOS and the overlap between LOS are both larger.

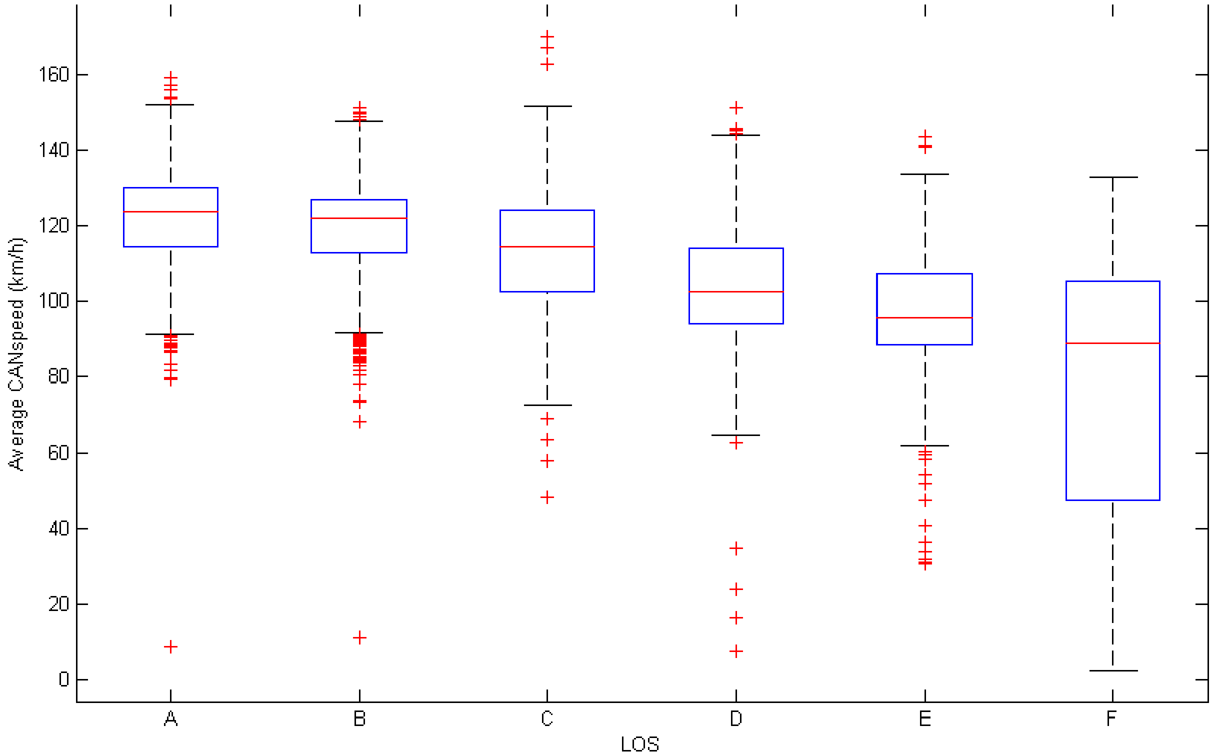

Figure 13: LOS versus average CAN bus speed. The average CAN bus speed of the uCAN equipped vehicles passing a loop detector is plotted against the Level of Service ground truth in the corresponding minute, determined with the loop detector data. The central red mark inside the blue box indicates the median. The edges of the box represent the 25th and 75th percentiles. The whiskers extend to the most extreme data points not considered outliers and outliers are plotted individually with red crosses.

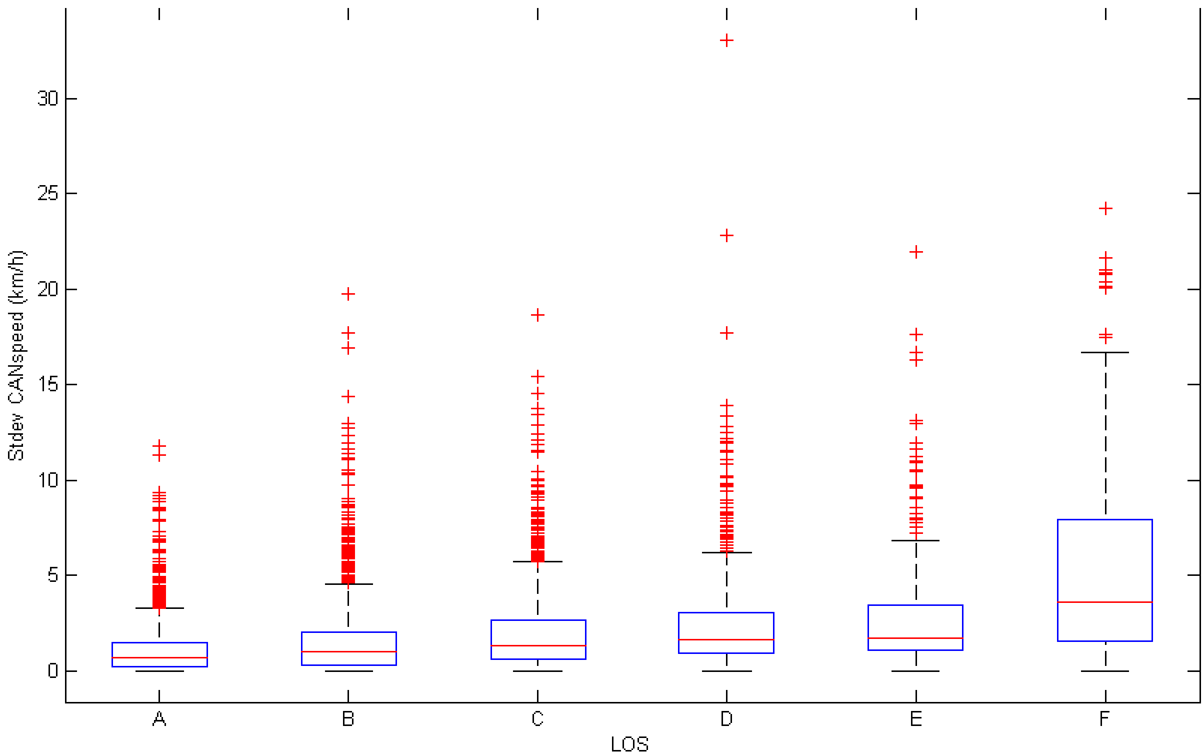

Figure 14: LOS versus st.dev. of CAN bus speed. The standard deviation of the CAN bus speed of the uCAN equipped vehicles passing a loop detector is plotted against the Level of Service ground truth in the corresponding minute.

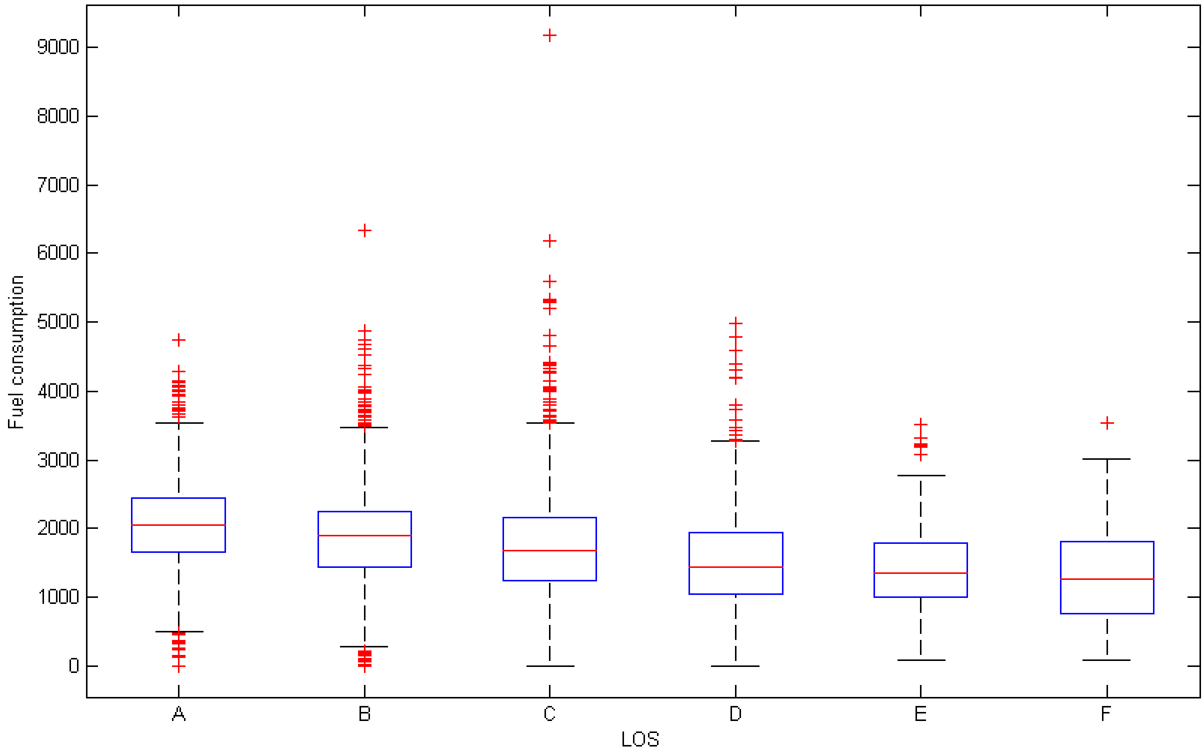

Figure 15: LOS versus average fuel consumption. The average fuel consumption of the uCAN equipped vehicles passing a loop detector is plotted against the Level of Service ground truth in the corresponding minute. The unit of fuel consumption is unknown.

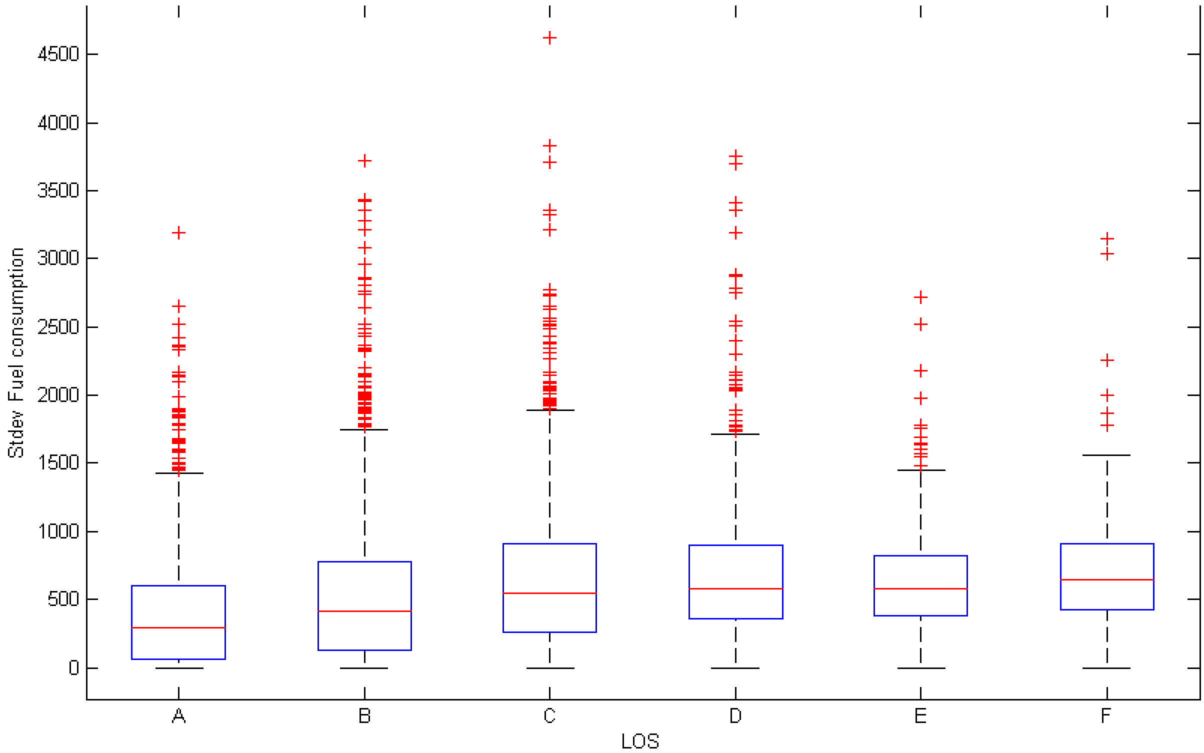

Figure 16: LOS versus st.dev of fuel consumption. The standard deviation of fuel consumption of the uCAN equipped vehicles passing a loop detector is plotted against the Level of Service ground truth in the corresponding minute. The unit of fuel consumption is unknown.

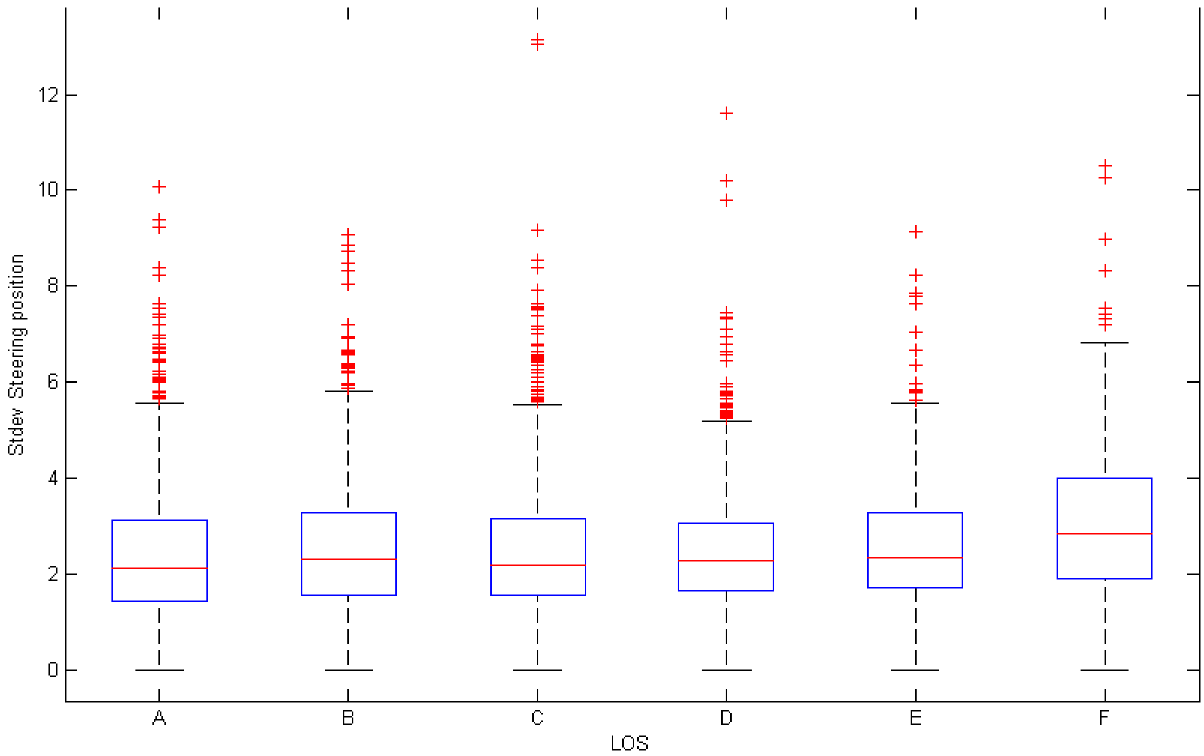

Figure 17: LOS versus st.dev of steering position. The standard deviation of the steering position of the uCAN equipped vehicles passing a loop detector is plotted against the Level of Service ground truth in the corresponding minute. The unit of steering position is unknown.

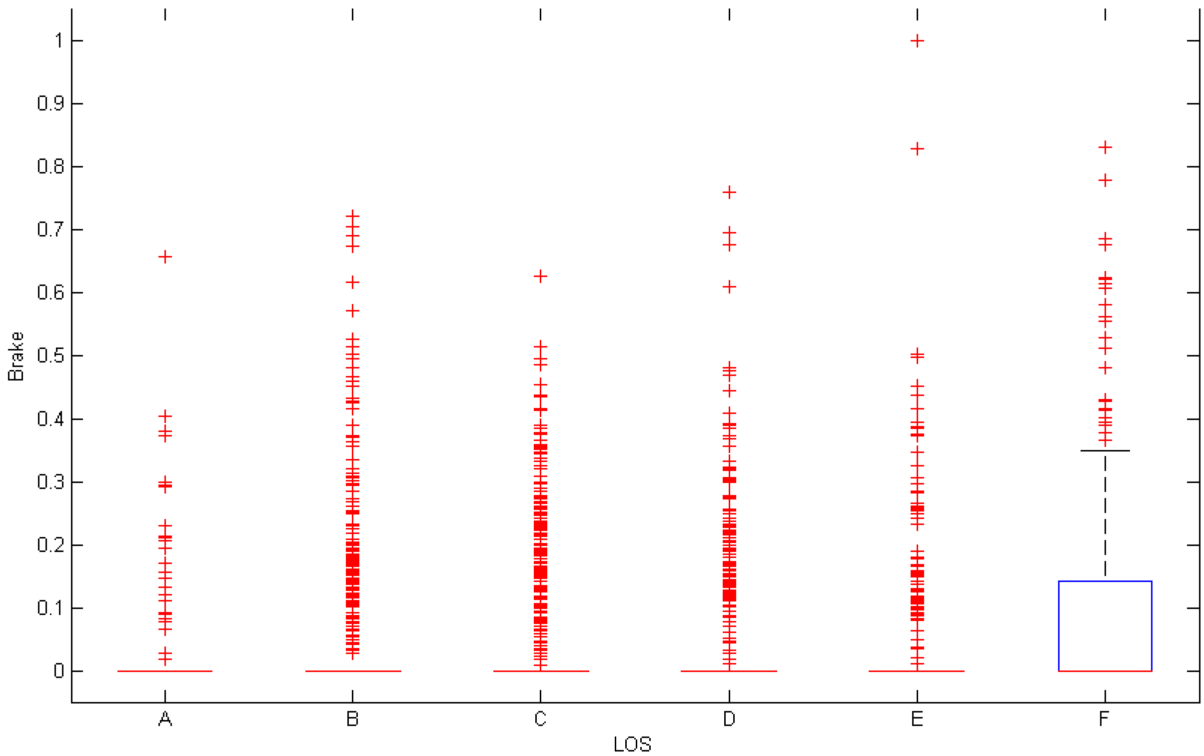

Figure 18: LOS versus average brake usage. The average brake usage (the number of seconds in which the brake was used during the vehicle passage devided by the total number of seconds of the vehicle passage) of the uCAN equipped vehicles passing a loop detector is plotted against the Level of Service ground truth in the corresponding minute.

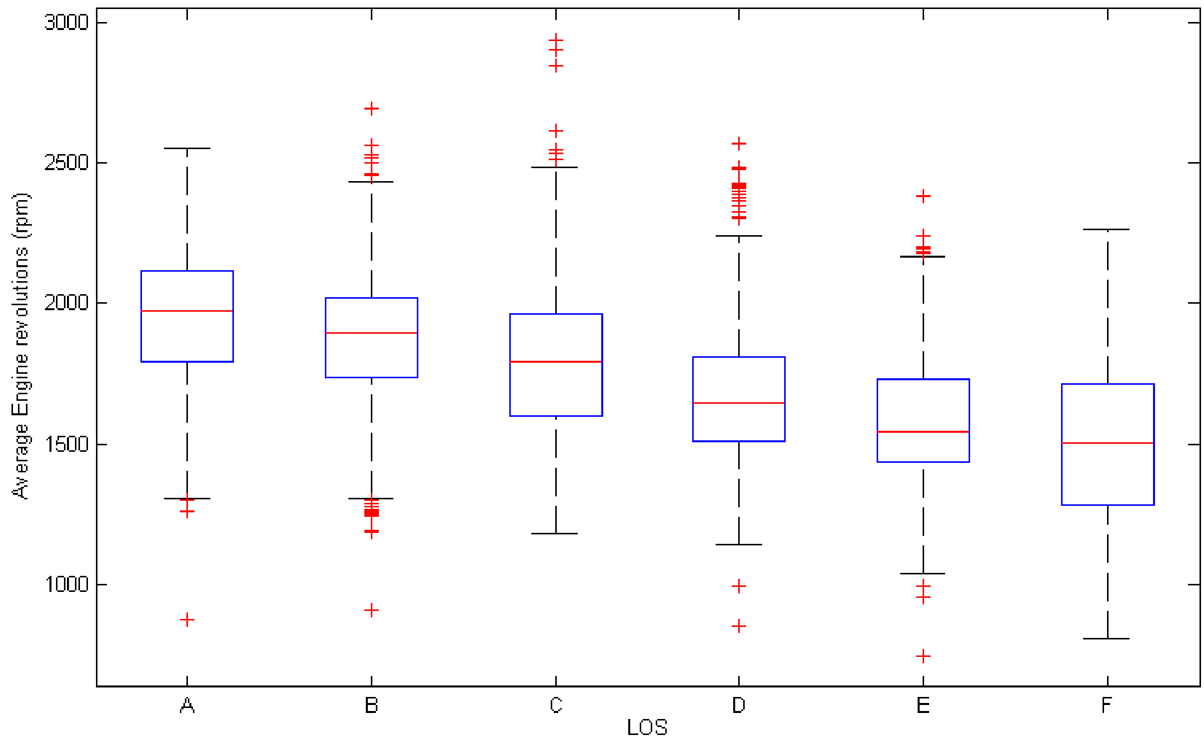

Figure 19: LOS versus average engine rotational speed. The average engine rotational speed of the uCAN equipped vehicles passing a loop detector is plotted against the Level of Service ground truth in the corresponding minute.

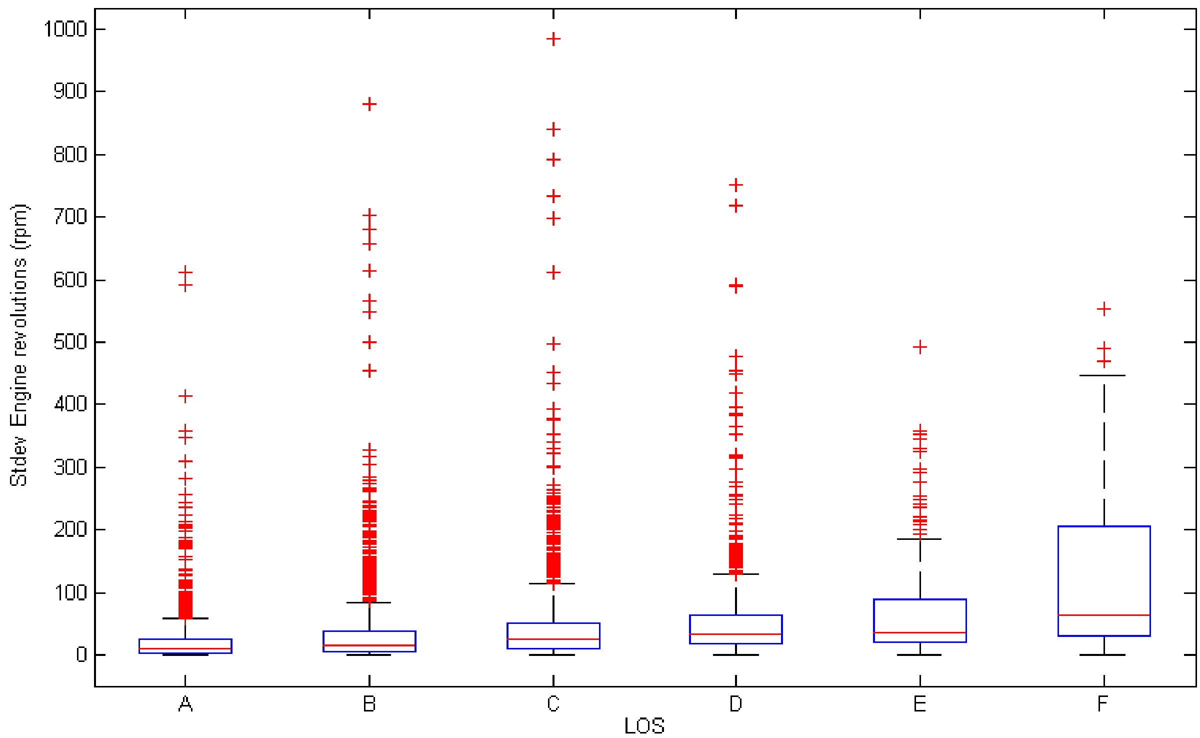

Figure 20: LOS versus st.dev of the engine rotational speed. The standard deviation of the engine rotational speed of the uCAN equipped vehicles passing a loop detector is plotted against the Level of Service ground truth in the corresponding minute.

5.1.7 Relationship between LOS and CAN bus parameters

The figures in the previous paragraph have shown that some of the parameters clearly differ at the various LOS, albeit with a lot of overlap. To assess which CAN bus parameters can be used to estimate the Level of Service and to assess which combination of the CAN bus signal transmission period and the radius is most effective, the correlation between the LOS and all CAN bus parameters is determined. Since the figures in the previous section clearly show that a correlation exists between LOS and the CAN bus parameters, but these do not appear to be linear, instead of the normal Pearson correlation coefficient, the Spearman rank correlation coefficient is determined. This is also done because the LOS has an ordinal scale, for which Pearson correlation is less suitable. The Spearman correlation coefficient assesses how well the relationship between two variables can be described using a monotonic function. In Appendix H, Table 25 - Table 28 show the R and P values for the Spearman correlation between the CAN bus parameters and the LOS, for several values of the radius and the signal period. Table 9 and Table 10 show the same R values, but only for the CAN bus parameters which show the strongest correlation with LOS and have a P-value near zero. An R value of 0 indicates no correlation, an R value of 1 or -1 relatively indicates perfect positive of negative correlation. The P-value indicates the chance of no correlation between the LOS and the specified parameter.

Table 9: R values Spearman correlation between LOS and CAN bus parameters, for 50m & 100m radii and 1-6s CAN bus signal transmission periods.

CAN bus parameter	50m						100m					
	1s	2s	3s	4s	5s	6s	1s	2s	3s	4s	5s	6s
rpm_avg	-0,45	-0,45	-0,45	-0,45	-0,46	-0,47	-0,45	-0,45	-0,45	-0,44	-0,44	-0,45
rpm_std	0,30	0,27	0,22	0,19	0,20	0,19	0,32	0,30	0,29	0,25	0,22	0,21
rpm_max	-0,46	-0,45	-0,46	-0,46	-0,46	-0,47	-0,47	-0,46	-0,46	-0,46	-0,45	-0,45
rpm_min	-0,42	-0,43	-0,44	-0,45	-0,45	-0,46	-0,40	-0,40	-0,42	-0,42	-0,42	-0,43
speed_avg	-0,46	-0,46	-0,47	-0,47	-0,47	-0,49	-0,46	-0,46	-0,46	-0,46	-0,46	-0,46
speed_std	0,28	0,26	0,23	0,19	0,20	0,19	0,30	0,29	0,28	0,25	0,22	0,22
speed_max	-0,46	-0,46	-0,47	-0,47	-0,47	-0,49	-0,47	-0,46	-0,47	-0,46	-0,46	-0,46
speed_min	-0,45	-0,46	-0,47	-0,47	-0,47	-0,49	-0,45	-0,45	-0,45	-0,45	-0,46	-0,46
brake_avg	0,12	0,12	0,11	0,11	0,12	0,10	0,16	0,15	0,15	0,14	0,14	0,13
brake_std	0,13	0,12	0,11	0,11	0,12	0,11	0,15	0,15	0,15	0,13	0,14	0,15
steeringposition_std	0,09	0,17	0,20	0,19	0,19	0,18	0,05	0,05	0,08	0,14	0,17	0,20
steeringofcenter_std	0,08	0,17	0,20	0,19	0,19	0,18	0,06	0,06	0,08	0,14	0,17	0,19

The CAN bus parameters which show the strongest correlations, for all period-radius combinations, are average RPM, maximum RPM, average speed, maximum speed and minimum speed. For many CAN bus parameters, it appears that the correlation doesn’t reduce for increasing periods. It appears that for many CAN bus parameters, a single data point provides enough information. For some CAN bus parameters, especially the standard deviation parameters, this doesn’t hold. For an adequate determination of the standard deviation, multiple data points are necessary. The effect of the radius on the correlation also seems to be limited, but larger radii generally show greater correlations than smaller radii. Based on these findings, in the rest of the analysis, a radius of

200m and a period of 1s will be used. 400m is not chosen, because this length is greater than the minimum distance between the selected loop detectors and on-ramps, off-ramps and weaving sections, as mentioned in section 5.1.2. Using a radius of 400m would contaminate the data.

Table 10: R values Spearman correlation between LOS and CAN bus parameters, for 200m & 400m radii and 1-6s CAN bus signal transmission periods.

CAN bus parameter	radius = 200m						radius = 400m					
	1s	2s	3s	4s	5s	6s	1s	2s	3s	4s	5s	6s
rpm_avg	-0,45	-0,45	-0,45	-0,45	-0,45	-0,45	-0,46	-0,46	-0,46	-0,46	-0,46	-0,46
rpm_std	0,33	0,33	0,32	0,31	0,30	0,28	0,34	0,34	0,33	0,32	0,32	0,32
rpm_max	-0,48	-0,48	-0,48	-0,47	-0,47	-0,47	-0,49	-0,49	-0,49	-0,49	-0,49	-0,49
rpm_min	-0,37	-0,38	-0,39	-0,39	-0,40	-0,42	-0,32	-0,33	-0,35	-0,35	-0,36	-0,38
speed_avg	-0,46	-0,46	-0,46	-0,46	-0,46	-0,47	-0,47	-0,47	-0,47	-0,47	-0,47	-0,47
speed_std	0,32	0,32	0,32	0,30	0,29	0,27	0,33	0,33	0,32	0,31	0,31	0,31
speed_max	-0,48	-0,48	-0,48	-0,47	-0,47	-0,47	-0,49	-0,49	-0,49	-0,49	-0,49	-0,49
speed_min	-0,44	-0,44	-0,44	-0,44	-0,44	-0,45	-0,42	-0,42	-0,43	-0,43	-0,43	-0,44
brake_avg	0,19	0,19	0,18	0,16	0,18	0,16	0,22	0,22	0,21	0,20	0,21	0,19
brake_std	0,19	0,19	0,18	0,16	0,18	0,16	0,22	0,22	0,21	0,19	0,21	0,19
steeringposition_std	0,05	0,05	0,05	0,05	0,06	0,08	0,04	0,05	0,04	0,05	0,04	0,04
steeringofcenter_std	0,07	0,07	0,06	0,06	0,05	0,09	0,06	0,06	0,06	0,05	0,04	0,05

Besides the correlation between the CAN bus parameters and LOS, also the Pearson correlation between the CAN bus parameters and the 1-minute aggregated speed and intensity and density have been determined. The R and P values can be respectively found in Table 29 and Table 30 in Appendix H. Generally, the CAN bus parameters have the strongest correlation with speed. The correlations with the density are weaker and the correlations with the intensity are the weakest.

5.1.8 LOS estimation

The Level of Service is indicated with a letter, meaning it is measured on an ordinal scale. Therefore, normal regression analysis is not suitable. Instead, since the LOS uses an ordinal scale, multinomial logistic regression is used. With multinomial logistic regression, the β's of the LOS estimation function are determined:

$$P(LOS=y) = e^{(\beta y * Xi)} / \Sigma_k e^{(\beta k * Xi)}$$

First, this is done with using single CAN bus parameters for x_i. Using the resulting functions, for every passage the LOS with the highest probability is chosen as the LOS estimation. The percentage of correct LOS estimations is subsequently calculated. The results can be found in the second left column of Table 11. The CAN bus parameters which provide the highest LOS estimation accuracy are shown in bold letters. As expected, these are the parameters which also show the strongest correlation with the LOS. These CAN bus parameters are combined with the other CAN bus parameters and used as variables x_i to determine the β's of the LOS estimation functions. Again, the LOS is estimated for all passages and the correct LOS estimation percentage is determined. The results are shown on the right side of Table 11.

Table 11: Multinomial logistic regression LOS estimation accuracy using 1 or 2 variables. The percentage indicates the correct LOS estimation rate, when using the specified variable(s).

CAN bus parameter	single variable	2 variables				
		rpm_avg	rpm_max	speed_avg	speed_max	speed_min
rpm_avg	**34,6%**		36,4%	36,9%	36,9%	36,0%
rpm_std	31,1%	35,5%	36,1%	36,8%	36,9%	36,5%
rpm_max	**35,7%**	36,4%		36,5%	36,8%	35,9%
rpm_min	32,1%	35,4%	36,2%	36,9%	37,0%	36,2%
speed_avg	**37,1%**	36,9%	36,5%		36,9%	36,6%
speed_std	30,5%	35,5%	35,7%	36,7%	36,8%	36,8%
speed_max	**37,0%**	36,9%	36,8%	36,9%		36,7%
speed_min	**36,1%**	36,0%	35,9%	36,6%	36,7%	
brake_avg	30,0%	34,6%	35,4%	36,9%	36,9%	36,3%
brake_std	30,1%	34,7%	35,5%	36,8%	36,8%	36,3%
brake_max	29,8%	34,6%	35,7%	37,1%	37,0%	36,1%
brake_min	30,3%	34,8%	35,4%	36,7%	36,8%	36,4%
steeringposition_avg	27,0%	32,1%	33,6%	34,0%	34,1%	33,2%
steeringposition_std	27,1%	32,3%	33,4%	34,2%	34,0%	33,3%
steeringposition_min	26,1%	33,0%	33,9%	34,1%	33,9%	33,3%
steeringposition_max	27,1%	31,8%	33,3%	33,9%	34,0%	33,2%
steeringofcenter_avg	27,0%	32,2%	33,5%	33,9%	33,9%	33,2%
steeringofcenter_std	26,4%	32,5%	33,5%	34,0%	33,7%	33,4%
steeringofcenter_min	27,0%	32,0%	33,2%	33,8%	33,9%	33,0%
steeringofcenter_max	26,3%	32,7%	33,7%	33,9%	33,8%	33,3%
RDI_avg	29,8%	34,7%	35,9%	37,2%	37,0%	36,1%
RDI_std	29,8%	34,7%	35,8%	37,1%	36,9%	36,1%
RDI_min	29,8%	34,7%	35,7%	37,2%	36,9%	36,2%
RDI_max	29,8%	34,6%	35,8%	37,0%	36,9%	36,1%
LDI_avg	29,8%	34,8%	35,8%	37,0%	36,9%	36,1%
LDI_std	29,8%	34,8%	35,7%	37,0%	36,9%	36,1%
LDI_max	29,8%	34,7%	35,7%	37,1%	36,9%	36,1%
FrontFogLight_avg	30,0%	34,7%	35,7%	36,9%	36,8%	36,1%
FrontFogLight_min	30,0%	34,7%	35,7%	36,9%	36,8%	36,1%
FrontFogLight_max	30,0%	34,7%	35,7%	36,9%	36,8%	36,1%
WiperFast_avg	29,6%	34,6%	35,7%	36,9%	36,9%	36,2%
WiperFast_std	29,7%	34,5%	35,7%	36,7%	36,8%	36,1%
WiperFast_min	29,6%	34,6%	35,7%	36,9%	36,8%	36,1%
WiperFast_max	29,6%	34,4%	35,8%	36,7%	36,7%	36,1%

Adding more variables to the multinomial logistic regression does generally not improve the correct LOS estimation percentage. The multinomial logistic regression with the average speed as single variable gives us the best LOS estimation functions, with an accuracy of 37.1%. These functions are visualized in Figure 21. From the figure, several things can be observed. First of all, as was also shown in Figure 13, there is a lot of overlap between subsequent LOS. Secondly, from the figure, the boundary values of the average speed between subsequent LOS can be observed. Finally, the

LOS can only be estimated with large accuracy when the speed is very low (LOS F) or very high (LOS A). For all other speeds, the accuracy is very limited, because those speeds are more likely to occur in several LOS.

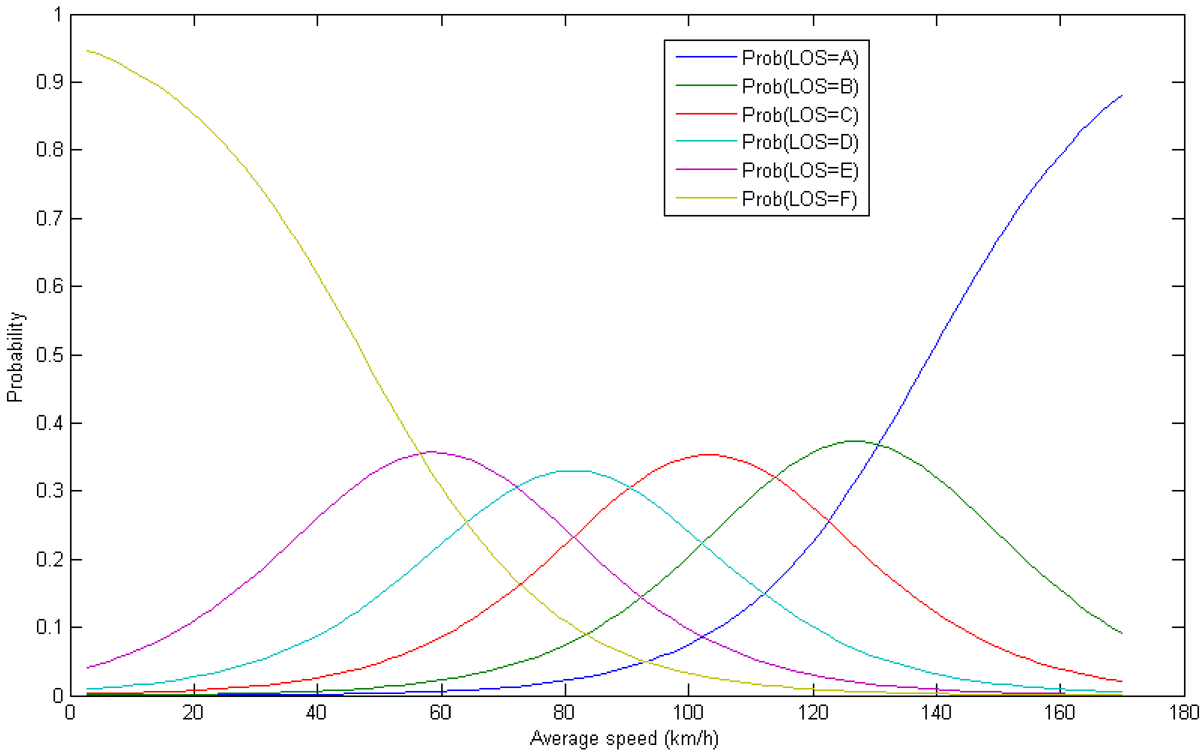

Figure 21: LOS probability functions. The functions indicate, for every value of the average speed of the passing uCAN equipped vehicle, the probability of occurrence of every LOS. The LOS with the highest probability is selected as LOS estimation.

5.2 LOS estimation with simulation

In the previous chapter, we have shown that CAN bus signals can provide relevant data to assess the LOS. However, the penetration rate of the uCAN equipped was too low to estimate the LOS with an acceptable accuracy. In this chapter, simulation data will be used to determine how the accuracy of LOS estimation with CAN bus data improves for higher penetration rates.
Since LOS estimation with only the average speed was found to be the most accurate in the previous section, and the speed is one of the parameters which is available in the simulated data set, we will be able to estimate the LOS estimation accuracy at higher penetration rates, instead of only giving a lower boundary of this accuracy which was expected in section 2.6.
First, the single vehicle passages will be used again to estimate the LOS with the single vehicle passages. Secondly, for every minute and detection point for which a LOS has been determined, the LOS will be estimated with the simulated CAN speed signal from multiple vehicles, for different penetration rates. Finally, the LOS will also be estimated by only using the positional data (simulating the GPS signal) and the results of the different LOS estimation methods will be compared.

5.2.1 LOS ground truths

As was defined in section 1.6.2, LOS is determined by the density. For the real data, the LOS ground truth was determined using the loop detector 1-minute aggregated speed and intensity data to estimate the density, using $k=q/u$. This was the only applicable method to determine the LOS, since no other information was available. In the case of simulation, there is more information available which can be used to determine the LOS. The LOS ground truth in the simulation is determined through four different methods. These are described in this paragraph.

Similar to the real situation, in the simulation detection points are present which provide 1-minute aggregated speed and intensity data which can be used to estimate the density and the corresponding LOS. This method corresponds with LOS determination with perfect functioning loop detectors in real life.
The second LOS ground truth is determined using the detector data with an introduced error. Assuming the same MAPE as in section 5.1.3 of 5%, a random normally distributed error is added to the perfect simulation detector measurements to obtain measurements with an accuracy which is similar to the accuracy of the real loop detector measurements, using:

$$V_{error} = v + \mathrm{normrand}(\mu,\sigma)$$

$$I_{error} = I + \mathrm{normrand}(\mu,\sigma)$$

with: normrand(μ,σ) is a function which randomly selects a point from the normal distribution

$\mu = 0$

$\sigma = 0.0625$

The third LOS ground truth is based on the observed density. Since the simulation data also includes all vehicle trajectories, it is also possible to determine the 1-minute aggregated LOS using the actual density. This is done by measuring the average density in a minute over a 200m road section, 100m upstream and 100m downstream of the detection points.

The fourth LOS ground truth is similar to the third LOS ground truth, but is not 1-minute aggregated. Instead, for every vehicle passage, the LOS at that moment and at that position is determined by the average density in the same 200m road section, but only during the 5 seconds closest to the moment of the detection point passage. This final ground truth gives the highest certainty that the determined LOS ground truth really reflects the traffic conditions that the vehicle encounters during the detection point passage.

5.2.2 LOS estimation with single vehicle passages

LOS ground truths versus simulated CAN bus parameters

In this paragraph, the values of some of the simulated CAN bus parameters will be plotted against the four different LOS ground truths. These figures are consequently compared with the corresponding figures in section 5.1.6.

Figure 22 and Figure 23 respectively show the LOS, determined using the detection point data, versus the average speed and the LOS, determined with the error included detection point data, versus the average speed. Although the LOS determination inaccuracy was found to be quite high, no large differences are found between the figures. This could mean that the loop detector accuracy has no significant effect on the LOS estimation accuracy. This will be further investigated in section 5.2.

In case of a correct estimation of the loop detector accuracy and a perfect simulation model, Figure 23 should look a lot like Figure 13. In general it can be stated that the simulated data shows less overlap between the values at different LOS. Furthermore the values found in LOS A-D are similar in both figures, but the average speed values in LOS E and F are quite different. Both properties contribute to better LOS differentiation possibilities based on the average speed of the simulated CAN bus data. The difference in LOS E and F is thought to be mainly caused by the difference in real driving behaviour and the simulated driving behaviour. To make the simulated model correspond more with reality, the simulation model parameters will have to be altered toward more risk prone behaviour. This will result in smaller headways and higher average speeds at high densities. Model calibration has not been executed in this study and is left as a recommendation for further research.

Figure 24 depicts the average speed for LOS ground truth 3. The average speed values for LOS A-D are similar to the values for other LOS ground truths. Differences occur for LOS E. The average speed at LOS E is higher when using LOS ground truth 3 and more concentrated.

LOS ground truth 4 provides similar results, illustrated in Figure 25, but the amount of outliers for LOS A-C is strongly reduced, compared to the amount of outliers for the other LOS ground truths in Figure 22 - Figure 24. This is an indication that LOS ground truth 4 is indeed the best LOS ground truth determination method.

In the following paragraph, the average speed dependency on LOS will be further discussed, by calculating the correlation and the LOS estimation accuracy using multinomial logistic regression. Also, based on those findings, the model validity will be further discussed.

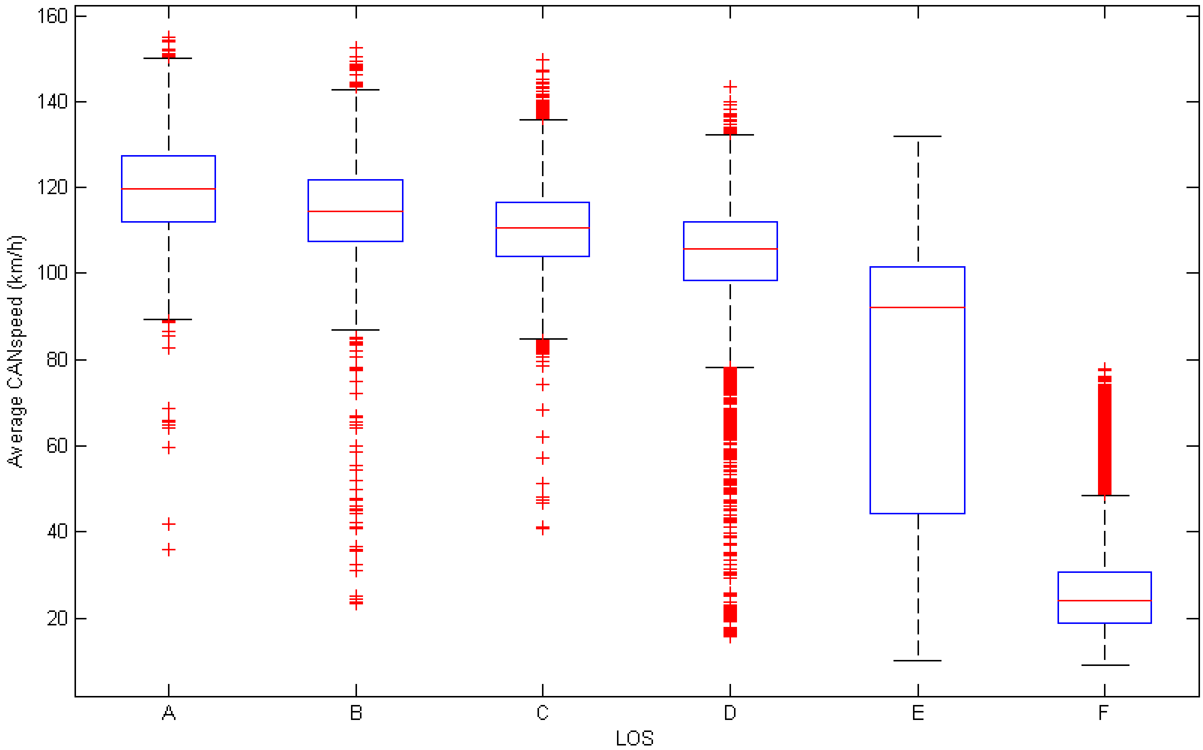

Figure 22: LOS ground truth 1 versus average speed. The average simulated CAN bus speed of the simulated vehicles passing a detection point is plotted against the Level of Service ground truth in the corresponding minute, determined with the detection point data.

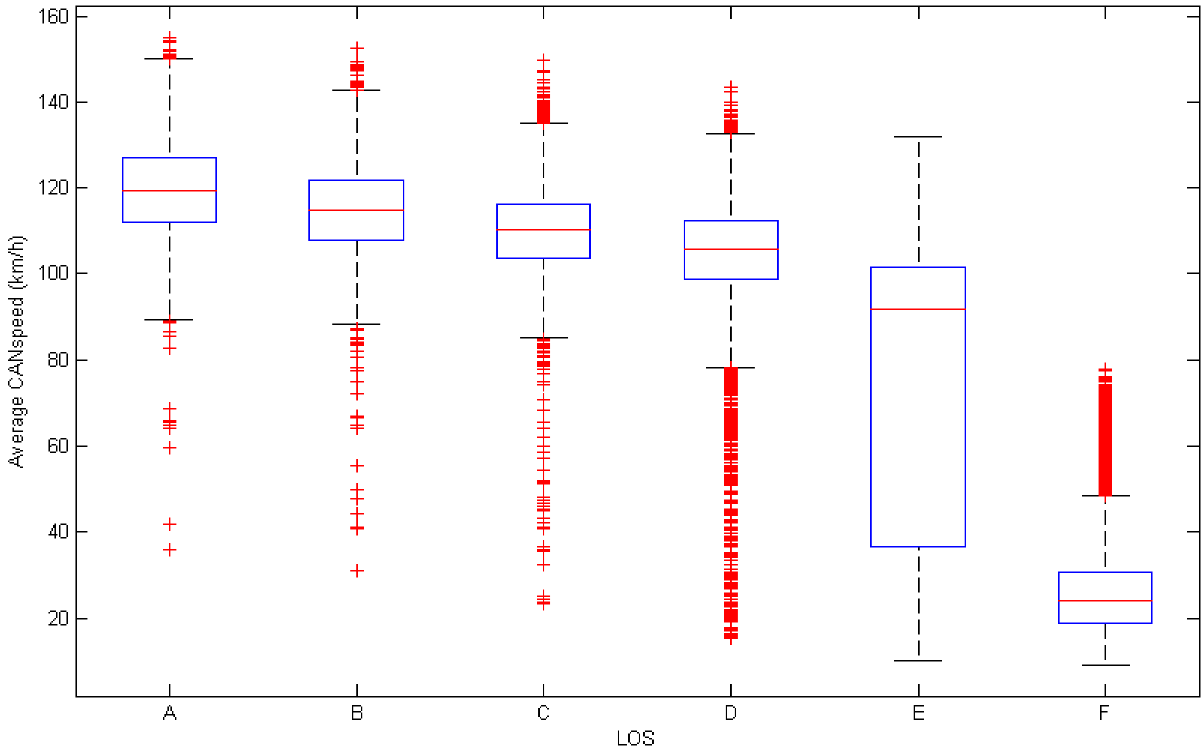

Figure 23: LOS ground truth 2 versus average speed. The average simulated CAN bus speed of the simulated vehicles passing a detection point is plotted against the Level of Service ground truth in the corresponding minute, determined with the detection point data, including an error, simulating real loop detector performance.

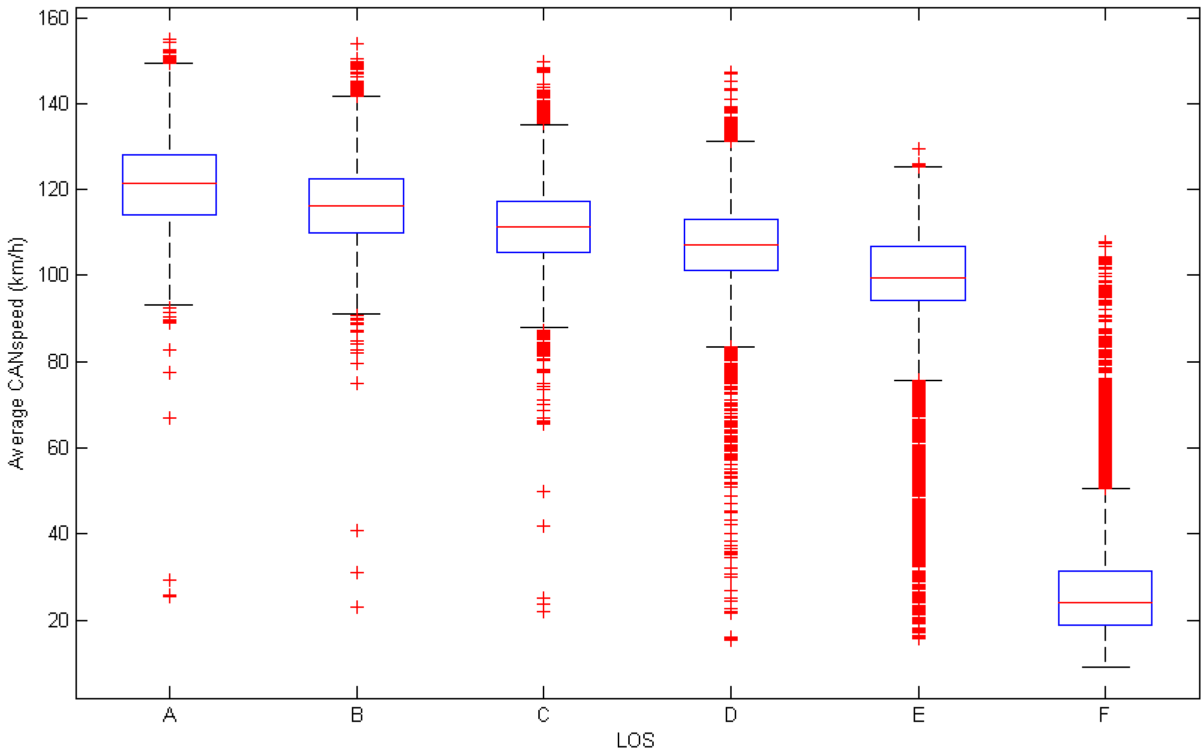

Figure 24: LOS ground truth 3 versus average speed. The average simulated CAN bus speed of the simulated vehicles passing a detection point is plotted against the Level of Service ground truth in the corresponding minute, determined with the presence of the simulated vehicles near the detection point.

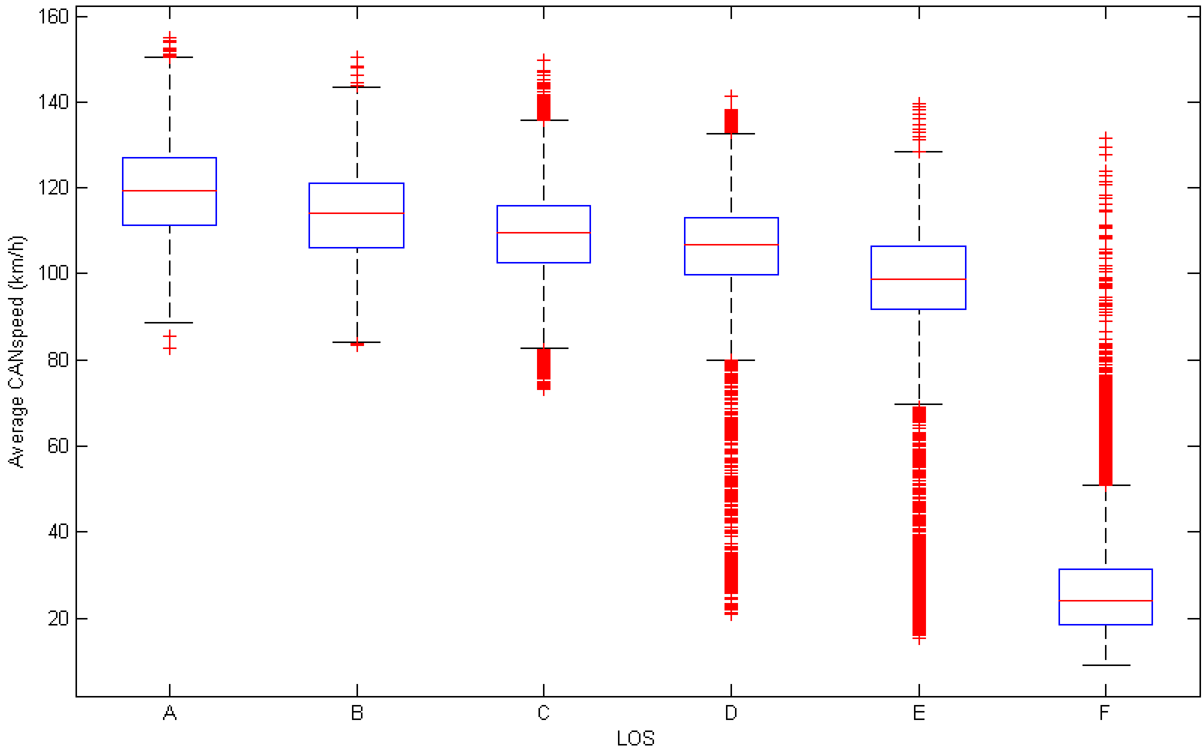

Figure 25: LOS ground truth 4 versus average speed. The average simulated CAN bus speed of the simulated vehicles passing a detection point is plotted against the Level of Service ground truth at that moment, determined with the presence of the simulated vehicles near the detection point.

Relationship between LOS and CAN bus parameters

The Spearman rank correlation coefficient has been calculated for the available simulated CAN bus parameters and the different LOS ground truths. The results are presented in Table 12. When comparing the R values in this table with the R values which were found for the correlation between the LOS and the real CAN bus parameters, in Table 9 and Table 10, it can be observed that the simulation data shows a stronger correlation between LOS and the CAN bus parameters. This is in line with the findings of the previous paragraph.

As expected, the correlation between the speed parameters and LOS ground truth 2 (based on the erroneous detector values) is slightly weaker than the correlation of LOS ground truth 1 (based on the perfect detector values), although the differences are very limited.

Surprisingly, the LOS ground truth based on the 1-minute aggregated density shows weaker correlations with the speed parameters than LOS ground truth 1 and 2. The R values for LOS ground truth 4 (which was thought to be the best LOS determination method) are however indeed larger than the other LOS ground truths.

Table 12: R values Spearman correlation between simulated CAN bus signals and several LOS ground truths

variable	LOS ground truth			
	1	2	3	4
avg_speed	-0,7221	-0,71402	-0,6872	-0,7638
stdev_speed	0,6108	0,60513	0,6078	0,6173
min_speed	-0,7285	-0,71986	-0,6949	-0,7673
max_speed	-0,7005	-0,69296	-0,6651	-0,7404

LOS estimation

Analogous to the process in 5.1.8, the simulated average speed data is used to estimate the LOS with single vehicle passages using multinomial logistic regression. The correct LOS estimation percentage per LOS ground truth is presented in Table 13.

Table 13: Correct LOS estimation percentage using average simulated CAN bus speed, for several LOS ground truths.

LOS ground truth			
1	2	3	4
52,4%	50,8%	51,5%	53,3%

In accordance with the correlation findings, the average speed from a passing vehicle is best suited to estimate LOS ground truth 4 (determined with the density at the moment of passage).
The correct LOS estimation with the real data was found to be 37.1% (section 5.1.8). The corresponding correct LOS estimation percentage (using LOS ground truth 2) with simulation is 50.8%. The consequences of this difference are discussed in the next paragraph. The visualizations of the LOS estimation functions, using multinomial logistic regression, are shown in Figure 70 - Figure 73, in Appendix H.

Model validity

In the previous paragraphs, we have shown that the correlation between LOS and the average speed is larger for the simulated data than for the real CAN bus data. Consequently, we have also seen that the correct LOS estimation percentage is larger with the simulated data. These differences are thought to be mainly caused by the difference in real driving behaviour and the simulated driving behaviour.
Another source of the differences between the simulation data and real data results is the inaccuracy of the real loop detectors. This inaccuracy will lead to erroneous LOS ground truth determination and is likely to cause a negative effect on the LOS estimation accuracy. The average size of the error is however not known. We have seen that the requirements of the NDW and findings in literature do not match. Using the greatest allowable MAPE, by NDW standards, we found only a small negative effect on the LOS estimation accuracy. But the actual error of the loop detectors could be larger, leading to a greater negative effect in the LOS estimations.
Another difference between the model and reality is the vehicles types. The real CAN bus data only contains data from one type of vehicle and a specific driver type (taxi driver), on roads with

multiple vehicles types. The model only produces data from one vehicle type, on a road with just that vehicle type.

Because of these differences, the LOS estimation with the traffic model performs better than the LOS estimation with real data. Therefore, we will consider the outcomes of the analysis with the traffic model as an upper boundary for the LOS estimation accuracy in real life.

5.2.3 LOS estimation with CAN bus data at variable penetration rates

In the preceding sections, the LOS has been estimated and the LOS estimation accuracy has been determined using simulated and real CAN bus parameters from single vehicle passages. In this and following sections, the LOS will be estimated using the simulated CAN bus average speed from multiple vehicle passages. This is done for several penetration rates.

Although in real life the overall penetration rate may be known (eg. the number of equipped vehicles in the Netherlands divided by the total amount of vehicles in the Netherlands), the penetration rate on a certain road section at a certain moment will not be known. With the assumption that the equipped vehicles are randomly distributed over the network, this is taken into account in the simulation analysis as follows. Given a certain penetration rate, the data from a passage is randomly included with a chance equal to the penetration rate (and otherwise excluded). All data from the passages which were included during a minute at a detection point are used to calculate the average speed during that minute at that detection point. Using this average value of the average speed of the considered (simulated uCAN equipped) vehicles, again LOS estimation functions are determined with multinomial logistic regression. The used LOS ground truths are 1 and 3. The LOS estimation functions for several penetration rates are visualized Figure 74 - Figure 86, in Appendix H. The certainty with which the LOS can be estimated, increases for increasing penetration rates. When an average speed below 85 km/h or above 125 km/h is measured, it can be stated with very large certainty that the occurring LOS is F/E or A respectively. For values in between, the accuracy of LOS estimation is much smaller.

The percentage of correct LOS estimations for different penetration rates, using both LOS ground truths, is displayed in Figure 26. In this figure only the minutes in which at least one of the equipped vehicle passes the detection point are considered. The correct LOS estimation percentage at very low penetration rates approaches the values which were found with the LOS estimation with single passages, which is as expected, since at very low penetration rates, the LOS estimation will often be based on a single vehicle passage. For increasing penetration rates, the LOS estimation accuracy increases up to over 70% for LOS ground truth 3 (determined with the average density in a minute). LOS ground truth 1 (determined with the 1-minute aggregated speed and intensity measurements from detection points) can be less accurately determined for all penetration rates.

In case of low penetration rates, it will often occur that during a minute none of the equipped vehicles will pass the detection point. In that case, no measurement values are available to estimate the LOS. The percentage of minutes for which no data is available is shown in Figure 27.

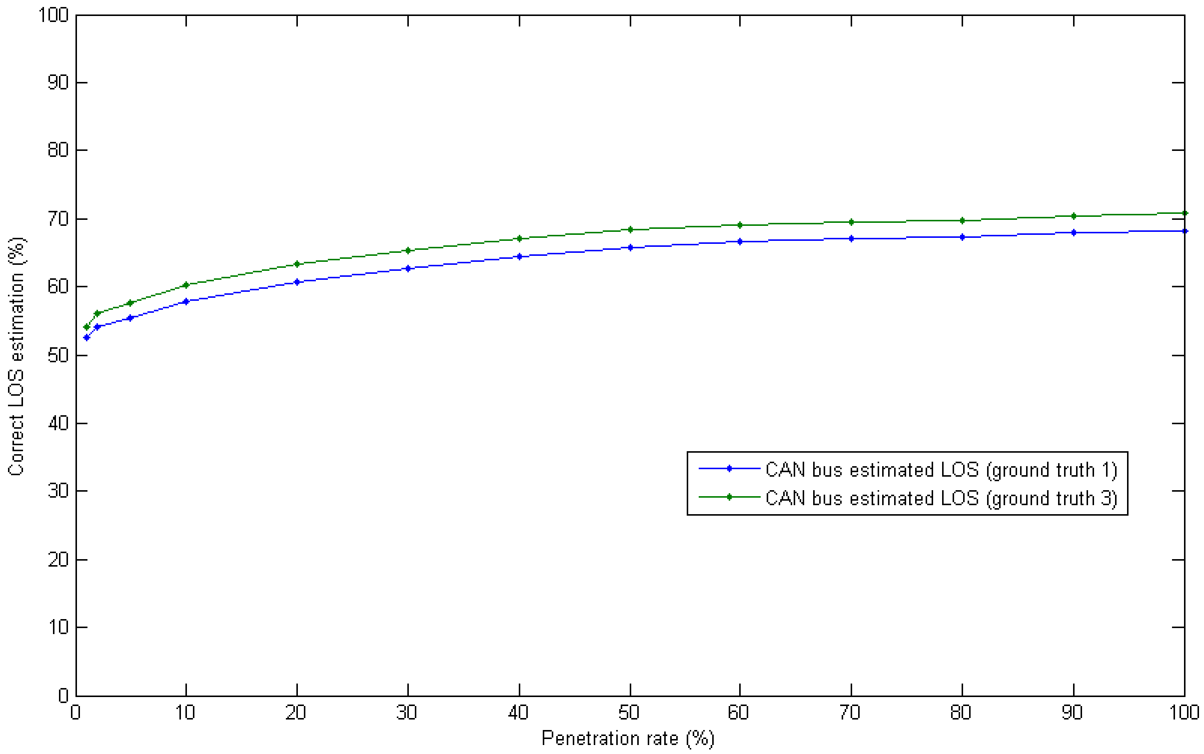

Figure 26: Penetration rate versus correct LOS estimation with simulated average CAN bus speed, for ground truth 1, determined with simulation loop detector data, and ground truth 3, determined with the vehicle presence, for minutes with available data.

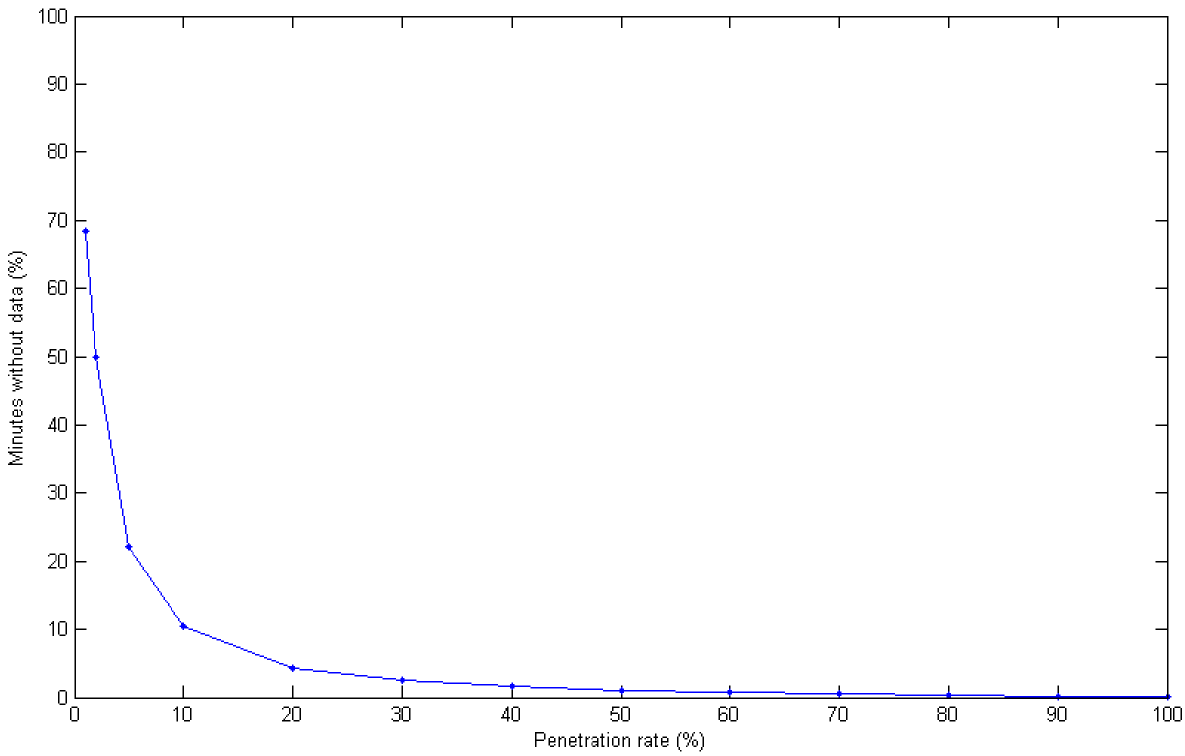

Figure 27: Penetration rate versus percentage of minutes without information

When determining the LOS estimation accuracy for all minutes, including the minutes for which no data is available, we obtain Figure 28. In this figure, all minutes without any equipped vehicle passages are considered to have an incorrect LOS estimation.

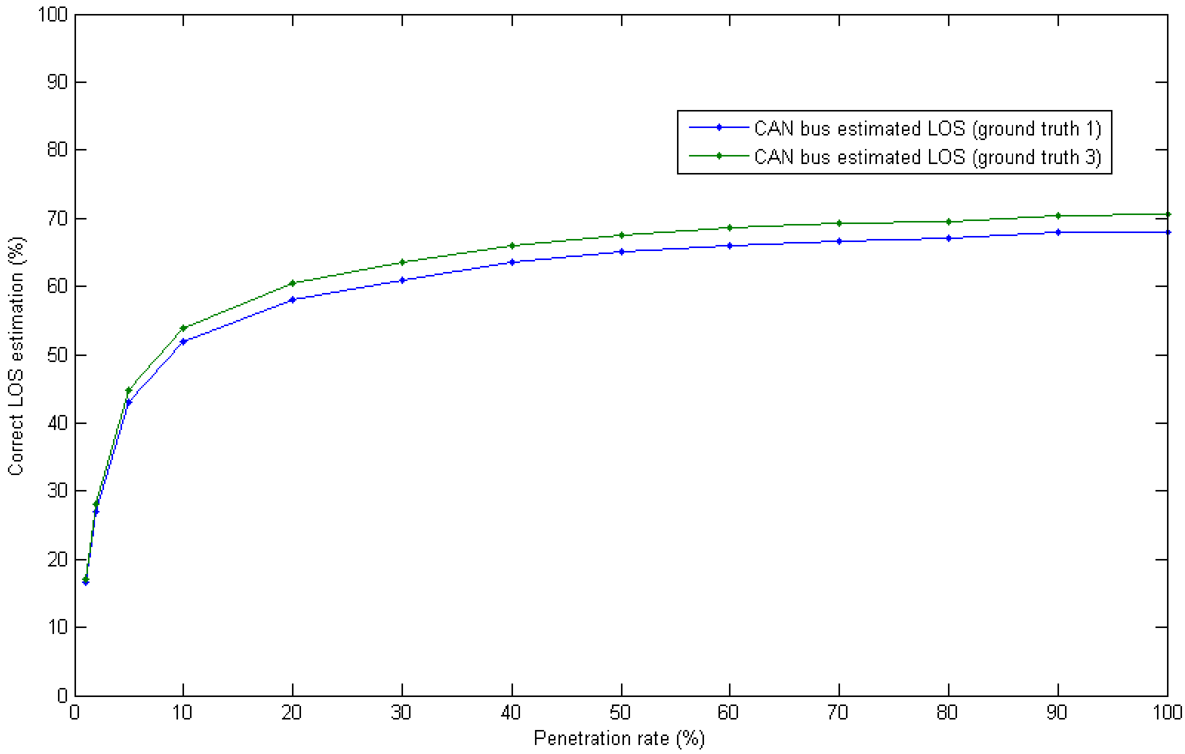

Figure 28: Penetration rate versus correct LOS estimation with simulated average CAN bus speed, for ground truth 1, determined with simulation loop detector data, and ground truth 3, determined with the vehicle presence, for all minutes.

5.2.4 LOS estimation with plain floating car data at variable penetration rates

When the penetration rate is sufficiently large, the density and thus the LOS can be estimated by observing the amount of vehicles on a road section and dividing it with the penetration rate:

$$k = m / L / n / P$$

with: k = density (vehicles/km/lane)

m = amount of observed vehicles on road section

L = length of road section (km)

n = number of lanes

P = penetration rate

As mentioned before, the exact penetration rate at a certain road section at a certain moment is not known. Again we assume the equipped vehicles are perfectly randomly distributed over the network, meaning the chance of a vehicle being an equipped vehicle is unrelated to the location of the vehicle and the presence of other equipped vehicles. With this assumption, we can use the simulation trajectory data to determine the accuracy of the LOS estimation with plain floating car data by again randomly drawing vehicles which will be considered as equipped vehicles with a chance equal to the penetration rate. The average density that these vehicles cause near a detection point (100m upstream and 100 downstream) during a minute is divided by the penetration rate to obtain the estimated density and LOS.

The LOS estimation accuracy for LOS ground truth 1 and 3, when only considering minutes in which at least one equipped vehicle was present on the road section, can be seen in Figure 29. The LOS estimation accuracy at very low penetration rates is less than 40% for both LOS ground truths. For increasing penetration rates, the accuracy increases, up to 100% for LOS ground truth 1 estimation and over 94% for LOS ground truth 1 estimation. LOS ground truth 3 estimation reaching an accuracy of 100% is trivial, since the LOS determined with this measurement method and LOS ground truth are determined in the same way. They are both based on the average density, caused by all vehicles, during a minute.

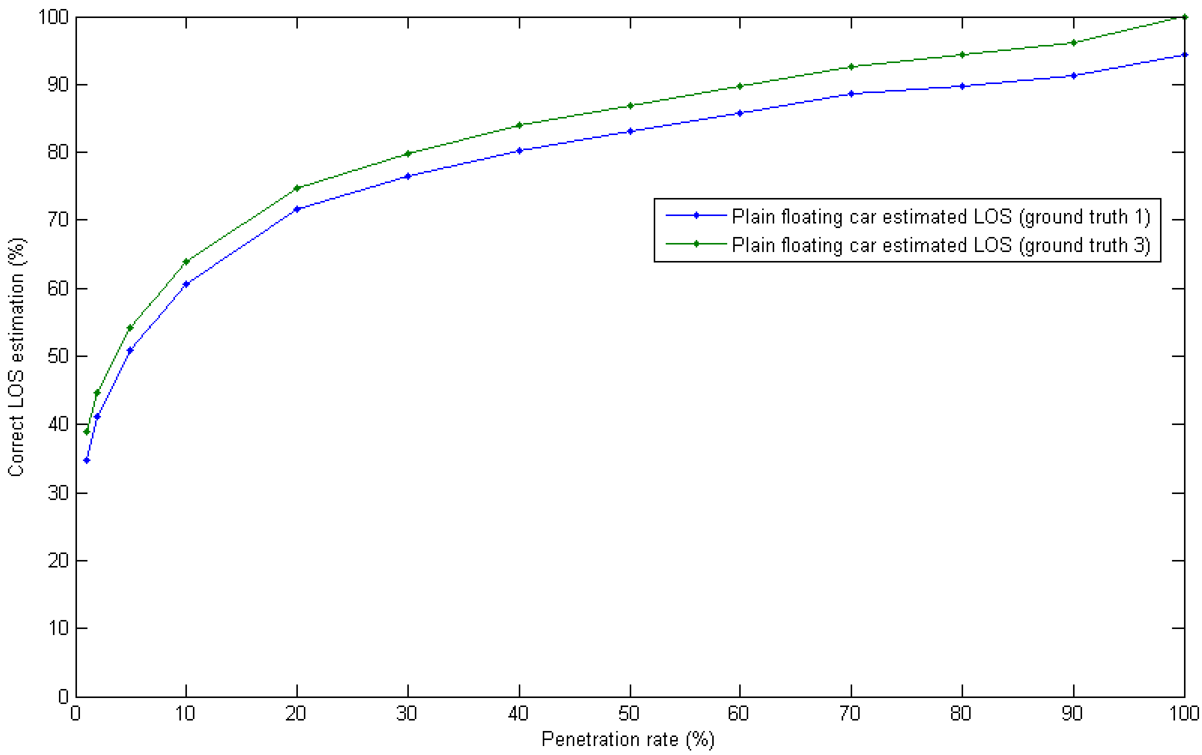

Figure 29: Penetration rate versus correct LOS estimation with simulated plain floating car data (containing only positional data), for ground truth 1, determined with simulation loop detector data, and ground truth 3, determined with the vehicle presence, for minutes with available data.

Also using this LOS estimation method, there will be minutes without any data from equipped vehicles. When these minutes are considered to result in an incorrect LOS estimation, the LOS estimation accuracies in Figure 30 are obtained.

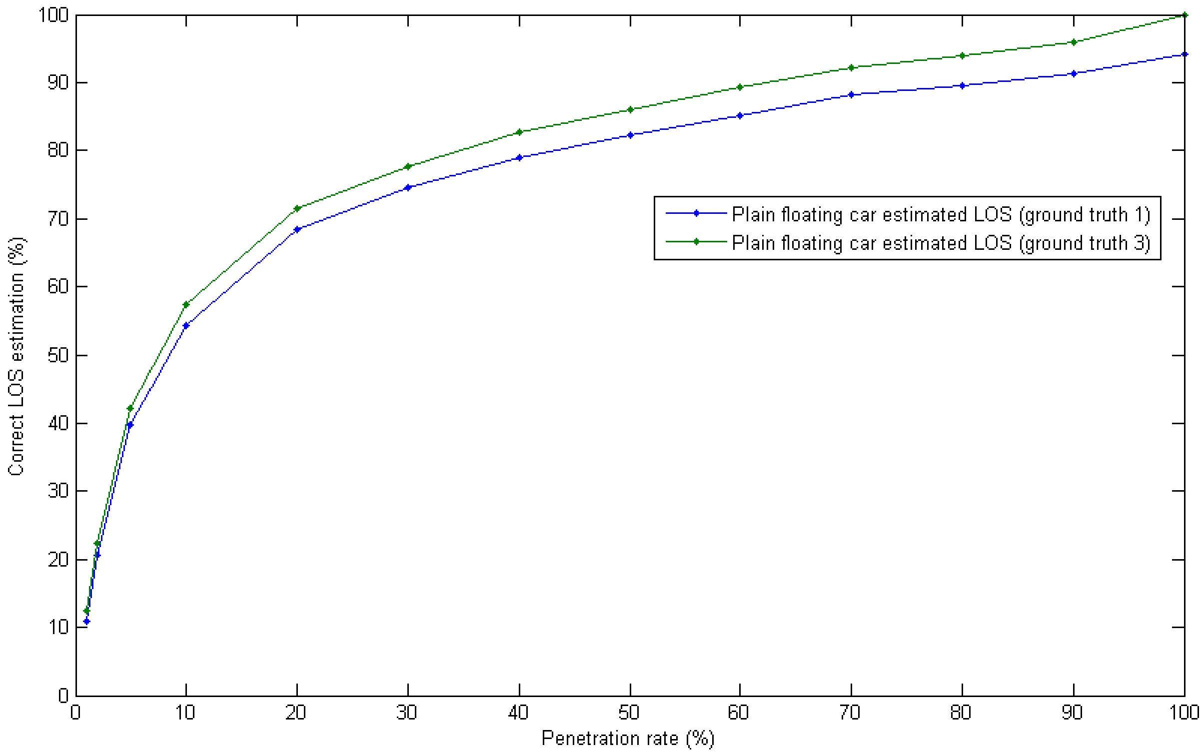

Figure 30: Penetration rate versus correct LOS estimation with simulated plain floating car data (containing only positional data), for ground truth 1, determined with simulation loop detector data, and ground truth 3, determined with the vehicle presence, for all minutes.

5.2.5 Method accuracy versus penetration rate

The figures from the previous paragraph have been combined in Figure 31 and Figure 32. Figure 31 shows the LOS estimation accuracy for both methods and for both LOS ground truths, for all minutes with available data. Figure 32 shows the same information, but for all minutes, also the minutes where no data is available.

Both figures show that LOS estimation with plain floating car data performs better than LOS estimation with CAN bus data for penetration rates over 8%. In fact, since the LOS estimation accuracy with simulated CAN bus data is considered to be an upper boundary of the real CAN bus LOS estimation accuracy, as mentioned in the section about Model validity in 5.2.2, it is likely that plain floating car data also outperforms CAN bus LOS estimation at penetration rates under 8%.

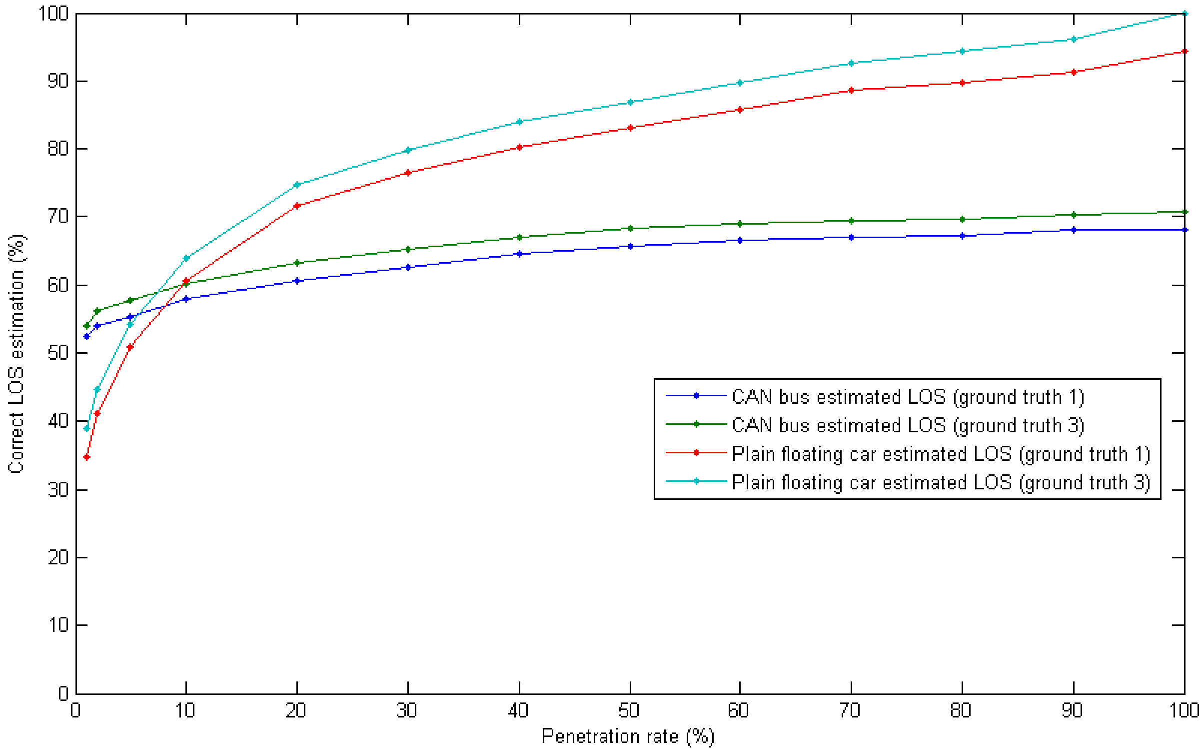

Figure 31: Penetration rate versus correct LOS estimation with simulated plain floating car data (containing only positional data) and average simulated CAN bus speed, for ground truth 1, determined with simulation loop detector data, and ground truth 3, determined with the vehicle presence, for minutes with available data.

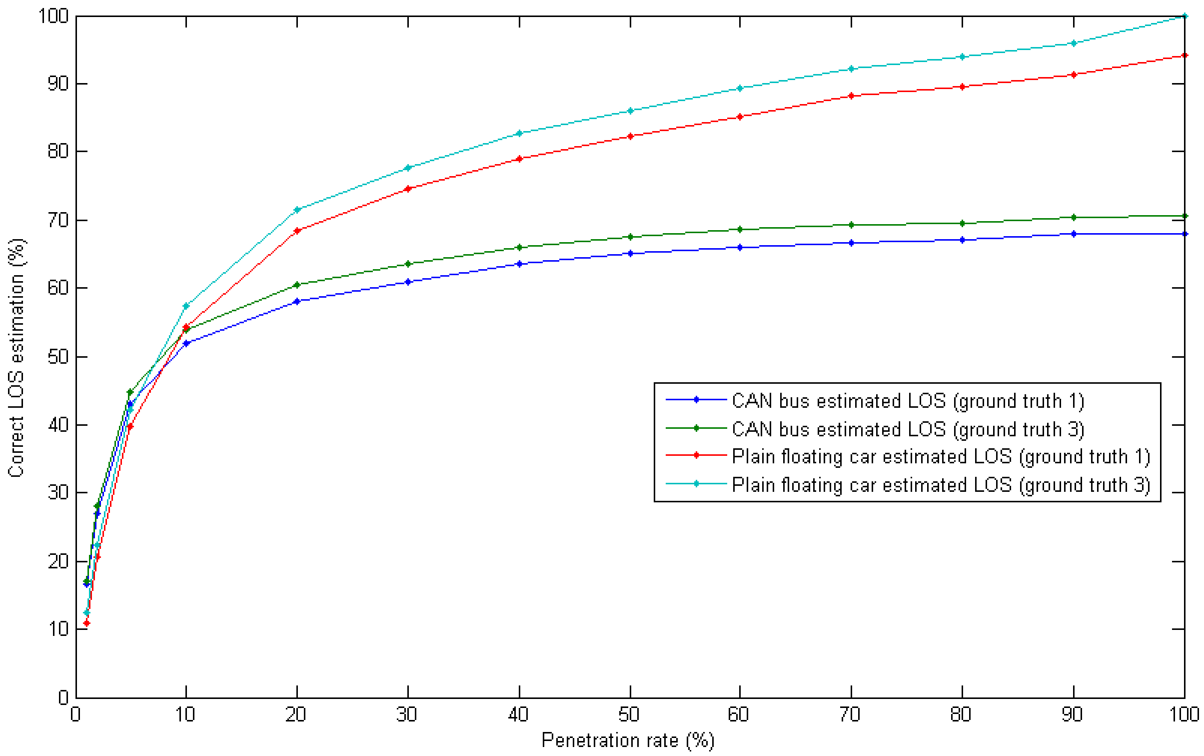

Figure 32: Penetration rate versus correct LOS estimation with simulated plain floating car data (containing only positional data) and average simulated CAN bus speed, for ground truth 1, determined with simulation loop detector data, and ground truth 3, determined with the vehicle presence, for all minutes.

6 Conclusions and recommendations

CAN bus data, containing data from in-vehicle sensors, combined with GPS signals, may offer a new method to measure traffic flow. In this study, the Level of Service estimation accuracy of CAN bus signals is investigated.

After identifying several driving behaviour factors which depend on Level of Service and exploration of the processed CAN bus data, several vehicles are selected to be used in this research. Using the CAN bus data of 35 vehicles of the same type over 2 months and loop detector reference data over the same period, the Level of Service estimation accuracy with CAN bus data from single vehicles is determined. A traffic simulation model is used to assess the Level of Service estimation accuracy with simulated CAN bus data from multiple vehicles, for various penetration rates. Also, the traffic simulation model is used assess the Level of Service estimation using only the simulated positional data, to be able to compare the Level of Service estimation performance of CAN bus data and GPS data.

Using the findings of the previous chapter, in this section, the research questions will be answered and recommendations are provided.

6.1 Conclusions

The main research question of this master thesis is:

Is it possible to accurately estimate the Level of Service with in-vehicle sensor floating car data?

To answer this main question, first the answers to the sub questions will be provided.

6.1.1 Sub questions

1. *Which main driver behavior characteristics are dependent on Level of Service?*

In chapter 3, a short literature survey was used to identify the following main driver behavioural factors depending on Level of Service: speed, speed variation, lane change rate and distance headway. This knowledge was used to select data from vehicles which contain CAN bus signals which are most likely to be able to be used for LOS estimation.

2. *Which CAN bus signals are provided by the uCAN equipped vehicles?*

The complete list with CAN bus signals which are provided by the uCAN equipped vehicles is found in Table 20 in Appendix E. A total of 19 different CAN bus signals are provided by the uCAN equipped vehicles in the Smart In-Car project, but not every uCAN equipped vehicle provides all of these signals. The main CAN bus signals provided by the 35 vehicles selected for this study are: speedometer, brake usage, steering position, rpm, indicator usage and fuel consumption.

3. *What is the accuracy of Level of Service estimation with in-vehicle sensor floating car data from single vehicles?*

In section 5.2.2, the Level of Service was estimated using the real CAN bus data from 7185 passages of single vehicles and loop detector data from the specific locations to determine the LOS ground truth. The most accurate LOS estimation functions were found by using only the average CAN bus speed of the single vehicles. A correct LOS estimation rate of 37% was found.

4. *What is the accuracy of Level of Service estimation with in-vehicle sensor floating car data from multiple vehicles?*

In section 5.2.3, using only the average speed of multiple vehicles from simulation data , the found LOS estimation accuracy with multiple vehicles ranged from 52% at 1% penetration rate and 70% at 100% penetration rate. Because the simulated CAN bus data was found to provide a higher correct LOS estimation rate with single vehicles than the real CAN bus data (52% versus 37% respectively), these found accuracies are considered to be an upper boundary for the accuracy which could be reached with real CAN bus data.

5. *Does in-vehicle sensor floating car data provide a better LOS estimation than plain floating car data?*

In section 5.2.4, the LOS estimation accuracy at various penetration rates was determined by using the vehicle trajectories from simulation data. When comparing the results with the LOS estimation with in-vehicle sensor floating car data, we found that CAN bus data may provide a better LOS estimation than plain floating car data at penetration rates lower than 8%. Generally, LOS estimation with plain floating car data, containing only the vehicle position, outperforms LOS estimation with CAN bus data.

6. *Does CAN bus data provide a better LOS estimation than loop detector data?*

In section 5.1.3, we estimated, using the NDW requirements, that the correct LOS estimation rate with loop detectors is around 89%. Since the maximum correct LOS estimation rate using CAN bus data was found to be 70% at most in section 5.2.3, the CAN bus data does not provide a better LOS estimation than loop detector data.

7. *Does the transmission frequency of the CAN bus signals influence the LOS estimation accuracy?*

In section 5.1.7, we found that the frequency of the CAN bus signals does have an effect on the correlation with LOS, but that this influence is small when the amount of data points during a passage is sufficiently large.

6.1.2 Main research question

With the answers from the sub questions, also an answer for the main research question can be formulated.
No, it is not possible to accurately estimate LOS with the CAN bus signals which were present in the available dataset. The correct LOS estimation rate with CAN bus signals ranges up to 70% at 100% penetration, which does not comply with NDW requirements. Also LOS estimation with loop detectors and plain floating car data both perform better, especially at high penetration rates.
Note that the this conclusion relates to, and is only valid for, two-lane motorways with a 120 km/h speed limit and the CAN bus signals which were available in the data set, since only that road type and those CAN bus signals where considered in this research.

6.2 Discussion

In this section, we will reflect on the performed study.

6.2.1 Restrictions

In this study, several assumptions have been made, several conditions have been encountered and several decisions have been made which affect the study results and restrict the validity of the study. These will be discussed in this paragraph.

First of all, the available CAN bus data only contained a small part of the total CAN bus signals which are used in modern vehicles (compare Table 1 with Table 19). If more CAN bus signals are available, the LOS estimation might be more accurate, although none of these new CAN bus signals are expected to provide a significant improvement for LOS estimation. There is however at least one in-vehicle signal which is not mentioned in Table 1, but is measured by some vehicles and should be suitable to be used to estimate density. As stated in section 3.1.2, the average distance headway of vehicles is directly related to the density and Level of Service. If this distance headway can be obtained from the sensors of adaptive cruise control or collision warnings systems, this data can be used to estimate the LOS.
Also, the loop detector data which was used to determine the LOS ground truth is known to be inaccurate, as was mentioned in section 1.2. In section 5.1.3, we estimated the correct LOS estimation rate of loop detectors at 89% using the requirements of the NDW, which are mentioned in Appendix B. The effect of this inaccuracy on the correct LOS estimation rate was found to be

small by using simulation data in section 5.2.2, although the real inaccuracy of the loop detectors is unknown and may be larger than the assumed inaccuracy.

The simulation model is another point of attention. In section 5.2.2 we discussed the model validity. The correct LOS estimation rate with single vehicles was found to be 37% and 52%, using real CAN bus data and simulated CAN bus data respectively. Three sources for this difference were identified:

- The differences in real driving behaviour and the simulated driving behaviour.
- The difference in the simulated and real loop detector accuracy.
- The usage of only one vehicle type (passenger cars) in the model, while multiple vehicle types are present in reality.

Because no detection point data was available on secondary roads near Eindhoven, only road sections on the motorways were included in this research. This, together with the finding that the uCAN equipped vehicles used the motorways sporadically, led to the decision to combine the loop detector data and uCAN equipped vehicle passages for all road sections. Although only road sections with the same cross section were used (2 lanes with a shoulder lane), we cannot exclude the possibility that there are small differences in the road sections which influence the local driving behaviour. Should this be the case, then a larger accuracy of the LOS estimation with CAN bus data at a single location may be found. Also, because only one type of road section was used, the LOS estimation capabilities of CAN bus data for other road sections is not known, although it is not expected that these will differ much.

On all road sections of the motorways near Eindhoven which were included in the research, the speed limit is set at 120 km/h. The LOS estimation accuracy at roads with another speed limit is therefore not known. But it is expected that the LOS estimation accuracy will be higher for roads with a higher speed limit and lower for roads with a lower speed limit, since the free flow speed at the latter will be lower and closer to the speed at capacity, making differentiation of Level of Service based on speed less accurate.

Other conditions which affect the study results are the availability of only one vehicle type (Mercedes E-class 212) and only one driver type (taxi driver).

During the research, it was decided to adopt the Level of Service boundaries from the 2010 Highway Capacity Manual. When some of the CAN bus parameters values was plotted against the LOS (section 5.1.6), it was found that the average speed at LOS F is higher than 80 km/h. This should have made us realize that the used LOS boundaries yielded LOS categories which do not reflect the conditions which these categories are supposed to reflect (Table 2). When more appropriate LOS boundaries would have been determined, this could have improved the LOS estimation accuracy.

6.2.2 Level of Service and other traffic quantities

In this study, the suitability of CAN bus data for LOS estimation was investigated. In this paragraph, the findings of this study are used to discuss the plausibility of suitability of CAN bus data for estimation of other traffic quantities: density, intensity and speed.

At the start of this study, it was decided to investigate the estimation capabilities of CAN bus signals for Level of Service, instead of intensity, because it was expected the intensity estimation accuracy with CAN bus data from single vehicles would be very small. If the results from this study indicate that LOS can be accurately estimated with CAN bus signals, the study could be extended to investigate the CAN bus signal suitability for intensity estimation. In this study, it was found that the LOS estimation accuracy with single vehicle CAN bus data is small. To provide insight in the accuracy of density estimation, in the next section, the results of density estimation with simulated data are briefly discussed. It will be shown that the corrected mean average percentage error (MAPE) of the density estimation ranges from 31% at 1% penetration to 24% at 100% penetration.

In section 5.1.7 and Table 29 in Appendix H, we found that the CAN bus signals show the strongest correlation with speed, followed by density, while the correlation with intensity was the weakest. Based on these findings, it is expected that intensity estimation will perform less accurate than density estimation and speed estimation will perform more accurate than density estimation. Therefore, it is plausible that CAN bus data will prove to be unsuitable for intensity estimation, while it could be suitable for speed estimation. However, since the GPS data can also be used to measure speed, the CAN bus data is not likely to significantly contribute to speed estimation.

6.2.3 Density estimation

In this section, the density estimation accuracy with simulated CAN bus data of multiple vehicles for different penetration rates is presented.

The results for the density estimation were obtained with the same methods as the LOS estimation in this report. However, instead of multinomial logistic regression, which was used because of the ordinal scale of Level of Service, the curve fit tool, included in Matlab, is used to determine a function which best fits the found points in the CAN bus speed and density plane. The fitted function is used to estimate the 1 minute aggregated density with the average CAN bus speed of the vehicles passing the detection point during that minute.

To assess the density estimation accuracy using CAN bus signals and plain floating car data, the mean absolute percentage error of the density estimation is calculated. The absolute percentage errors (APE's) are found to be quite large, especially when the density is low. The regular APE definition is:

$$APE = abs(k_{estimated} / k_{groundtruth} - 1) * 100\%$$

where: APE is the absolute percentage error

$k_{groundtruth}$ is the density ground truth

$k_{estimated}$ is the estimated density

Using this definition, $k_{estimated}=1$ and $k_{groundtruth}=8$ yields an APE of abs(1/8-1)*100% = 87.5 %

and $k_{estimated}=8$ and $k_{groundtruth}=1$ yields an APE of abs(8/1-1)*100% = 700 %.

When $k_{groundtruth}$ is small, or even zero, the found APE is very large and can even be infinite when $k_{groundtruth}=0$. These extreme APE values strongly influence the MAPE and give results which are hard to interpret. Therefore, an alternative MAPE definition is used, which corrects this aspect. The absolute percentage error of the density estimations is calculated with:

$$APE_{corrected} = abs(k_{groundtruth} / k_{estimated} - 1) * 100\% \qquad \text{for } k_{estimated} > k_{groundtruth}$$

$$APE_{corrected} = abs(k_{estimated} / k_{groundtruth} - 1) * 100\% \qquad \text{for } k_{estimated} \leq k_{groundtruth}$$

where: $APE_{corrected}$ is the corrected absolute percentage error

$k_{groundtruth}$ is the density ground truth

$k_{estimated}$ is the estimated density

Using the corrected APE definition, in both cases ($k_{estimated}$=1, $k_{groundtruth}$=8 and $k_{estimated}$=8, $k_{groundtruth}$=1) an APE of 87.5 % is found. The average value of $APE_{corrected}$ gives us the corrected mean absolute percentage error ($MAPE_{corrected}$).

In Figure 33 the $MAPE_{corrected}$ of the density estimation using CAN bus signals and plain floating car data is presented, using two density ground truths and for two estimation methods, similar to the approach in section 0.

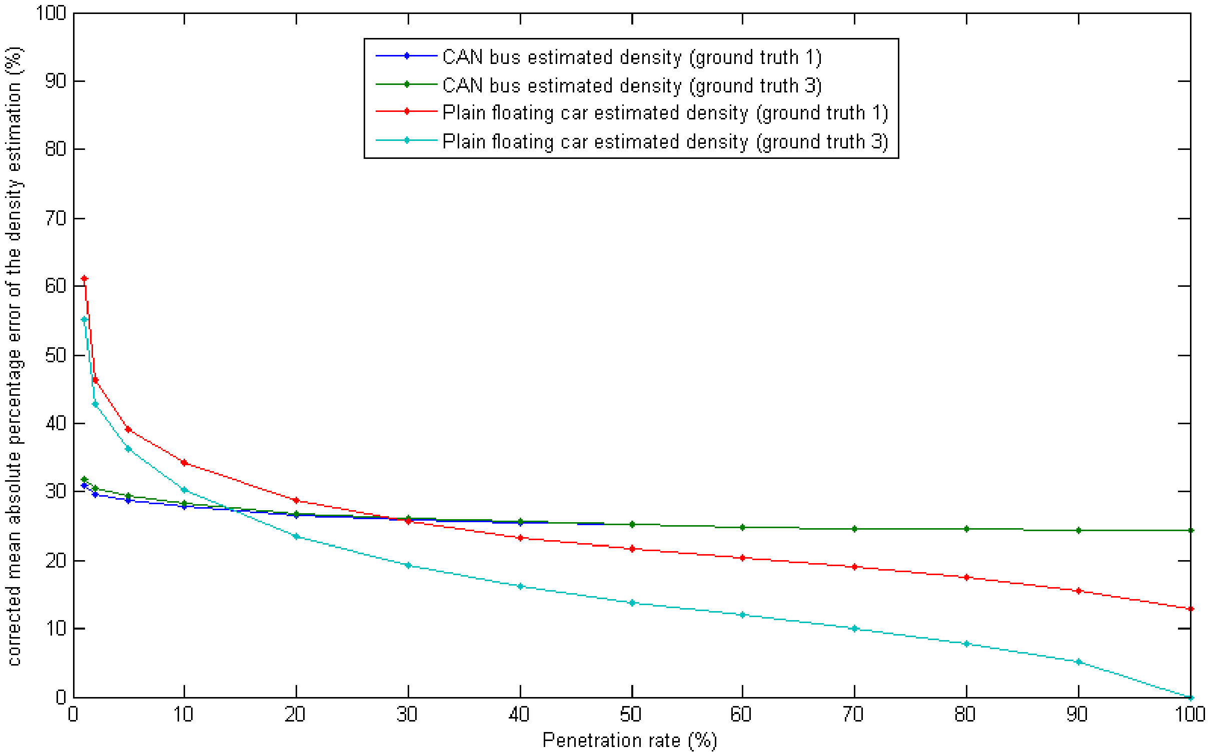

Figure 33: Corrected Mean Absolute Percentage Error of the density estimation with simulated CAN bus data and simulated plain floating car data (only containing positional data) using multiple vehicles, for different penetration rates. The 1-minute aggregated density ground truth is determined with the simulated detection point data (ground truth 1) and with the real density near the detection point, using the simulated vehicle trajectories (ground truth 3).

6.3 Recommendations

6.3.1 Further research on CAN bus signal usability for other use cases

It is recommended to also examine the other use cases which are mentioned in the introduction of this thesis and to look for other applications which could benefit from CAN bus signals. The processed and analyzed CAN bus data in this study show that CAN bus signals provide a wealth of information which may be used for many applications.

6.3.2 Further research on the CAN bus signals

Supporting the further research on the other use cases, it is recommended to do more research on the properties of the several CAN bus signals. Knowledge of the accuracy and reliability of the CAN bus signals is valuable. Also differences between the CAN bus signals of different vehicle models should be investigated.

6.3.3 Further research on the costs of a CAN bus signal traffic measurement system

Once applications have been found which could benefit from CAN bus signals, it is useful to compare the costs of a CAN bus signal traffic measurement system with other traffic measurement systems.

6.3.4 Further research with other CAN bus signals

Only a small part of the CAN bus signals which could theoretically be available, as shown in Table 1, is available in the CAN bus data set from the Smart In-Car project. Most of the CAN bus signals which are not available are unlikely be strongly correlated with LOS, but one of them might be able to significantly improve the LOS estimation. As stated in section 3.4, the average distance headway of vehicles is directly related to the density and Level of Service. If this distance headway can be obtained from the sensors of adaptive cruise control or collision warnings systems, this data can be used to estimate the LOS. Therefore, it is recommended to investigate the LOS and density estimation capabilities of the distance headway CAN bus signal, when this data becomes available.

6.3.5 Suggestions for improvements and extensions of the research

Some aspects of this research could be improved or be extended. Because the expectancy that these improvements will significantly change the outcome is low, it is only recommended to actually implement these improvements when CAN bus signals have become available which are able to estimate the LOS much more accurately, like described in the previous paragraph.

More accurate ground truth data

As we have shown in section 5.1.3, the accuracy of the LOS ground truth determined with loop detector data is likely to be small. The LOS estimation capabilities of CAN bus signals can be better assessed if the LOS ground truth is determined more accurately.

Other LOS boundaries

Figure 13 shows the average speeds in LOS F are higher than 80 km/h. This is not according to the definition of LOS F, which should reflect a congested traffic state. Apparently, the definition of the boundaries of the LOS is not optimal for the conditions encountered in this research. With redefined LOS boundaries, the CAN bus signals might show larger correlations and improved LOS estimation accuracy.

Other vehicle types and drivers

As mentioned before, the used CAN bus data set only includes taxis. Other vehicle types may show different results.

Improved simulation model validity

In section 5.2.2, the model validity was discussed. The research can be improved by improving the model validity. This can be done by calibration of the model parameters and by adding more vehicle types.

Field test with higher penetration rates

A field test with higher penetration rates can be performed to check for compliance of the results with the simulation results.

Other speed limits

The considered road sections in this research only included roads with a speed limit of 120 km/h. As was already mentioned in section 2.5.1, the LOS estimation accuracy may be affected by the occurring speed limit. Therefore, also road sections with other speed limits should be included in the research.

Other road types

While LOS estimation at motorways is useful, LOS estimation at secondary roads where normally no traffic detectors are present may be even more valuable. Therefore, it is recommended to also investigate the LOS estimation capabilities at other road types. Also, the LOS at motorways with more than 2 lanes or at weaving sections or on-ramps and off-ramps can be considered.

6.3.6 Level of Service category boundaries in the Netherlands

As we have seen in section 5.1.6, the definition of Level of Service categories as given in Table 2 and the corresponding densities, does not seem to match with the occurring traffic conditions at the observed motorways. The found average vehicle speeds at LOS E and LOS F were much higher than would be expected according to their definitions. Apparently, traffic in the Netherlands is able to maintain higher speeds at high densities. This is in line with the fact that the capacity of Dutch motorways are generally higher than the capacity values of motorways given in the Highway Capacity Manual [27]. Therefore, it is recommended to more accurately determine the Level of Service category boundaries for the Netherlands, to obtain Level of Service categories which properly represent the traffic conditions they should represent.

References

[1] Klein, L.A., Mills, M.K., Gibson, D.R.P., Traffic Detector Handbook: Third Edition - Volume I, Federal Highway Administration, 2006.

[2] Briedis, P., Samuels, S., The accuracy of inductive loop detectors, ARRB Group Limited, 2010.

[3] Heidemann, J., Silva, F., Wang, X., Giuliano G., Hu, M., Sensors for Unplanned Roadway Events - Simulation and Evaluation, University of Southern California, 2005.

[4] Robert Bosch GmbH, CAN: das Netzwerk für die Elektronik im Kraftfahrzeug, www.semiconductors.bosch.de/media/pdf/canliteratur/can.pdf.

[5] Robert Bosch GmbH website, www.semiconductors.bosch.de/en/ipmodules/can/can.asp.

[6] CAN in Automation website, www.can-cia.de.

[7] LaQuSo, Project report Smart In-Car, Stan Ackermans Institute, 2012.

[8] Swinkels, J., Presentation continuous road condition monitoring using in-car sensors in (potentially) any commodity car, Rijkswaterstaat, 2012.

[9] Highway Capacity Manual, Transportation Research Board, 2000.

[10] Transportation Research Board, Highway Capacity Manual, Special Report 209, 1994, pp. 3-9.

[11] Transportation Research Board, Highway Capacity Manual 2010, TR News 273, 2011, pp. 45-48.

[12] Choocharukul, K., Sinha, K.C., Mannering, F.L., User perceptions and engineering definitions of highway level of service: an exploratory statistical comparison, Transportation Research Part A, Vol. 38, 2004, pp. 677–689.

[13] Brilon, W., Traffic Flow Analysis Beyond Traditional Methods, 4th International Symposium on Highway Capacity, TRB Circular EC018, 2000, pp. 26-41.

[14] Cameron, R., G3 F7: An Expanded LOS Gradation System, ITE Journal, Vol. 66, No. 1, 1996, pp. 40-41.

[15] Landman, R.L., Hegyi, A., Hoogendoorn, S.P., Service level-oriented route guidance in road traffic networks, Proceedings of the 2011 IEEE Conference on Intelligent Transportation Systems and Control, 2011, pp. 1120-1125.

[16] Schakel, W.J., Knoop, V.L., Van Arem, B., LMRS: An Integrated Lane Change Model with Relaxation and Synchronization, Transportation Research Board Annual Meeting Technical Papers, 2011, pp. 1-21.

[17] Hoogendoorn, R.G., van Arem, B., Hoogendoorn, S.P., Brookhuis, K.A., Applying the Task-Capability-Interface Model to the Intelligent Driver Model in Relation to Complexity, Transportation Research Board Annual Meeting Papers, 2013, pp. 1-21.

[18] Treiber, M., Hennecke, A., Helbing, D., Congested traffic states in empirical observations and microscopic simulations, Physical Review E, Vol. 62, pp. 1805-1824, 2000.

[19] Kesting, A., Treiber,M., Helbing, D., Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity, Philosophical Transactions of the Royal Society A:

Mathematical, Physical and Engineering Sciences, Vol. 368, No. 1928, 2010, pp. 4585-4605.

[20] Drake, J.S., Schofer, J.L., May, A.D., A Statistical Analysis of Speed Density Hypothesis, Highway Research Record 154, Highway Research Board, 1967, pp. 53-87.

[21] Tang, C.F., Jiang, R. and Wu, Q.S., Extended speed gradient model for traffic flow on two-lane freeways, Chinese Physics, Vol. 16, No. 6, 2007, pp. 1570-1576.

[22] Knoop, V.L., Hoogendoorn, S.P., Shiomi, Y., Buisson, C., Quantifying the Number of Lane Changes in Traffic: Empirical analysis, Transportation Research Board 91st Annual Meeting, 2011.

[23] Ko, J., Guensler, R., Hunter, M., Li, H., Instrumented Vehicle Measured Speed Variation and Freeway Traffic Congestion, Applications of Advanced Technology in Transportation, 2006, pp. 356-361.

[24] K. Tindell, K., Burns, A., Wellings, A.J., Calculating Controller Area Network (CAN) Message Response Times, Control Engineering Practice, Vol. 3, No. 8, 1995, pp. 1163-1169.

[25] Bureau Onderweg, Onnauwkeurigheid en onbetrouwbaarheid van voertuigdetectiedata, RWS Dienst Verkeer en Scheepvaart, 2009.

[26] Van Rij, M., Henkens, N., Memo Conclusies kwaliteitsmetingen NDW, Grontmij, 2009.

[27] Hoogendoorn, S.P., Traffic Flow Theory and Simulation reader, TU Delft, 2007, pp. 108.

[28] Popov, A., Babuska, R., Hegyi, A., Werner, H., Distributed Controller Design for Dynamic Speed Limit Control Against Shock Waves on Freeways, Proceedings of the 17th International Federation of Automatic Control World Congress, 2008, pp. 14060-14065.

[29] Tampère, C., Corman, F., Himpe, W., Invloed van datakwaliteit op Dynamisch Verkeersmanagement, 2011.

[30] Miltenburg, P., Bijlage 2 bij kwalitatieve analyse functioneren AID: Eerste verkenning functioneren filestaart waarschuwing, RWS Intranet, 2009.

[31] Floor, M., Bedienhandleiding Dynamax webportal en regelingen, RWS Intranet, 2012.

[32] Klein, L.A., Mills, M.K., Gibson, D.R.P., Traffic Detector Handbook: Third Edition, Volume I, Federal Highway Administration, Publication No. FHWA-HRT-06-108, 2006.

[33] Barto, M., Handleiding voorbereiding en uitvoering toeritdoseerinstallaties van de 3e generatie, RWS Intranet, 2007.

[34] The Vehicle Detector Clearinghouse, A Summary of Vehicle Detection and Surveillance Technologies used in Intelligent Transportation Systems, 2007.

[35] Minnesota Department of Transportation, NIT Phase II Evaluation of Non-Intrusive Technologies for Traffic Detection, Final Report, 2002.

[36] Middleton, P.E., Parker, R., Longmire, R., Vehicle Detector Evaluation, Texas Transportation Institute, Report 2119-1, 2002.

[37] Minge, E., Kotzenmacher, J., Peterson, S., Evaluation of Non-Intrusive Technologies for Traffic Detection, Minnesota Department of Transportation, 2010.

[38] Grone, B.W., An Evaluation of Non-Intrusive Traffic Detectors at the NTC/NDOR Detector Test Bed, Civil Engineering Theses, Dissertations, and Student Research, Paper 47, 2012.

[39] Middleton, D., Parker, R., Initial Evaluation of Selected Detectors to Replace Inductive Loops on Freeways, Texas Transportation Institute, Report 1439-7, 2000.

[40] Cheung, S.Y., Coleri, S., Dundar, B., Ganesh, S., Tan, C.-W., Varaiya, P., Traffic Measurement and Vehicle Classification with a Single Magnetic Sensor, Transportation Research Record: Journal of the Transportation Research Board, No. 1917, 2005, pp. 173–181.

[41] Robert, K., Video-based traffic monitoring at day and night: Vehicle Features Detection and Tracking, Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems, 2009, pp. 285-290.

[42] Urazghildiiev, I., Ragnarsson, R., Wallin, K., Rydberg, A., Ridderström, P., Öjefors, E., A Vehicle Classification System Based on Microwave Radar Measurement of Height Profiles, Proceedings of the RADAR Conference, 2002, pp. 409-413.

[43] Zwahlen, H.T., Russ, A., Oner, E., Parthasarathy, M., Evaluation of Microwave Radar Trailers for Nonintrusive Traffic Measurements, Transportation Research Record: Journal of the Transportation Research Board, No. 1917, 2005, pp. 127-140.

[44] NDW website, www.ndw.nu.

[45] Heijligers, B., Vukovic, D., Telman, J., NDW Toetsingsmethode, TNO-rapport TNO-034-DTM-2009-01939, 2009.

[46] Modsching, M., Kramer, R., ten Hagen, K., Field trial on GPS Accuracy in a medium size city: The influence of builtup, Proceedings of the 3rd Workshop on Positioning, Navigation and Communication, 2006, pp. 209-218.

[47] Schultz van Haegen, M.H., Eindbeeld verhoging maximumsnelheid, letter to the Dutch Parlement,www.rijksoverheid.nl/onderwerpen/wegen/documenten-en-publicaties/kamerstukken/2012/06/26/eindbeeld-verhoging-maximumsnelheid.html, 2012.

Appendix A

1 Common applications

In this paragraph, a description of commonly used traffic management applications will be given. The necessary measured traffic flow variables will be discussed. Like in the previous paragraph, the contents of this paragraph is meant as an overview, not as a complete summary of all available traffic management applications. With the obtained information, we will be able to propose several use cases in the next chapter. Besides the traffic management applications that will be described in the following subparagraphs, there are many other applications. Some of them will be briefly mentioned here. Traffic jam detection can, besides for warning upstream road users which will be discussed in the next paragraph, also be used for traffic information applications. Shockwave detection can be used to dissolve the shockwave by adjusting maximum speeds with variable message signs. [28] Besides warning for traffic jams, variable message signs can also be used in combination with local traffic systems to warn road users for opening bridges. Also, in case of adverse weather conditions like rain, snow, ice formation and fog variable message signs can be used to warn the road users. These systems use local weather sensors as a data source. Other traffic management applications on motorways include: peak lanes, plus lanes, obstacle detection, high occupancy vehicles lanes and road pricing. [29]

1.1 Traffic jam detection

Traffic jam detection is one of the most common traffic management applications. In the Netherlands, Automatic Incident Detection (AID), which is part of the Motorway and Traffic Managementsystem (MTM2), is used to detect the traffic jams. AID warns road users for downstream traffic jams. AID does not use the common available 1-minute aggregated data. It uses the individual vehicles observations from inductive loop detectors to calculate a running average of the speed. If this speed drops below certain boundaries (often 35 km/h) the AID gets activated and the variable message signs upstream of the loop detectors will start showing adapted speed limits. When the speed exceeds a certain boundary again (often 50 km/h) the AID will be turned off and the adapted speed limits will disappear. [30]

The formula that is used to calculate the running average of the speed is:

$P_{nieuw} = \alpha * P_{gemeten} + (1-\alpha) * P_{oud}$

with: $\alpha = \alpha_{acc}$ if $P_{gemeten} < P_{oud}$

$\alpha = \alpha_{dec}$ if $P_{gemeten} > P_{oud}$

common values for α_{acc} and α_{dec} are 0.15 and 0.40 respectively.

1.1 Dynamic maximum speeds

Dynamic maximum speeds (Dynamax) can be used for several reasons. Maximum speeds can be increased to reduce travel times or to increase capacity. Maximum speeds can be decreased to reduce noise or pollution, to increase capacity or to increase safety. Listed below are some of the Dynamax control types which are used in the Netherlands [31]:

- Increasing the maximum allowed speed when the intensity is low, to decrease travel times (A1 between Bussum and Muiderberg).
- Decreasing the maximum allowed speed when the intensity gets higher than a certain level or the average speed below a certain level.
- Increasing the maximum allowed speed during the night (22:00 – 7:00, A10 West).
- Increasing the maximum allowed speed from 80 km/h to 100 km/h during the rush hour to increase capacity.
- Decreasing the maximum allowed speed from 130 km/h to 100 km/h when the intensity downstream gets to high or a traffic jam occurs downstream (A2 Everdingen-Deil, A16 Klaverpolder-Galder).
- Decreasing the maximum allowed speed when it rains (A12 Bodegraven-Woerden).
- Decreasing the maximum speed in case of shockwaves (A12).

In most cases, the 1-minute aggregated speed and intensity is used for Dynamax. In some cases, like the rain-based Dynamax control, more specific data is needed. The rain-based Dynamax control uses rainfall radar data with a 30 second interval from a meteorological service. The shockwave-based Dynamax control needs data that is more actual than the NDW 1-minute aggregated data. It uses individual vehicle passages per 30 seconds from inductive loop detectors.

1.2 Ramp metering

A ramp meter is a traffic light just upstream of an onramp, which controls the traffic flow entering the motorway. The traffic light has a short cycle time. The green light indicates only one (or two) vehicle is allowed to enter the motorway. Ramp metering is used to prevent congestion on the motorway, by limiting and spreading the inflow from onramps. By limiting the inflow, breakdown of the vehicular flow on the motorway can be prevented. Vehicles tend to arrive at onramps in platoons, because of signalized intersections near the onramp. The merging if the platoon can cause disturbances at the motorway. The ramp meter can reduce the disturbances on the motorway by breaking up these platoons.

The ramp meter uses traffic flow conditions on the motorway to determine the maximum allowable inflow from the onramp without causing congestion on the motorway. In the Netherlands, speed and intensity information is obtained from the 1-minute aggregated data of inductive loop detectors, upstream and downstream of the onramp, to determine the upstream demand and the downstream capacity. [32][33]

Appendix B

1 Common data collection systems

1.1 Introduction

In this chapter, several common traffic flow data collection systems will be discussed. The content of this chapter is intended as an overview, not as a complete summary of all existing traffic data collection systems. Besides a description of the traffic detectors, their performance will be discussed. In literature, there are numerous studies that focus on the performance of a particular sensor. Since there are several different types of all sensors, each with its own characteristics, these studies do not provide a full overview for the whole sensor type. Therefore, multiple sources will be mentioned, in order to provide a broader picture of the performance of each sensor type. Finally, also the requirements of the NDW, the Dutch National Data Warehouse for Traffic Information, will be mentioned.

1.2 Inductive loop detectors

The inductive loop detector is the most widely used traffic detection sensor. An electric current flows through a coil of wire that is embedded in the pavement. When a metal vehicle is present or passes over the loop, the inductance of the loop is reduced. The decreased inductance leads to an increase in the frequency of the oscillation. The detector oscillator detects this higher frequency and the vehicle’s presence is registered. Single loop detectors are not able to directly measure speed and vehicle classification. These can be estimated with an algorithm or be determined more accurately with a double loop. [34]

Two studies have been performed two assess the quality of inductive loop detectors in the Netherlands in 2009. In the first study [25] a mean error of 2.0% for the 1-minute aggregated intensity measurements of 50 locations was found. Based on the accuracy, 12% of the measurements did not comply with the NDW requirements, which will be mentioned in section 1.7. Based on the reliability, 8% of the measurements did not meet the NDW requirements.
In the second study [26] traffic measurements have been performed at 75 locations in the Netherlands with radar and infrared detectors, to assess the accuracy and reliability of the traffic measurement with inductive loop detectors. The results showed that 41% of the 1-minute aggregated intensity measurements and 11% of the 1-minute aggregated speed measurements have an error greater than 8%. For the 15-minute aggregated data, the amount of measurements with an error greater than 8% were 17% and 8% respectively.

The Minnesota Department of Transportation [35] compared manual counts of 46 hours of videotapes with loop detector data. They found differences for 1-hour aggregation periods between 0.1% and 3% at motorways and between 2.8% and 8.6% at an intersection. Using probe vehicle data, the speed performance of the detector loops was also determined. An average speed error between 1.2% and 3.3% was found.

Middleton et al. [36] examined the classification accuracy of the PEEK ADR-6000, which uses four inductive loops per lane. The classification accuracy was found to be 98.9%, using 12 vehicle classes. Also the count accuracy performed very well. Only one out of 1923 vehicles was missed.

1.3 Magnetic detectors

Magnetic detectors are, contrary to loop detectors, passive sensors that sense the vertical component of the Earth's magnetic field. When the Earth's magnetic field is disturbed by the metal of a passing vehicle, the vehicle detection is registered. When two magnetic detectors are used, the speed and length of vehicles can also be registered. [34][37]

Grone [38] found a correct detection rate of 92.7% for the 3M Microloop 702. Its mean absolute percent error (MAPE) at a 1-minute aggregation interval was 6.1%. The volume detection was found to be negatively affected by the combined effect of dusk lighting and rain. The Microloop 702 showed a correct vehicle classification rate of 94.9%, using 3 classes based on length (0-24 ft, 25-40 ft, 40+ ft). The 15-minute aggregate classification error was 2.1%.

Minge et al. [37] reported a volume error of 2.5% for the GTT Canoga Microloop. The speed error was found to be smaller than 1 mph. The observed vehicle length error was 4 ft.

The Minnesota Department of Transportation [35] tested the 3M Canoga and found a 15-minute aggregate volume error of 2.4% and a 15-minute aggregate volume error between 1.4% and 4.9%.

The 3M Microloop was also mentioned in a report by Middleton et al. [39]. They observed a volume error of 5%. The speed estimate error was found to be -0.25 mph on average with a standard deviation of 3.6 mph.

Cheung et al. [40] performed two experiments on an urban road to assess the detection capabilities of magnetic sensors. They found a vehicle detection rate better than 99% and the accuracy of the speed and vehicle length estimates appeared to be larger than 90%.

1.4 Video

Video image processor systems use a microprocessor to analyze the video images of one or multiple cameras. The images are interpreted by software and converted to traffic flow data. The video image processor is able to detect vehicles by comparing successive frames. Video-based detectors are able to detect many traffic flow parameters, like volume, speed, density, qualification, acceleration and lane changes. Weather conditions can negatively affect the video image quality. For example, rain, fog, snow and low light conditions reduce the performance of video-based detectors. [34][37]

Grone [38] tested the Autoscope Solo Pro II and found a correct detection rate of 90.5% and a MAPE at a one-minute aggregation interval of 6.5%. The Autoscope Solo Pro II was able to correctly classify 85,4% of the vehicles, using 3 classes based on length (0-24 ft, 25-40 ft, 40+ ft). Its 15-minute aggregate classification error was 10,4%.

Minge et al. [37] found that the volume error for the Miovision is smaller than 2%.

The Minnesota Department of Transportation [35] assessed the performance of the Autoscope Solo and the Traficon VIP D. Both systems were tested in sidefire and overhead position. The 15-minute aggregate volume error of the Autoscope Solo was reported to be 1,5%-2,2% in overhead position and 2,0%-2,7% in sidefire position. The 15-minute aggregate speed errors were 2,5%-7,0% and 5,7%-7,4% respectively. The Traficon VIP D 15-minute aggregate volume error ranged from 1,9% to 3,7% in sidefire position and from 2,7% to 4,8% in overhead position. The 15-minute aggregate speed errors were reported to be 2,3% to 7,7% and 3,3%-7,2% respectively.

In another study, by Middleton et al. [39], the volume error of the Autoscope Solo was estimated at 2.1-3.5%. The reported speed error was 0.8-3.1%.

Robert [41] focussed on the performance of a video image processor in low light conditions. A correct detection rate of 97% was found at night and 94.8% at dusk.

1.5 Microwave radar

Microwave radar uses electromagnetic waves to detect vehicles. It emits the waves in the direction of the road. A portion of these waves is reflected back to the sensor by passing vehicles. Microwave radars can be installed in overhead or sidefire position. There are two types of microwave radar sensors used as traffic detectors.

CW Doppler radar transmits continuous wave (CW) Doppler waveforms and detects vehicles using the Doppler principle. The frequency of the reflected wave will be higher than the frequency of the emitted wave for vehicles approaching the radar and lower for vehicles moving away from the radar. The CW Doppler radar is not able to detect stationary vehicles, since they do not change the frequency of the reflected wave.

Frequency modulated continuous wave (FMCW) radar transmits an electromagnetic wave with a frequency which is continuously being changed with time. The use of different frequencies allows the radar to determine the distance to the vehicle. By using successive measurements, the vehicle speed can also be obtained. The FMWC is also able to detect stationary vehicles. [34][37][32]

Grone [38] assessed the performance of two FMCW radars in sidefire position, the ISS RTMS G4 and the Wavetronix Smartsensor 105. The former was found to have a correct detection rate of 92.8% and a MAPE at a one-minute aggregation interval of 5.5%. The correct detection rate of the

latter was just 83,7% and its MAPE at a one-minute aggregation interval 8,2%. The RTMS G4 showed a correct detection rate of 96,2%, using 3 classes based on length (0-24 ft, 25-40 ft, 40+ ft), and a 15-minute aggregate classification error of 1,6%. For the Smartsensor 105, these percentages were 95,2% and 2,1% respectively.
Minge et al. [37] found a volume error for the Wavetronix Smartsensor HD, another FMCW radar, in sidefire position smaller than 2% at LOS A-D, but ranging up to 20% for LOS E-F. They reported a speed error smaller than 1 mph for LOS A-E and a length error between 1,6 and 2 ft.
Urazghildiiev et al. [42] found a classification accuracy of 99% for a down-looking FMCW radar, using 5 different vehicle classes.
Zwahlen et al. [43] found that the Wavetronix SmartSensor 105 in sidefire position counted more than 95% of all vehicles and measured vehicle speeds to within 3 mph. Vehicle classification was found to be very unreliable, even though just 3 classes were used (0-20 ft, 20-40 ft and more than 40 ft).

1.6 Infrared

There are two types of infrared traffic detectors, passive infrared and active infrared (laser).
Active infrared sensors operate by transmitting low-energy laser beams to the road surface and measuring the time for the reflected signal to return to the sensor. Presence of a vehicle reduces this time and makes detection possible. By transmitting two beams, vehicle speeds can be measured by comparing the times at which the vehicle enters the detection area of each beam.
Depending on the particular sensor and how it is deployed, detection capabilities of active infrared sensors include count, presence, length and speed. [35][37]
Passive infrared sensors detect energy from graybody emission. Graybody emission occurs at all objects with a temperature higher than 0 K. When a vehicle enters the sensor’s field of view, the observed graybody emission changes, caused by the difference in thermal radiation between the road surface and the vehicle, and the vehicle is detected. Passive infrared sensors are able to detect vehicle volume, speed , classification by length and presence. Adverse weather conditions are known to have a negative effect on the detector performance. [34][35][37]

In a study of Minge et al. [37] two active infrared detectors were tested in sidefire position. The volume error in LOS A-D for the PEEK Axlelight and the TIRTL were 5.4% and 3.8% respectively. The speed error of the PEEK Axlelight was reported to be 2 mph and 1.2 mph for the TIRTL. The Axlelight’s length error was 1-2 ft, while the TIRTL length error was smaller than 1 ft.
In a study by the Minnesota Department of Transportation [35] the volume and speed performance of an active infrared, the Autosense II, and a passive infrared, the ASIM IR 254, sensor were reported, both in overhead position. The 15-minute aggregate volume error of the active infrared sensor of 0.7% was found to be much smaller than the passive infrared sensor’s volume error of 10.0%. The difference between both 15-minute aggregate speed errors were smaller, at 5.8% for the Autosense II and 10.8% for the ASIM IR 254.

1.7 NDW requirements

In the Netherlands, the National Data Warehouse for Traffic Information (NDW) operates a database containing traffic data. The data in this database is collected by various road authorities. Every minute, data from more than 20,000 measuring sites in the Netherlands is collected by the database and distributed to users of the data. The data is processed and made available to the customers within 75 seconds. [44]

The NDW has set specific requirement for the data to be accepted in the database. Measuring sites must be operational an average of 97% of the time. These requirements have been summarized in Table 14. In the table, ON stands for the mean absolute percentage error (MAPE) of the 1-minute aggregated data. OB is the percentage of the 1-minute aggregated values, greater than the specified maximum error.

Table 14: NDW accuracy and reliability requirements [45]

	Accuracy	Reliability	
	ON	**OB**	**maximum error**
Intensity	<5%	<2%	20%
Speed	<5%	<2%	20%
Estimated travel time	<20%	<3%	30%
Realized travel time	<10%	<3%	20%

The NDW has also set requirements for the vehicle classification. Within every 24-hour period, over 95% of the vehicles have to be correctly classified.

Appendix C

1 Use cases

Adopted from [7]:

Government use cases

- Traffic information
    - Average travel time (rush hour)
    - Traffic jam
        - Start point
        - End point
        - Avg. speed
    - Throughput
        - Vehicle
        - People
    - Traffic density
    - Incidents

- Accidents
    - Where did it occur?
    - When did it occur?
    - People involved
    - Seatbelt used or not

- Road conditions
    - Roughness
    - Noise
    - Quality of the road
    - Fuel consumption
    - Foggy condition
    - Slippery road
    - Manholes
    - Potholes
    - Black holes
    - Bumps
    - Black spot

Fleet owner use cases

- Fuel consumption
    - Per driver
    - Per location
    - Per time of day
    - Per vehicle
    - Per road segment

- Routing optimization
    - Traffic density (real time)
    - Static road conditions (pot holes)
    - Dynamic road conditions

- Information for service and maintenance planning.
    - Deviation from the optimal vehicle performance

- Driver behavior monitoring
    - Vehicle wear
    - Position (GPS)
    - Breaks (frequency, duration)
    - Drunk or tired drivers
    - Truck heavier/lighter than expected

General public use cases

- Tactical level
    - Route planning
        - Estimated time of arrival
    - Economical route
        - Fuel efficiency
        - Car wear
    - Blocked roads
    - Weather
    - High risk routes
        - Risky corners
        - Other…
    - Team up routes (several people using one car)

- Strategic level
    - Dangerous situations coming up

    - Enough fuel
    - Expected traffic jam
        - Because of accident or alike
    - Traffic jam ahead
        - Alternative suggestions
    - Directed marketing
        - Seat belts
        - Speeding
        - Drunk
        - Sleepy

- Control level
    - Fuel
    - Emissions
    - Maintenance
    - Dynamic tolling
    - E-call
        - Accident severity
        - Cause
            - Heart attack
            - Drunk
            - Sleep
        - Doctor nearby (civilian)

Appendix D

1 Traffic management use cases

In this appendix, several CAN bus data use cases are described. The descriptions of the use cases contain:

- A measurable quantity
- The purpose of obtaining this information
- Definition of the measurable quantity
    - A formula that defines the measurable quantity
- CAN bus measurement method(s)
    - A description of one or several methods that use the CAN bus data to measure or estimate the required variable that was defined for the use case.
    - The CAN bus variables that are needed for these methods with an estimate of the required frequency.
- Norm
    - A quality requirement which describes the allowed difference between the estimate of the measurable quantity from the CAN bus data and the ground truth.
- Test method
    - A description of a method to test whether the estimate of the measurable quantity from CAN bus data meets the norm.

As much as possible, existing definitions of the quantities and norms as used by Rijkswaterstaat and the NDW (National Data Warehouse for Traffic Information) will be used.

1.1 Speed

There are numerous traffic management application which use speed measurements. Some examples are Dynamax, traffic jam detection, ramp metering and peak lanes. Since most of these applications use the 1-minute aggregated values of the average speed, as was shown in Appendix A, also in this use case the 1-minute aggregated average speed values will be considered.

1.1.1 Definition

$V_{average}(T, \tau, x) = \Sigma_{i=1\ldots M}: V_i(x)/M$

where: T is the time, in whole minutes

τ is the length of the time interval $<T, T+\tau>$

$i = 1\ldots M$ are the vehicles (M in total) passing location x during the interval

$V_i(x)$ is the speed of vehicle i which passes location x during the interval

1.1.2 Measurement method

$$V_{average;estimated}(T, \tau, x) = \Sigma_{i=1\ldots m}: v_i(\breve{x})/m$$

where: m is the amount of equipped vehicles passing location x within τ

The speed of vehicle i is obtained from the CAN bus data when it passes location x. $v_i(\breve{x})$ is the speed at the location nearest to location x. The amount of equipped vehicles passing location x within τ, m, is smaller than or equal to M, since not every car is equipped with uCAN.

Required CAN signals

- GPS position
- GPS speed or CAN speed

1.1.3 Norm

For this use case the requirements of the NDW for the 1-minute aggregated average speed will be used (see Table 15).

Table 15: NDW accuracy and reliability requirements for 1-minute aggregated speed measurements [45]

	Accuracy	Reliability	
	ON	**OB**	**maximum error**
Speed	<5%	<2%	20%

1.1.4 Test method

Obtaining the ground truth of $V_{average}(T, \tau, x)$ from loop detector data.

Obtaining the $V_{average;estimated}(T, \tau, x)$ from the passages of the vehicles in the CAN bus data set.

Determining the minimum penetration rate to meet the norm.

1.2 Intensity (for Dynamax, ramp metering, peak lanes)

There are numerous traffic management application which use intensity measurements. Some examples are Dynamax, ramp metering and peak lanes. Since most of these applications use the 1-minute aggregated values of the average speed, as was shown in Appendix A, also in this use case the 1-minute aggregated average intensity values will be considered.

1.2.1 Definition

$I_{average}(T, \tau, x) = M/\tau$

where: T is the time, in whole minutes

τ is the length of the time interval $\langle T, T+\tau \rangle$

M is the amount of vehicles passing location x during the interval

1.2.2 Measurement method

The intensity of the traffic flow can only be determined directly if all vehicles can be detected. This will be achieved when all vehicles are equipped with uCAN. To determine the intensity when only a part of the vehicles is equipped, other methods will have to be used to estimate the intensity.

1.2.3 Measurement method: Vehicle position with full penetration

$I_{average;measured}(T, \tau, x) = M/\tau$

where: T is the time, in whole minutes

τ is the length of the time interval $\langle T, T+\tau \rangle$

M is the amount of vehicles passing location x during the interval

Required CAN signals

- GPS position

1.2.4 Measurement method: Distance headway

The first method which can be used to estimate the intensity uses the vehicle time headways. When the vehicle speed and the gross distance headway (or the net distance headway and the vehicle length) are known, the intensity can be derived from these values. When using this method, it essential to know the amount of lanes.

$I_{average;estimated}(T, \tau, x) = V_{average}(T, \tau, x) / S_v * N$

where: S_v is de average gross distance headway of uCAN equipped vehicles which pass location x during interval τ

N is the amount of lanes

Required CAN signals

- GPS position
- Gross distance headway or
- Net distance headway and CAN speed or GPS speed

1.2.5 Measurement method: Estimation with CAN bus sensors which are related with intensity

Some CAN bus signals are expected to be correlated with intensity. Using regression analysis, these CAN bus values can be used to give an estimation of the intensity. The CAN bus signals which are expected to be correlated with intensity are:

- Speed: based on the fundamental diagram it can be derived that speed and intensity are correlated.
- Speed variation: It is probable that the speed variation will show are relation with intensity. It is likely that the speed variation will be low in low intensity conditions and high in high density conditions.
- Braking behavior: It is likely that the brake usage is low in low intensity conditions and higher in high intensity conditions.
- Steering behavior and direction indicator usage: It is also plausible that the intensity influences the amount of lane changes. The amount of lane changes is likely to be low in low intensity conditions and high when the intensity in higher. It is expected that this will thus also affect the steering behavior and usage of direction indicators.
- Usage of cruise control: It is expected that cruise control is used more in low intensity conditions than in high intensity conditions, since high intensity conditions restrict the possibilities to retain the same speed for a long time.

Required CAN signals

- GPS position
- GPS speed or CAN speed
- Brake
- Brakeforce
- LDI_Active
- RDI_Active
- AnyDirectionIndicator
- SteeringLeft
- SteeringPosition
- SteeringRight
- CruiseActive

- CruiseControl

1.2.6 Norm

For this use case the requirements of the NDW for the 1-minute aggregated average speed will be used (see Table 16).

Table 16: NDW accuracy and reliability requirements for 1-minute aggregated intensity measurements [45]

	Accuracy	Reliability	
	ON	**OB**	**maximum error**
Intensity	<5%	<2%	20%

1.2.7 Test method

Obtaining the ground truth of $I_{average}(T, \tau, x)$ from loop detector data.
Obtaining the $I_{average;estimated}(T, \tau, x)$ from the passages of the vehicles in the CAN bus data set.
Determining the minimum penetration rate to meet the norm.

1.3 Travel times (for travel information)

1.3.1 Definition

$$TT(x_1, x_n, T) = \Sigma_{i1=1\ldots M1}: (t_2(x_2) - t_1(x_1)) / M_1 + \Sigma_{i2=1\ldots M2}: (t_2(x_3) - t_1(x_2)) / M_2 + \ldots + \Sigma_{in-1=1\ldots Mn-1}: (t_2(x_n) - t_1(x_{n-1})) / M_{n-1}$$

with: $T > t_2 > T + 1$ minute

where: $x_1 \ldots x_n$ are the start and end points of the travel time sections

$t(x_a)$ is the time at which vehicle i passes location x_a

M_a is the Total amount of vehicles reaching the end of travel time section a between T and T + 1 minute

The travel time for a section between x_1 and x_n, existing of subsections x_1 to x_2, x_2 to x_3, ... , x_{n-1} to x_n, for time T is the sum of the average travel time of every subsection, given by the average travel time of the vehicles which covered the subsection from x_{a-1} to x_a where x_a was passed between T and T + 1 minute.

1.3.2 Measurement method

$$TT_{measured}(x_1, x_n, T) = \Sigma_{i=1\ldots M1}: (t_2(x_2) - t_1(x_1)) / M_1 + \Sigma_{i=1\ldots M2}: (t_2(x_3) - t_1(x_2)) / M_2 + \ldots + \Sigma_{i=1\ldots Mn-1}: (t_2(x_n) - t_1(x_{n-1})) / M_{n-1}$$

with: $T > t_2 > T + 1$ minute

where: $x_1 \ldots x_n$ are the start and end points of the travel time sections

$t(x_a)$ is the time at which vehicle i passes location x_a

M_a is the amount of equipped vehicles reaching the end of travel time section a between T and T + 1 minute

When there are no vehicle passages between T and T + 1 minute, the data from the most recent minute is taken with vehicle passages. When the last vehicle passage is longer than a certain amount of time, historical averages will be used.

For the purpose of comparing the travel times using floating car data with the travel times acquired with loop detectors it is recommended to use the same subsections. Travel times on a subsection between two location are only reliable if the subsection is the only possible route between both locations.

For every equipped vehicle passing location x_a (distance between x_a to their GPS position smaller than a certain value) between T and T + 1 minute, the time $t_1(x_{a-1})$ at which this vehicle passed location x_{a-1} is taken and subtracted from $t_2(x_a)$ and the travel time on the subsection for this vehicle is obtained. The average travel time of all vehicles on the subsection is taken as the travel time of the subsection. Adding all subsection travel times produces the travel time for the total section.

Required CAN signals

- GPS position

1.3.3 Norm

For this use case the requirements of the NDW for the 1-minute aggregated travel time will be used (see Table 17).

Table 17: NDW accuracy and reliability requirements for 1-minute aggregated realized travel time measurements [45]

	Accuracy	**Reliability**	
	ON	**OB**	**maximum error**
Realized travel time	<10%	<3%	20%

1.3.4 Test method

Obtaining the ground truth of $TT(x_1, x_n, T)$ from loop detector data.
Obtaining the $TT_{measured}(x_1, x_n, T)$ from the passages of the vehicles in the CAN bus data set.
Determining the minimum penetration rate to meet the norm.

1.4 Rain (for Dynamax)

1.4.1 Definition

WR(T, τ, x) : R(T, τ, x)

where: R is the nominative rain intensity (dry, light rain, rain, heavy rain)

T is the time

τ is the length of the time interval <T, T+ τ>

x is the location

1.4.2 Measurement method

$WR_{measured}$(T, τ, x) : RWS(T, τ, x)

"dry"	if $\Sigma_{i=1...M}$:		RWS(T, τ, x)/M	≤ 0.5
"light rain"	if $\Sigma_{i=1...M}$:	0.5 <	RWS(T, τ, x)/M	≤ 1.5
"rain"	if $\Sigma_{i=1...M}$:	1.5 <	RWS(T, τ, x)/M	≤ 2.5
"heavy rain"	if $\Sigma_{i=1...M}$:	2.5 <	RWS(T, τ, x)/M	≤ 3.0

where: RWS is the windscreenwiper speed

M is the amount of equipped vehicles

with: RWS = 0 if Wiperon = 0

RWS = 1 if Wiperinterval = 1

RWS = 2 if Wiperslow = 1

RWS = 3 if Wiperfast = 1

This measurement method might be improved by adding the speed signal. AT higher speeds, more rain will be catched by the windscreen, resulting in a higher windscreenwiper speed.

Required CAN signals

- GPS position
- Wiperon
- Wiperinterval
- Wiperslow
- Wiperfast
- GPS speed or CAN speed

1.4.3 Norm

T.B.D.

1.4.4 Test method

Dependent on the available reference data.

1.5 Wind (for VMS warning)

1.5.1 Definition

$WW_i(T, \tau, x) : W_i(T, \tau, x)$

where: W_i is the average windspeed [m/s]

T is the time

τ is the length of the time interval $<T, T+ \tau>$

x is the location

1.5.2 Measurement method

$$wwi(T, \tau, x) : \sqrt{(wwi_x^2 + wwi_y^2)}$$

where: wwi_x is the average wind speed in x direction

wwi_y is the average wind speed in y direction

The x direction is the driving direction of the vehicle.

In x direction

$$wwi_x = ABG(v) / GBG(v) * SF(v)$$

where: ABG(v) is the current fuel consumption at the current speed v

GBG(v) is the average fuel consumption at speed v. The values of this function differ per vehicle and will have to be known.

$SF_x(v)$ is a scaling factor to convert the difference between the current fuel consumption and the average fuel consumption in a windspeed. The values of $SF_x(v)$ differ per vehicle and will have to be known.

In y direction

Hypothesis: Because crosswinds will cause a counter directed steering manouvre, the wind speed in y direction can be deduced from the steering position of the vehicle.

$$wwi_y = ASH(x_1,x_2)/HSH(x_1,x_2)*SF()$$

where: $ASH(x_1,x_2)$ is the current average steering position between x_1 and x_2

$HSH(x_1,x_2)$ is the historical average steering position between x_1 en x_2. The values of this function differ per vehicle and will have to be known.

$SF_y(v)$ is a scaling factor to convert the difference between the current and historical steering position in wind speed. The values of $SF_y(v)$ differ per vehicle and will have to be known.

Required CAN signals

- GPS position
- GPS speed or CAN speed
- SteeringPosition
- FuelConsumption or Throttle

1.5.3 Norm

T.B.D.

1.5.4 Test method

Obtaining the ground truth of $WW_i(T, \tau, x)$ from available road side wind measurement stations.

Obtaining wwi(T, τ, x) from the passages of the vehicles in the CAN bus data set.

Determining the minimum penetration rate to meet the norm.

1.6 Fog (for VMS warning)

1.6.1 Definition

$WM(T, \tau, x)$:	"very dense fog"	if		$M(T, \tau, x) \leq 50$
	"dense fog"	if	$50 <$	$M(T, \tau, x) \leq 200$
	"fog"	if	$200 <$	$M(T, \tau, x) \leq 1000$
	"no fog"	if	$1000 <$	$M(T, \tau, x)$

where: M is visibility [m]

T is the time

τ is the length of the time interval $<T, T+ \tau>$

x is the location

1.6.2 Measurement method

$$WM_{measured}(T, \tau, x): MLG(T, \tau, x)$$

with boundaries (example):

"no fog"	if		$\Sigma_{i=1...M}$: $MLG(T, \tau, x)/M$	≤ 0.01
"fog"	if	$0.01 <$	$\Sigma_{i=1...M}$: $MLG(T, \tau, x)/M$	≤ 0.4
	or	$0.4 <$	$\Sigma_{i=1...M}$: $MLG(T, \tau, x)/M$	≤ 0.6
		&	$V_{average;estimated}(T, \tau, x)$	> 80
"dense fog"	if	$0.4 <$	$\Sigma_{i=1...M}$: $MLG(T, \tau, x)/M$	≤ 0.6
		&	$V_{average;estimated}(T, \tau, x)$	≤ 80
	or	$0.6 <$	$\Sigma_{i=1...M}$: $MLG(T, \tau, x)/M$	≤ 0.8
	or	$0.8 <$	$\Sigma_{i=1...M}$: $MLG(T, \tau, x)/M$	≤ 1.0
		&	$V_{average;estimated}(T, \tau, x)$	> 60
"very dense fog"	if	$0.8 <$	$\Sigma_{i=1...M}$: $MLG(T, \tau, x)/M$	≤ 1.0
		&	$V_{average;estimated}(T, \tau, x)$	≤ 60

where: MLG is rear fog light usage

with: $MLG = 0$ if $RearFogLight = 0$

$MLG = 1$ if $RearFogLight = 1$

The initial boundary values can be optimized with the CAN bus data set and reference data.

Required CAN signals

- GPS position
- FrontFogLight
- AnyFogLight
- RearFogLight
- GPS speed / CAN speed

1.6.3 Norm

T.B.D.

1.6.4 Test method

Obtaining the ground truth of WM(T, τ, x) from available road side weather measurement stations or from other nearby meteorological stations.

Obtaining $WM_{measured}$(T, τ, x) from the passages of the vehicles in the CAN bus data set.

Determining the minimum penetration rate to meet the norm.

1.7 Obstacle detection

1.7.1 Definition

$BD(M,x,y,\tau)$: $I_{average}(T, \tau, x,y)) = 0$ & $I_{average}(T, \tau, x+s,y)) > 0$ & $I_{average}(T, \tau, x-s,y))$

where: T is the time, in whole minutes

τ is the length of the time interval <T, T+ τ>

x is the considered position where no flow is possible if there is an obstacle

s is the distance upstream and downstream of the obstacle where flow is possible

y is the lateral position (the lane)

1.7.2 Measurement method: Speed difference

Hypothesis:

At the location of the obstacle, the speed of the vehicles will be lower than upstream and downstream of the obstacle.

$BD_{estimated}(M,x,y,\tau)$: $V_{average;estimated}(T, \tau, x) * 1.2 < V_{average;estimated}(T, \tau, x+s)$ &
$V_{average;estimated}(T, \tau, x) * 1.2 < V_{average;estimated}(T, \tau, x-s)$

Required CAN signals

- GPS position
- GPS speed / CAN speed

1.7.3 Measurement method: Steering behavior

If an obstacle is present, it is expected that drivers will steer a lot to avoid the obstacle. The amount of lane changes will be high and the direction indicators will be used more. This will mostly be the case upstream of the obstacle. Obstacles on the right lane will be easier to detect using this driving behaviour than obstacles on the left lanes, because there will not always be traffic on the left lanes.

Hypothesis: In case of an obstacle, the changes in steering position will be larger at the position of the obstacle than further upstream and downstream of the obstacle.

$BD_{estimated}(M,x,y,\tau)$: $GSV(x-2s, x-s, T, \tau) * 2 < GSV(x-s, x, T, \tau)$ &
$GSV(x, x+s, T, \tau) * 1.5 < GSV(x-s, x, T, \tau)$

with: average change in steering position $GSV = \Sigma_{i=1...M}: SV(x_1,x_2)/M$

where: SV is the sum of the change in steering position between location x_1 and x_2

$i = 1 \ldots M$ are the vehicles (M in total) passing location x during the interval

$GSV(x_1, x_2, T, \tau)$ is the average sum of the change in steering position of the vehicles between x_1 and x_2. These vehicles pass x_2 in interval $<T, T+\tau>$.

Required CAN signals

- GPS position
- SteeringPosition

1.7.4 Measurement method: Direction indicators

Hypothesis: In case of an obstacle, the usage of the direction indicators will be larger at the position of the obstacle than further upstream and downstream of the obstacle.

$BD_{estimated}(M,x,y,\tau)$: $GRG(x-2s, x-s, T, \tau) * 2 < GRG(x-s, x, T, \tau)$ &

$GRG(x, x+s, T, \tau) * 1.5 < GRG(x-s, x, T, \tau)$

with: average indicator usage $GRG = \Sigma_{i=1 \ldots M}: RG(x_1,x_2)/M$

where: RG is:

- Time of direction indicator usage (float)
- Amount of times of direction indicator usage (int)
- The usage of the direction indicators (bool)

between location x_1 en x_2.

$i = 1 \ldots M$ are the vehicles (M in total) passing location x during the interval

Required CAN signals

- GPS position
- RDI_Active
- LDI_Active
- AnyDirectionIndicator

1.7.5 Measurement method: Combinations of CAN bus sensor data

Analogous to measuring the intensity, detecting an obstacle is difficult when only a part of the vehicles provide data. Therefore, the combined use of several of the previously mentioned method might offer a solution. Each method can give an indication of the chance that an obstacle is present. By combining these indicators, a more reliable obstacle detection can be provided. Other signals which might show a relation with obstacle presence include cruise control and brake related signals.

Required CAN signals

- GPS position
- GPS speed / CAN speed
- Brake
- BrakeForce
- HazardLights
- LDI_Active
- RDI_Active
- AnyDirectionIndicator
- CruiseActive
- CruiseControl

1.7.6 Norm

T.B.D.

1.7.7 Test method

To test the obstacle detection, the available CAN bus data from the equipped vehicles can be used. However, the occurrence of obstacles is rare and not always registered. Therefore it is unlikely that the obstacle detection can be tested sufficiently using only the available CAN bus data.
There are however other possibilities to be sure enough data will be available to test the obstacle detection. Intentionally creating an obstacle on the public road will cause danger and is therefore not an option. Creating an obstacle in a controlled environment is however possible. Another option would be to study the effects of an obstacle on the driving behavior with a driving simulator.

Appendix E

1 CAN bus data

In this chapter the CAN bus data from the Smart In-Car project will be discussed. In the first paragraph, the initial dataset is described and converted to Matlab. In paragraph 1.2, several visualizations are presented, to provide insight into the contents of the dataset. In paragraph 1.3, de quality of the data will be discussed.

1.1 Raw CAN bus data from Smart In-Car

The Smart In-Car project can deliver CAN bus data of around 130 taxis and 40 vehicles of the ANWB (the Dutch emergency roadside assistance) between April 2012 and October 2012. The format and amount of the data made it impossible to process and use all this data for this research. Before deciding which use case to investigate, a relatively small dataset was processed and analyzed. With the knowledge obtained from this dataset, it could be better assessed which use case could be investigated with the available data.
This initial dataset was around 8.5 Gigabyte, containing 60 million lines of CAN bus events from around 60 taxis. On average, 4 CAN bus signals were stored per second. So this dataset contained around 4000 hours of data (15 million seconds). In this chapter, first it will be explained how the raw data was processed to a MATLAB format.


```
{ "timestamp" : 1335874566255000 , "gpsdata" : { "lat" : 51.46685527777779 , "lon" : 5.473162500000002} ,
    "uvm" : { "id" : 119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874566511100 , "candata" : { "Speed" : 22.579200744628906} , "uvm" : { "id" :
    119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874566656500 , "candata" : { "RPM" : 1797.0} , "uvm" : { "id" : 119611000024168 ,
    "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874567031300 , "candata" : { "Speed" : 22.924800872802734} , "uvm" : { "id" :
    119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874567335000 , "gpsdata" : { "lat" : 51.466869444444455 , "lon" : 5.473101944444446} ,
    "uvm" : { "id" : 119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874567551200 , "candata" : { "Speed" : 23.443201065063477} , "uvm" : { "id" :
    119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874567666100 , "candata" : { "RPM" : 1822.0} , "uvm" : { "id" : 119611000024168 ,
    "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874568071100 , "candata" : { "Speed" : 23.50080108642578} , "uvm" : { "id" :
    119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874568332000 , "gpsdata" : { "lat" : 51.46687861111112 , "lon" : 5.473010833333335} ,
    "uvm" : { "id" : 119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874568590900 , "candata" : { "Speed" : 23.96160125732422} , "uvm" : { "id" :
    119611000024168 , "licenseplate" : "16-XN-FT"}}
{ "timestamp" : 1335874568667700 , "candata" : { "RPM" : 1848.0} , "uvm" : { "id" : 119611000024168 ,
    "licenseplate" : "16-XN-FT"}}
```


Figure 34: example of the raw CAN bus data

Ideally, the available data would be made available in a matrix format, containing all parameter values in single row. This however, was not the case. In Figure 34, a couple of lines from the raw data files are shown. Instead of containing all parameters, most lines only contain one CAN bus parameter. They are logged individually in so called json format. Since we wanted to be able to perform analyses on the data, the raw data had to be converted to a proper format which allows these analyses. The program that has been chosen for these analyses is Matlab. The raw data had to be imported in Matlab and converted to a matrix format.

1.1.1 Conversion to Matlab

The matrix code that has been written to import the raw data files is shown in Appendix I. Importing the text files to Matlab is time consuming. On a fast computer, the total processing speed, including the step in the next paragraphs, is around 5MB per minute. So the processing of this dataset will take around 30 hours.

Table 18 shows the rows of the matrix that were converted with the raw data of Figure 34. After importing, every CAN bus parameter has a separate column.

Table 18: Raw data imported and converted in Matlab

epoch timestamp	uvm	lon	lat	rpm	speed	brake	steeringposition	rdi_active	ldi_active
1335874566255000	119611000024168	5,4732	51,4669	NaN	NaN	NaN	NaN	NaN	NaN
1335874566511100	119611000024168	NaN	NaN	NaN	22,5792	NaN	NaN	NaN	NaN
1335874566656500	119611000024168	NaN	NaN	1797	NaN	NaN	NaN	NaN	NaN
1335874567031300	119611000024168	NaN	NaN	NaN	22,9248	NaN	NaN	NaN	NaN
1335874567335000	119611000024168	5,4731	51,4669	NaN	NaN	NaN	NaN	NaN	NaN
1335874567551200	119611000024168	NaN	NaN	NaN	23,4432	NaN	NaN	NaN	NaN
1335874567666100	119611000024168	NaN	NaN	1822	NaN	NaN	NaN	NaN	NaN
1335874568071100	119611000024168	NaN	NaN	NaN	23,5008	NaN	NaN	NaN	NaN
1335874568332000	119611000024168	5,4730	51,4669	NaN	NaN	NaN	NaN	NaN	NaN
1335874568590900	119611000024168	NaN	NaN	NaN	23,9616	NaN	NaN	NaN	NaN
1335874568667700	119611000024168	NaN	NaN	1848	NaN	NaN	NaN	NaN	NaN

The data only contains a limited number of CAN bus parameters. In total, 24 different CAN bus parameters were found in the total dataset (see Table 19).

The vehicles in the dataset only provide a couple of these parameters. The available parameters differ per vehicle. Table 20 shows the amount of uCAN equipped taxis per vehicle type that provide a particular parameter. Note that the amount of parameters per vehicle is only a small fraction of the total amount of existing CAN bus parameters, as was shown in Table 1.

Table 19: Parameters in the dataset [source: TASS uCAN API documentation]

Parameter name	Parameter description	Unit	Range	Remark
RPM	Rotational speed of the motor	rpm	0-10000	
Speed	Vehicle speed	km/h	0-1000	
Brake	Brake pedal	boolean	0 or 1	1 means brake is active
SteeringPosition	Steering position			
RDI_Active	Right direction indicator control	boolean	0 or 1	1 means right indicator is currently controlled
LDI_Active	Left direction indicator control	boolean	0 or 1	1 means left indicator is currently controlled
RDI_Switch	Right direction indicator switch	boolean	0 or 1	1 means switch is in the 'right' position
LDI_Switch	Left direction indicator switch	boolean	0 or 1	1 means switch is in the 'left' position
RDI_Pulse	Actuator signal for rdi light	boolean	0 or 1	1 means actuator is currently active
LDI_Pulse	Actuator signal for ldi light	boolean	0 or 1	1 means actuator is currently active
Hazardlights	Hazard lights	boolean	0 or 1	1 means hazard lights are on
RearFogLight	Rear fog light	boolean	0 or 1	1 means rear fog light is on
FrontFogLight	Front fog light	boolean	0 or 1	1 means front fog light is on
WiperMotor	Motor of front windscreen wipers	boolean	0 or 1	1 means wipers are moving
FuelConsumption	Fuel consumption	ml/h	>=0	
FuelLevel	Fuel level	l	>=0	
Odometer	Odometer	km	>=0	
OdometerHR	Odometer	hm	>=0	
Ignition	Ignition	boolean	0 or 1	1 means key position is 'ignition'
LowBeam	Low beam	boolean	0 or 1	1 means low beam is on
Lights	Low beam and / or parking light	boolean	0 or 1	1 means parking light and/or low beam is on
ParkingBrake	Parking brake status	boolean	0 or 1	1 means parking brake is active
ReverseGear	Reverse gear status	boolean	0 or 1	1 means reverse gear is on
OilPressureSwitch	Oil pressure switch	boolean	0 or 1	1 means oil pressure is low or faulty

Table 20: Amount of taxis with parameter data. [source: TASS uCAN API documentation]

Vehicle type	Periodic parameters							Event based parameters												Other	
	FuelConsumption	FuelLevel	OdometerHR	Odometer	RPM	Speed	SteeringPosition	Brake	FrontFogLight	Ignition	LDI_Switch	Lights	LowBeam	OilPressureSwitch	ParkingBrake	RDI_Switch	RearFogLight	ReverseGear	WiperMotor	GPS	Accelerometer
Mercedes E-class (211) 2006-2009					32	32		32		32	32		32		32	32		32		32	2
Mercedes E-class (212) 2009	35		35		35	35	35	35	35	35	35		35			35	35	35	35	35	3
Mercedes Sprinter (906) 2006	11				11	11				11		11						11		11	2
Opel Vivaro 2001-2006					2	2		2		2	2	2				2				2	
Opel Vivaro 2006-2010				35	35	35		35		35	35	35			35	35				35	
Opel Vivaro 2010				13	13	13		13		13	13	13				13				13	
Volkswagen Transporter 2003-2009		4		4	4	4				4		4		4	4			4		4	

The aggregated matrix contains, besides the 24 CAN bus parameter columns, the following columns:

- 7 time related columns:
  - year
  - month
  - day
  - hour
  - minute
  - second
- 1 vehicle identifier column:
  - uvm
- 2 GPS position columns:
  - longitude
  - latitude
- 3 accelerometer parameters:
  - x
  - y
  - z

1.1.2 Conversion of the epoch timestamp

The timestamp of the CAN bus data is given in microsecond epoch format. This is the time in microseconds since 0:00 on January 1st 1970. To make the timestamp understandable, the timestamp is converted to standard date and time. Columns with year, month, day, hour, minute and second are added to the matrix. Table 21 shows part of the same matrix as was shown in Table 18, after adding the date and time columns.

Table 21: Imported data with converted epoch timestamp in Matlab

year	month	day	hour	minute	second	Epoch timestamp	uvm	lon	lat	rpm	speed
2012	5	1	12	16	6,2550	1335874566255000	119611000024168	5,4732	51,4669	NaN	NaN
2012	5	1	12	16	6,5111	1335874566511100	119611000024168	NaN	NaN	NaN	22,5792
2012	5	1	12	16	6,6565	1335874566656500	119611000024168	NaN	NaN	1797	NaN
2012	5	1	12	16	7,0313	1335874567031300	119611000024168	NaN	NaN	NaN	22,9248
2012	5	1	12	16	7,3350	1335874567335000	119611000024168	5,4731	51,4669	NaN	NaN
2012	5	1	12	16	7,5512	1335874567551200	119611000024168	NaN	NaN	NaN	23,4432
2012	5	1	12	16	7,6661	1335874567666100	119611000024168	NaN	NaN	1822	NaN
2012	5	1	12	16	8,0711	1335874568071100	119611000024168	NaN	NaN	NaN	23,5008
2012	5	1	12	16	8,3320	1335874568332000	119611000024168	5,4730	51,4669	NaN	NaN
2012	5	1	12	16	8,5909	1335874568590900	119611000024168	NaN	NaN	NaN	23,9616
2012	5	1	12	16	8,6677	1335874568667700	119611000024168	NaN	NaN	1848	NaN

1.1.3 Filling the event based parameter columns

Event based parameters only occur sporadically in the raw data, since they are only stored when the parameter is switched on or off. So most of the rows don't contain event based parameter data, although the state of the parameters is known at that moment, because it will still be the same as it was in the last row which does contain the event based parameter. To make the event based parameters available in every second, the events based parameter columns are filled with the last known value.

1.1.4 Frequency of the periodic parameters

After the previous steps, the rows still only contain one type of periodic parameter. In the next step the parameters need to be combined. Therefore, the period of the periodic parameters, the GPS signal and accelerometer signal in the dataset are analyzed in this paragraph, to choose a proper aggregation period.

Figure 35 shows the occurrences of the GPS signal with a specific period. Each bar in the figure represents a 0.01s interval. It appears that 73% of the occurrences of the GPS signal have a period between 0.985s and 1.025s.

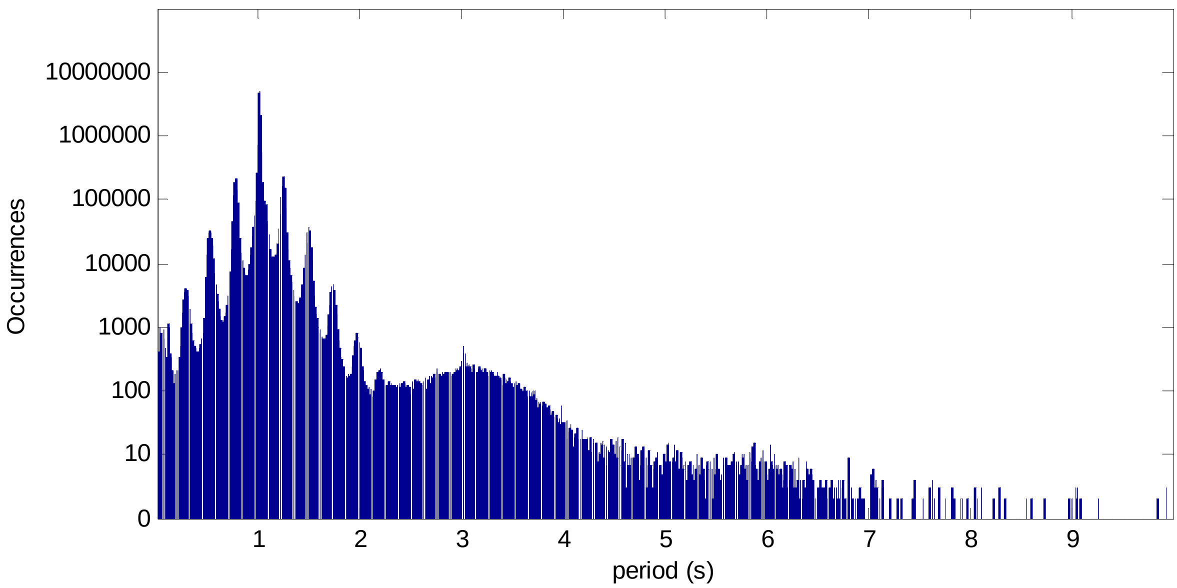

Figure 35: The occurrence of periods of the GPS signal (0.01s intervals)

Over 99% of the speed signals are given with a period between 0.495s and 1.005s, with peaks around 1.00s, 0.60s and 0.50s, representing 95% of the measurements (see Figure 36).

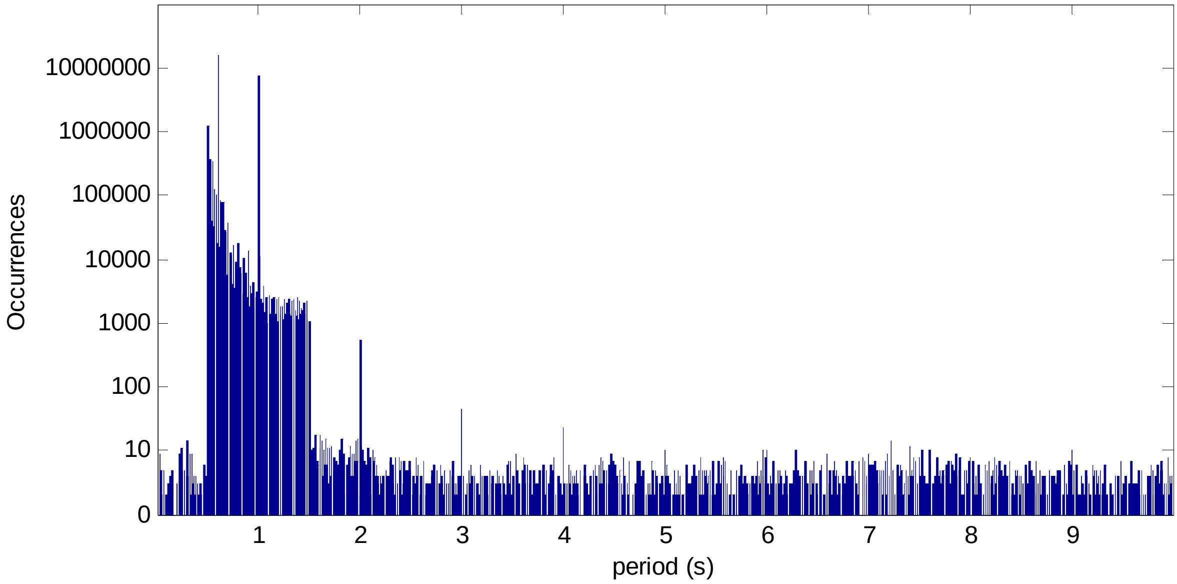

Figure 36: The occurrence of periods of the speed signal (0.01s intervals)

The other periodic parameters appear to have a more fixed period. The period of the RPM is between 1.015 and 1.025 in over 93% of the occurrences in the dataset. Over 98% of the steering position occurrences have a period between 1.005s and 1.015s. The period of the accelerometer was found to have a period between 0.095s and 0.115s in almost all occurrences. In nearly all occurrences of the fuel consumption, the period is between 0.995s and 1.005s or between 1.095s and 1.105s, while the period in almost all occurrences of the fuel level is between 0.995s and 1.005s or between 1.195s and 1.205s. This shows that the aggregation interval of 1 second is a good choice, since most parameters have a period near 1 second.

The Odometer and OdometerHR signal, which were called periodic signals in Table 20, appeared to be event-based signals instead of periodic signals. The Odometer and OdometerHR signals occur once every kilometer or once every hectometer respectively.

1.1.5 Aggregating data

Since most parameters are stored once per second, which was shown in the previous paragraph, the data has been aggregated with a 1 second period. Every row that occurs in the same second has been aggregated to a single row. In case a parameter occurs more than once a second, its average has been used in the aggregated data line. In Table 22 the data lines in the resulting matrix are shown which have been aggregated from the data in Table 21.

Table 22: 1 second aggregated data in Matlab

year	month	day	hour	minute	second	epoch timestamp	uvm	lon	lat	rpm	speed
2012	5	1	12	16	6,4742	1335874566474200	119611000024168	5,4732	51,4669	1797	22,5792
2012	5	1	12	16	7,3959	1335874567395900	119611000024168	5,4731	51,4669	1822	23,1840
2012	5	1	12	16	8,4154	1335874568415420	119611000024168	5,4730	51,4669	1848	23,7312

1.2 Visualisations

With the data in Matlab matrix format, the data is suitable for processing and analyzing. In this paragraph a number of visualizations of the dataset will be shown and discussed, to provide insight in the available data and the possibilities this data gives. Also the accuracy and reliability of the parameters will be discussed.

In the first paragraph, data from a single trip will be visualized. For a small period, parameters like position, speed, steering position and brake and indicator usage will be shown. In the second paragraph, data from all vehicles in the dataset will be used to show relations between several parameters. Finally, several parameters will be plotted on a map in the third paragraph, to show the driving characteristics per location.

1.2.1 Single vehicle visualizations

In this paragraph, CAN bus data from a single vehicle is visualized. A seven minute piece is taken from the data of one of the uCAN equipped taxis. In Figure 37 the GPS position is plotted on a map. The taxi travels from east to west.

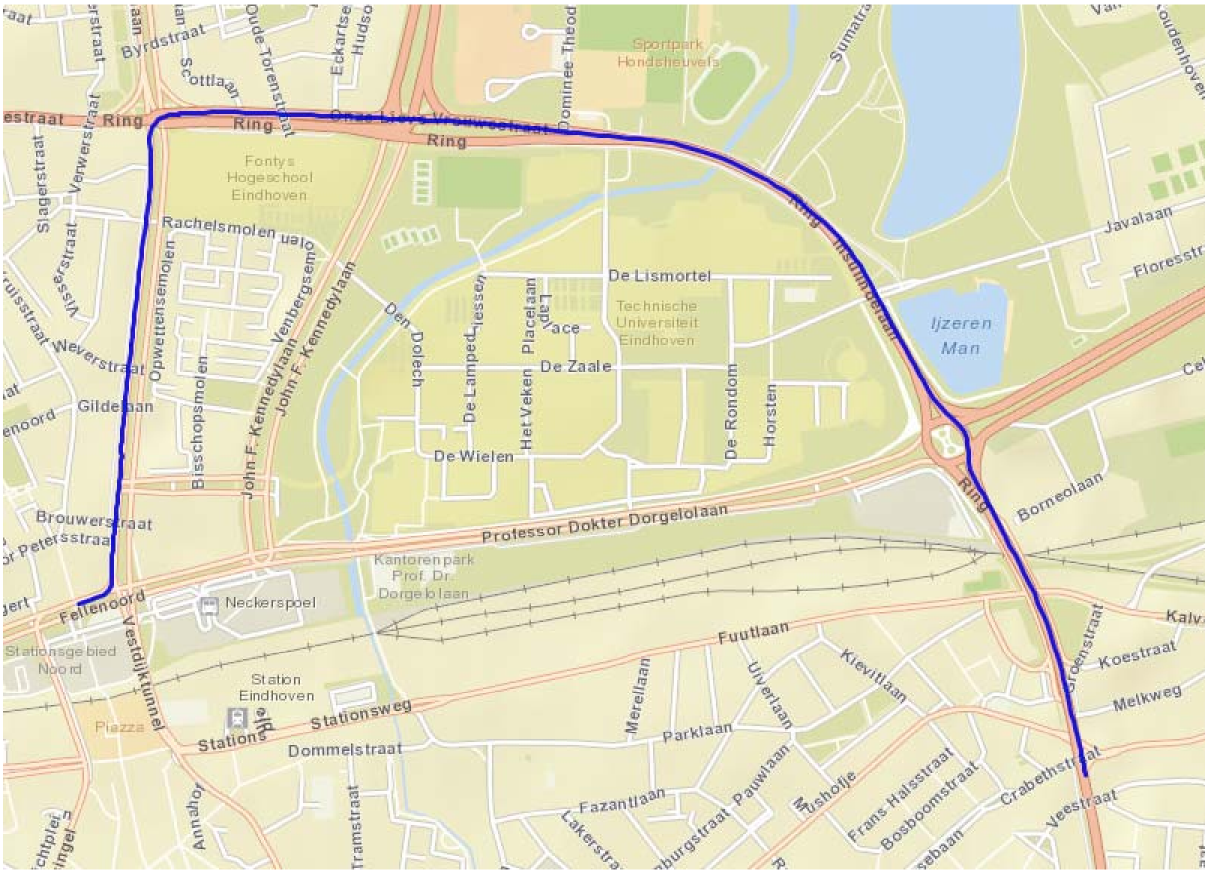

Figure 37: GPS track from a single trip. The blue line indicates the trajectory of the vehicle. The vehicle travels from east to west (right to left).

Besides the GPS position, the speed, RPM of the motor, steering position, fuel consumption, odometer, brake and direction indicator usage are also recorded during this trip. The values of these parameters are presented in Figure 38 - Figure 45.

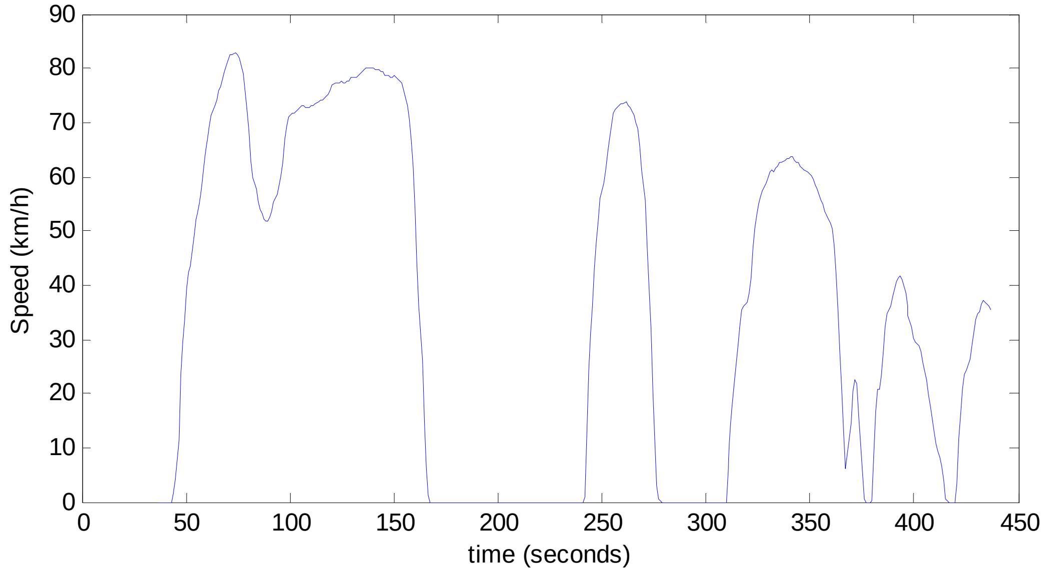

Figure 38: Speed during the trip

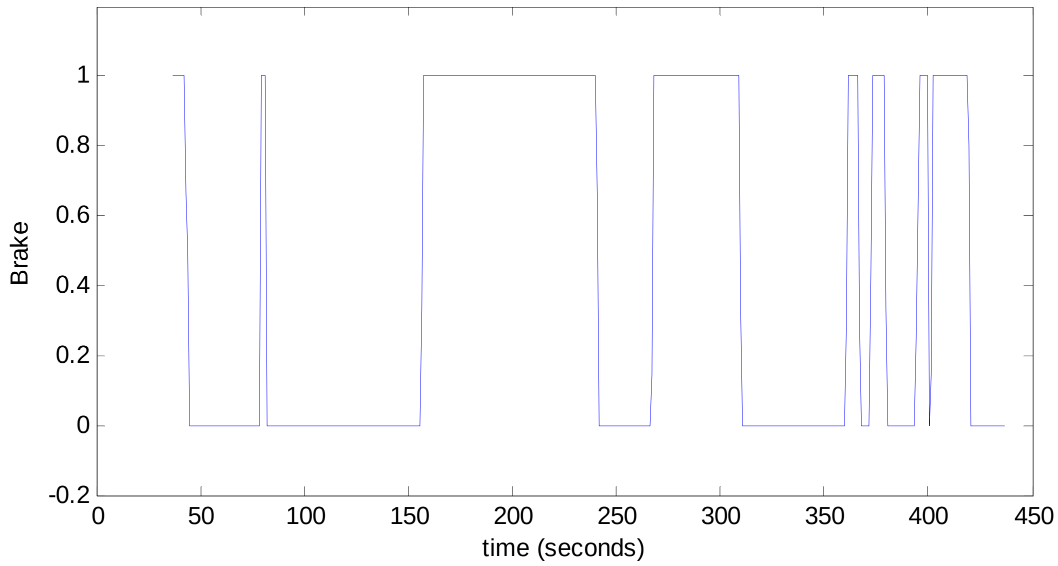

Figure 39: Brake usage during the trip (0=off,1=on)

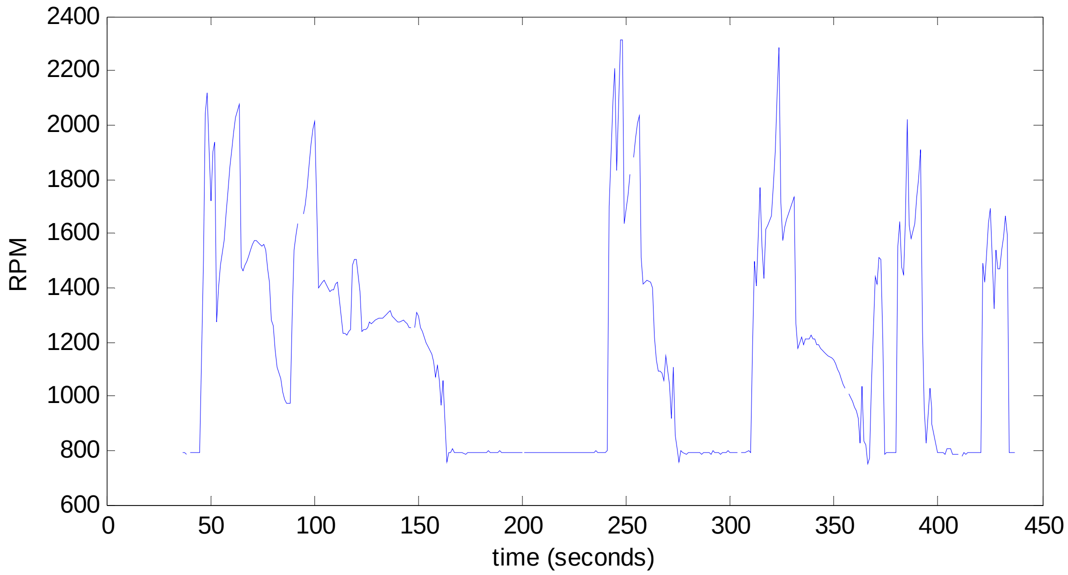

Figure 40: Engine revolutions during the trip

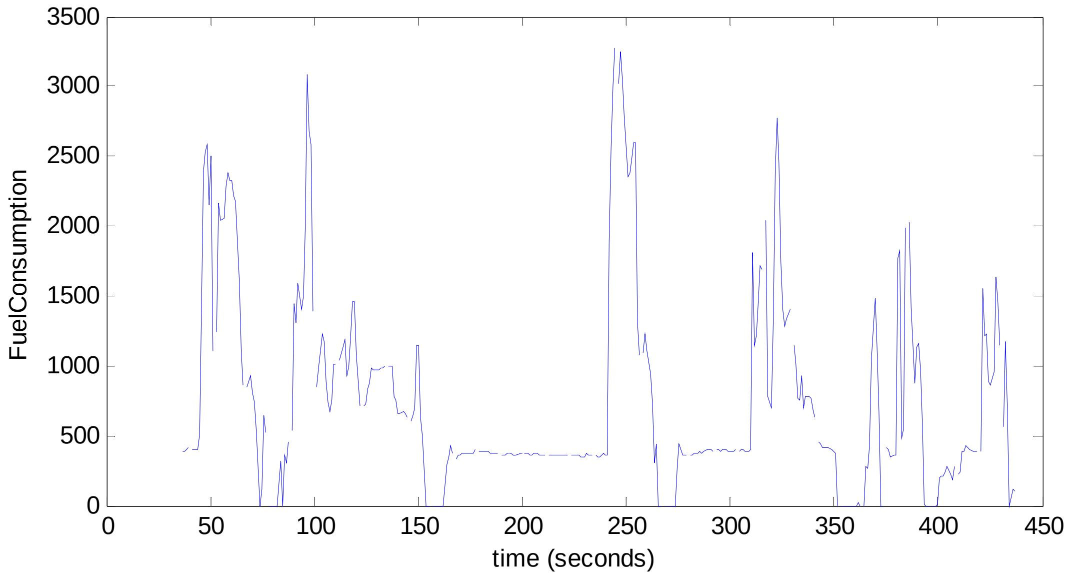

Figure 41: Fuel consumption during the trip. The unit of fuel consumption unknown.

The unit of the fuel consumption was given in Table 19 as ml/h. But the average value of the fuel consumption during this trip is just 758 ml/h. Since the average speed during the trip is 35.6 km/h, this would yield a fuel consumption of 47 km/litre. This is unrealistically high. Therefore, the fuel consumption unit is left out in the figure. The source of the error and the proper fuel consumption unit have not been further investigated.

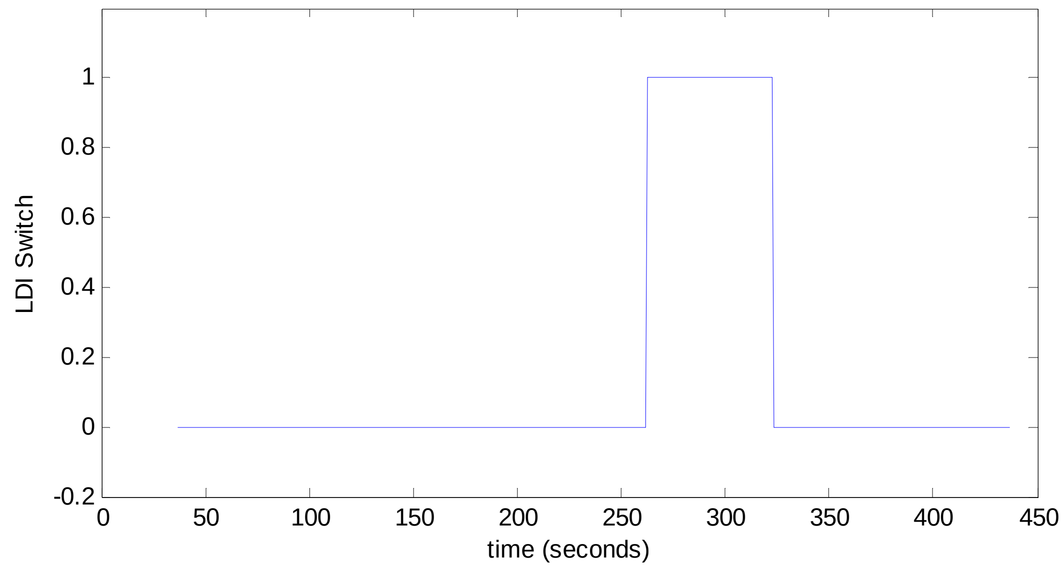

Figure 42: Left direction indicator usage during the trip (0=off, 1=on)

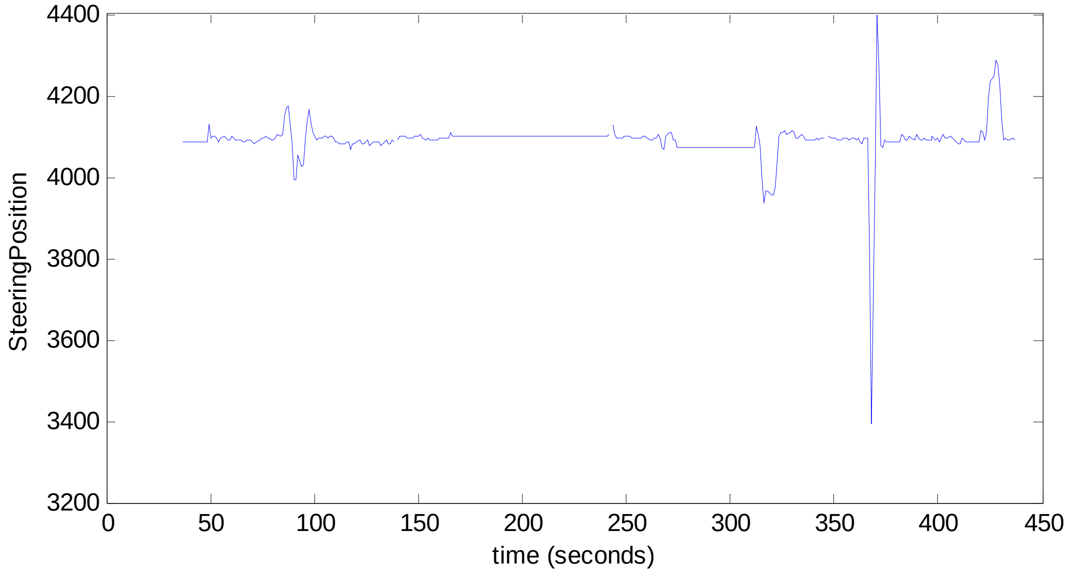

Figure 43: The steering position values during the trip. The unit of the steering position is unknown.

The unit of the steering position was not given in Table 19. When analyzing all steering position data in the dataset, the steering position shows values which range from 3297 to 4895, with an average and median of 4096. 4096 appears to be the center value of the steering position (the value when the vehicle drives in a straight line). When this number gets deducted from the steering position value, negative values represent steering to the left, positive values represent steering to the right. This has been done with the trip data to obtain Figure 44.

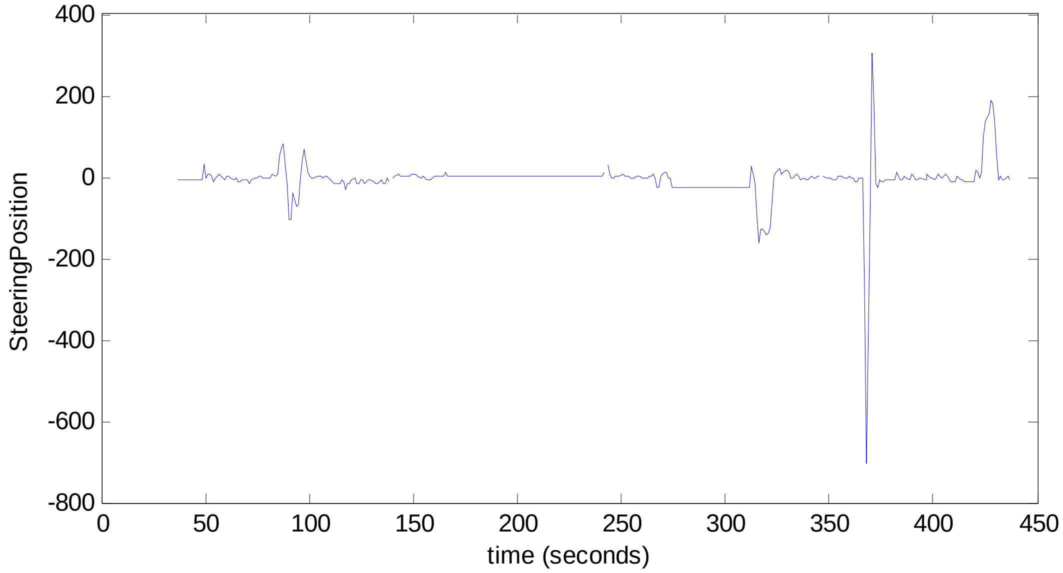

Figure 44: The adjusted steering position values during the trip

Using the GPS and steering position data, for some locations, the steering position values have been compared with the radius of the road curve. In Table 23, for 5 locations with different curves, the average steering position is shown. Using this data, a formula has been composed which can approximate the curve radius with the steering position:

$$R = 68000 / S^{1.5}$$

with: R = curve radius [m]

S = abs(SteeringPosition-4096)

The last column in the table shows the difference between the calculated and actual curve radius.

Table 23: Steering position and curve radius

radius (m)	average steering position	absolute steering position (abs_SP)	$R = 68000 / S^{1.5}$	difference
16	3820	276	15	-7,3%
41	3959	137	42	3,4%
192	4043	53	176	-8,2%
450	4123	27	485	7,7%
900	4113	17	970	7,8%

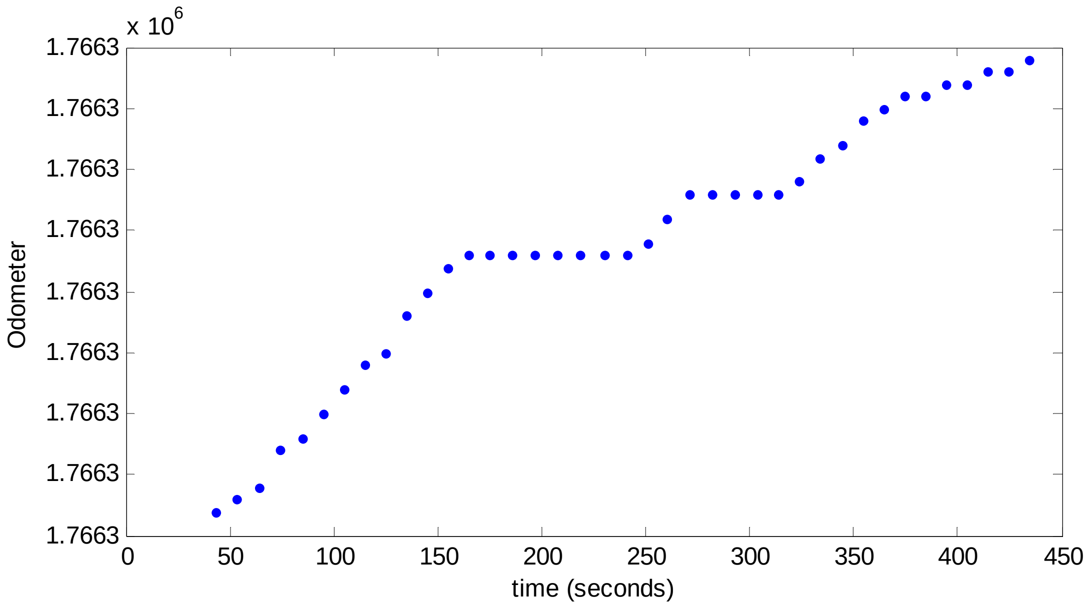

Figure 45: Odometer (hm) values during the trip

When combining the CAN bus parameters with the GPS positions, Figure 46 - Figure 51 can be made, showing the CAN bus parameters for every position during the trip.

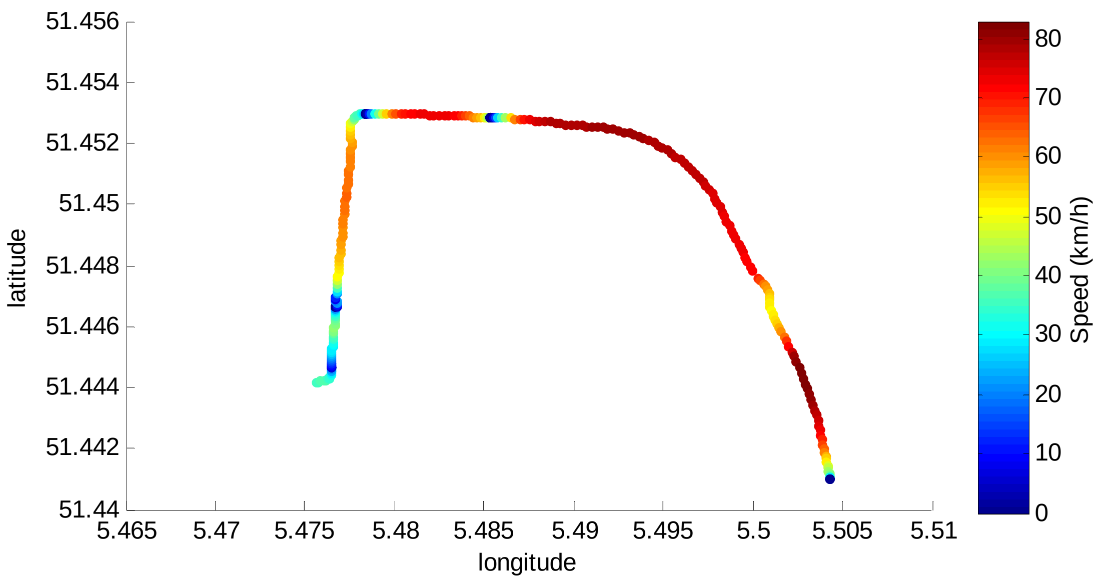

Figure 46: Position and speed during the trip. The color scale indicates the speed of the vehicle.

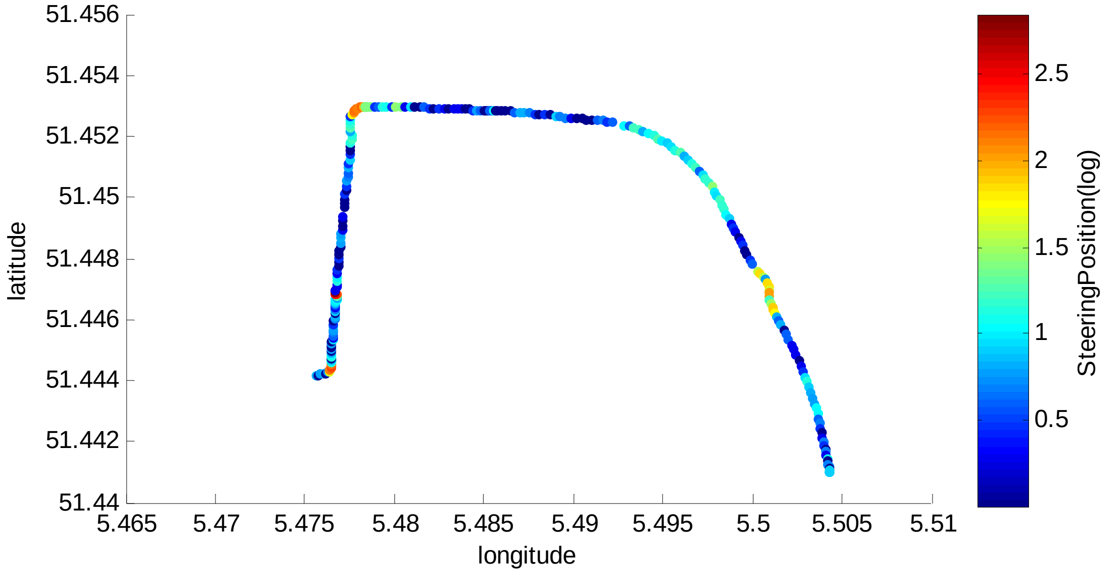

Figure 47: Position and log adjusted steering position during the trip. The (logarithmic) color scale indicates the difference between the steering position and the neutral steering position.

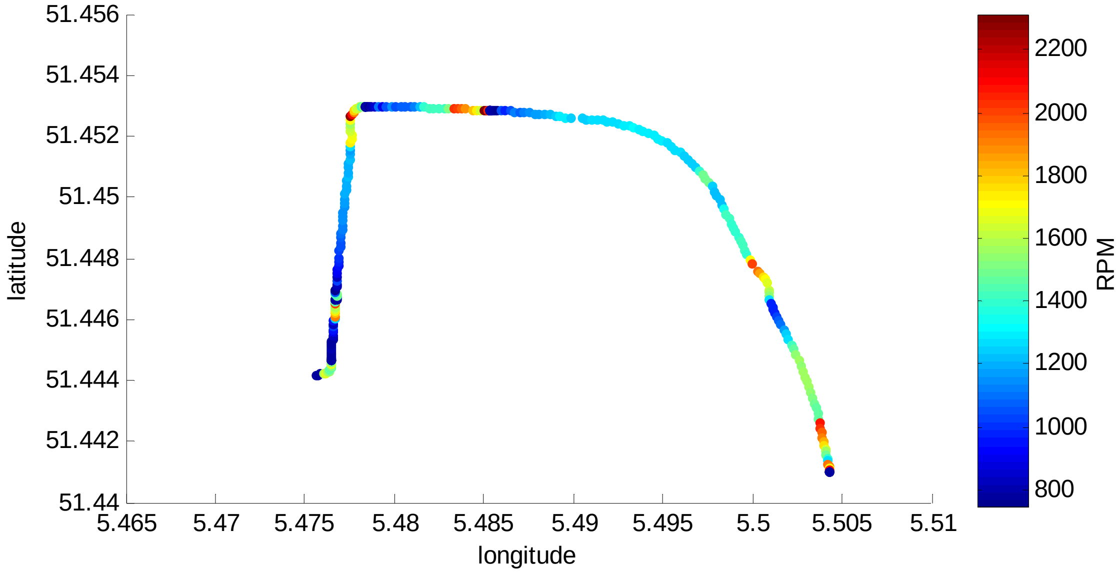

Figure 48: Position and RPM during the trip. The color scale indicates the engine revolution per minute.

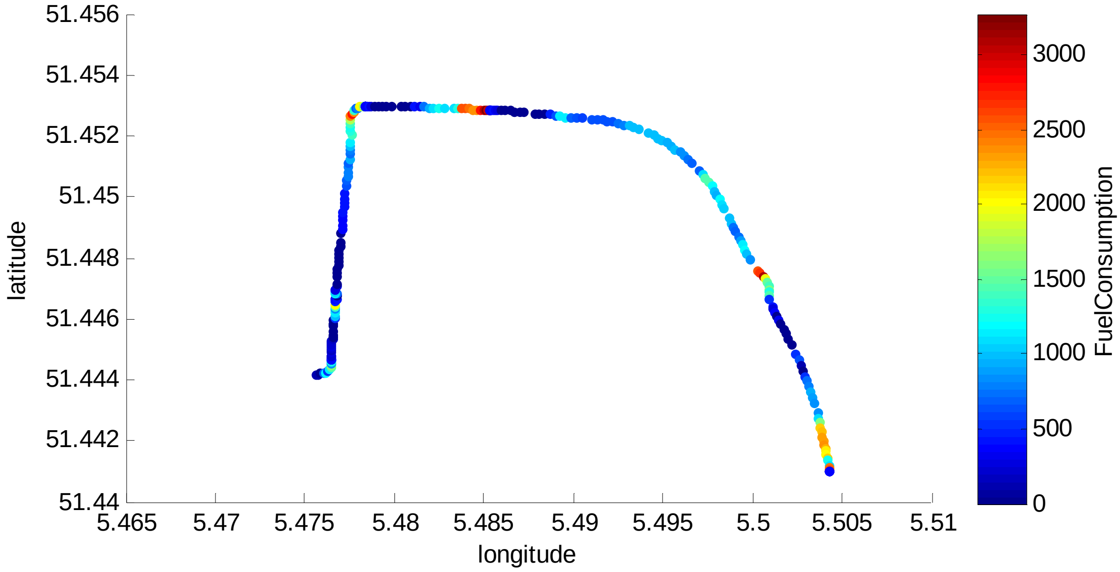

Figure 49: Position and fuel consumption during the trip. The color scale indicates the fuel consumption of the vehicle. The unit of the fuel consumption is unknown.

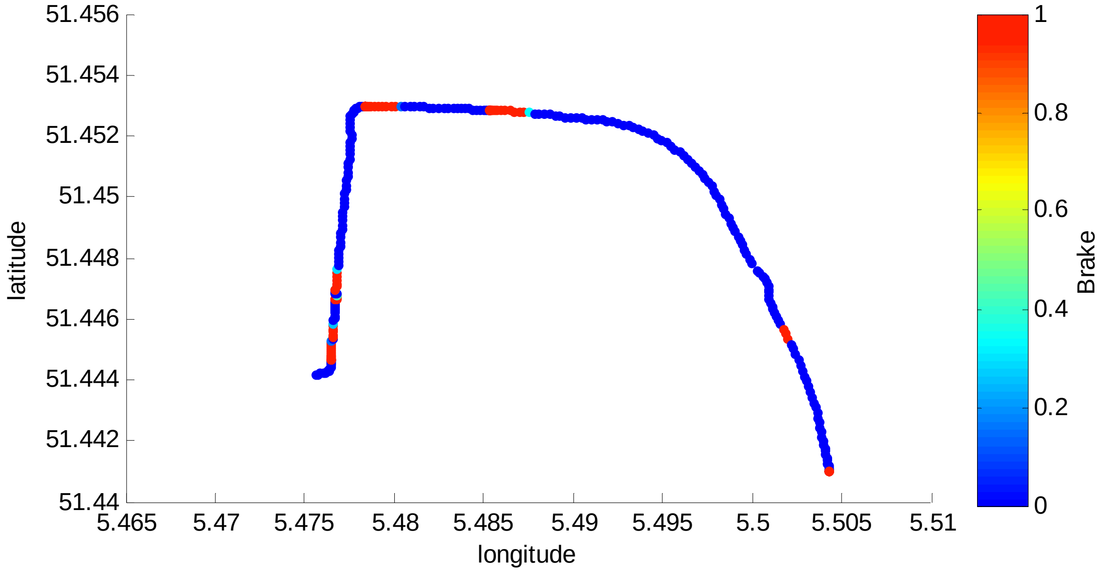

Figure 50: Position and brake usage during the trip. Red indicates the usage of the brakes. Blue indicates the brakes were not used.

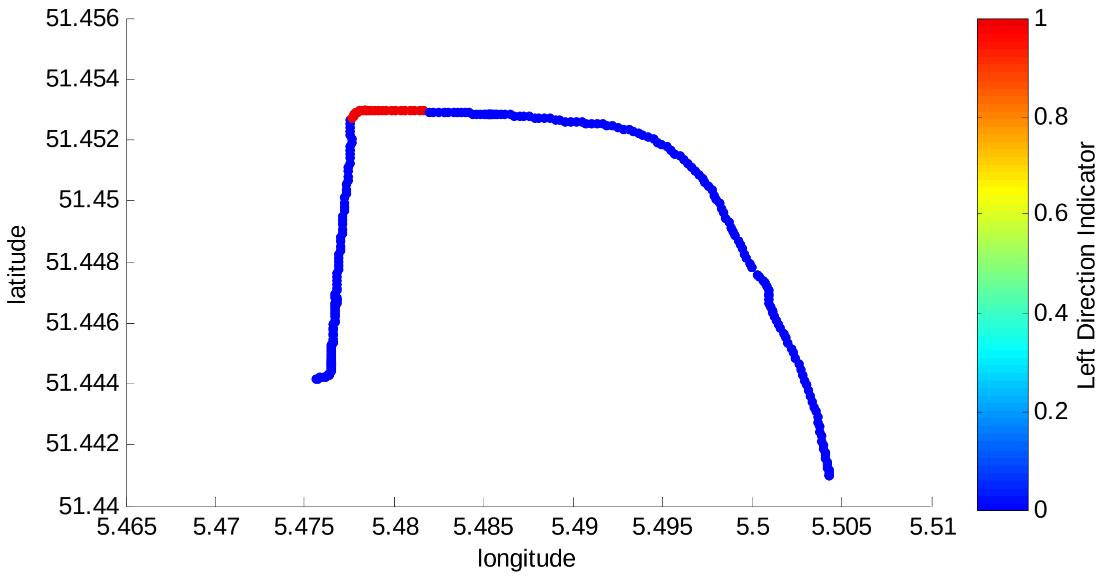

Figure 51: Position and left direction indicator usage during the trip. Red indicates the usage of the direction indicators. Blue indicates that the direction indicators were not used.

1.2.2 CAN bus parameter relationships

In this paragraph, several scatter plots will be presented, showing multiple parameters.

In Figure 52, the steering position of the vehicle is plotted against its speed. It appears, as can be expected, the steering angle decreases as the speed increases.

The relation between the revolutions per minute of the motor of the vehicle and its speed can be observed in Figure 53. Five distinct lines can be seen in the figure. Every line corresponds with a specific gear. So even if the gear is not part of the available CAN bus parameters, it can still be obtained if the speed and RPM are known. Some of the data points are not on one of the five lines which represent the five gears of the vehicles. A possible explanation is that the rpm and speed have changed during the second in which both signals were obtained. When a vehicle is accelerating, the rpm and speed will grow linearly. When the RPM signal is stored at the beginning of a second and the speed at the end of the second, this can cause deviations from the gear ratios.

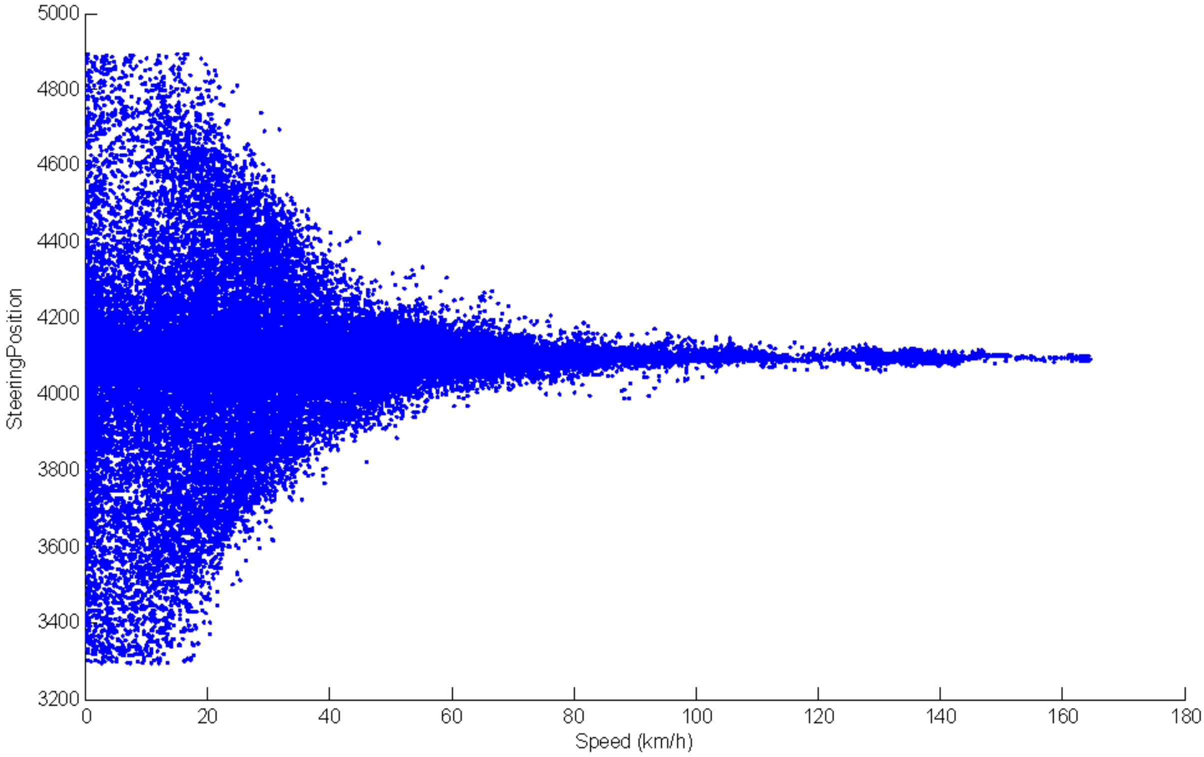

Figure 52: Speed versus steering position. The unit of the steering position is unknown.

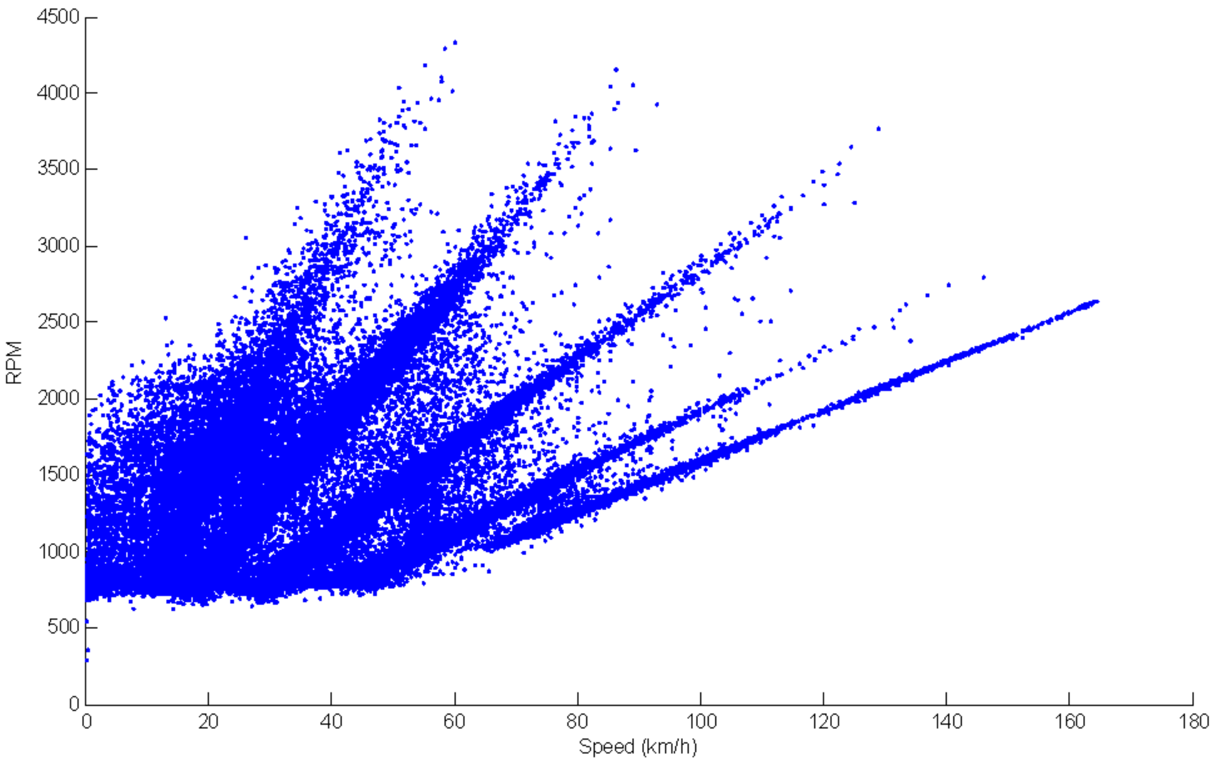

Figure 53: Speed versus RPM

In Figure 54 the average fuel consumption in shown per speed. As mentioned before, the unit of the fuel consumption is unknown. The local maximum at 40 km/h is notable. A possible explanation might be that around 40 km/h, the vehicles are often accelerating and not cruising at the same speed. But this could also be expected for the speeds below 40 km/h.

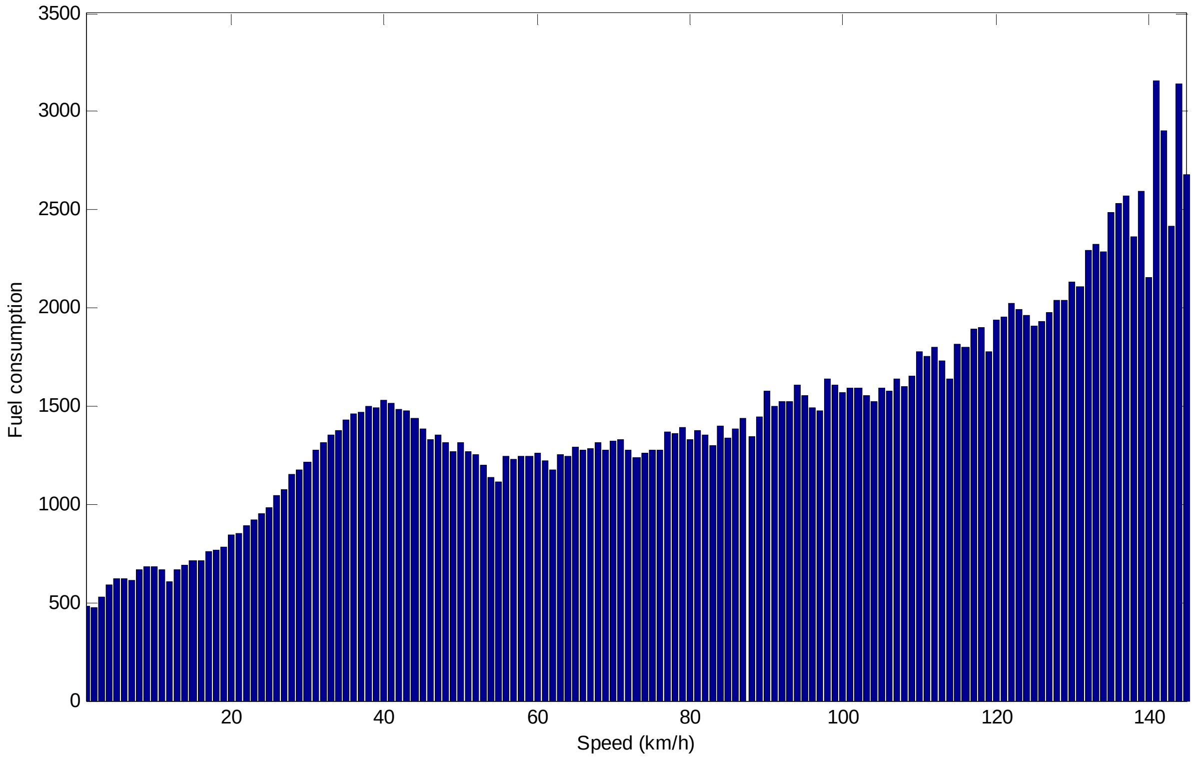

Figure 54: Speed versus average fuel consumption. The unit of the fuel consumption is unknown.

1.2.3 CAN bus parameter location plots

In this paragraph, the recorded GPS position will be used as a basis for several figures. The presented figures will provide insight in the available data and the area where the uCAN equipped vehicles are present. First of all, the presence of the uCAN equipped vehicles in the Eindhoven area is plotted. The colours in Figure 55 show how much time the vehicles have spend in a particular area. Every pixel in all images in this paragraph represents a 7 by 5 meter area. Notice that the colour scale in this figure is logarithmic. The uCAN equipped taxis are mainly present in the centre of Eindhoven and on the main urban roads. Their usage of the motorways around Eindhoven (at the bottom and left side of the figure) is limited.

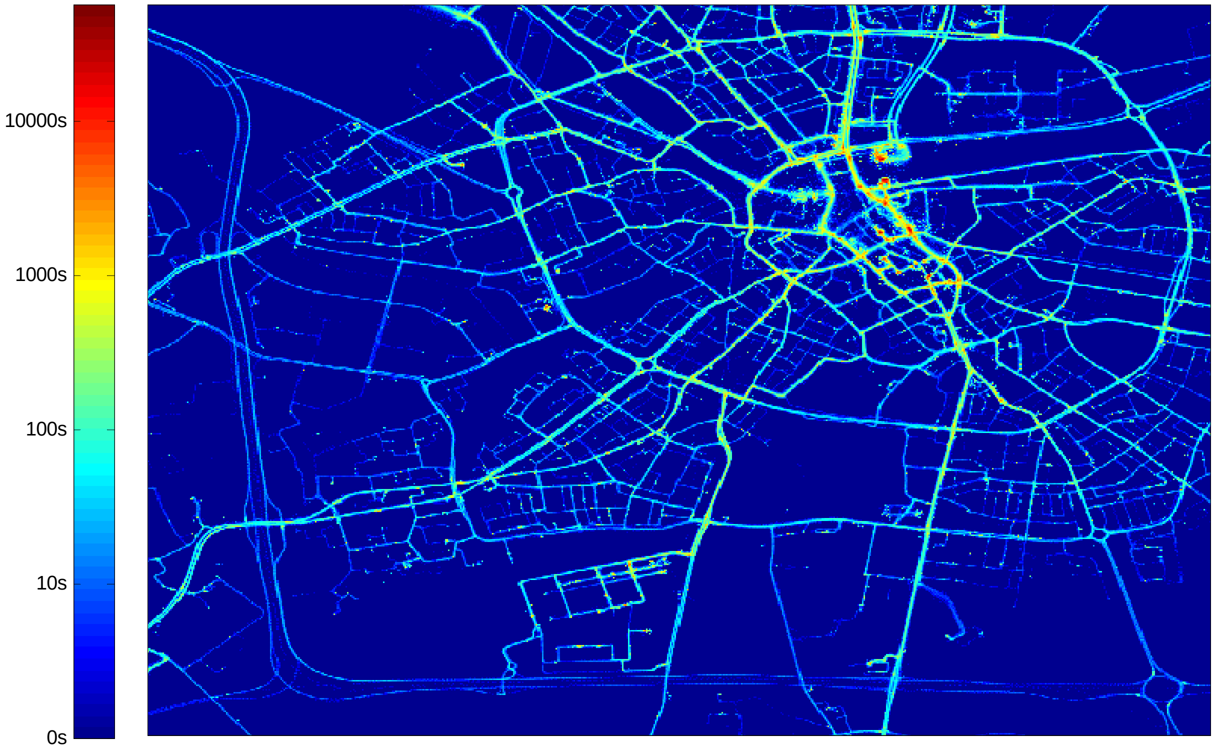

Figure 55: Presence of the uCAN equipped taxis in Eindhoven. The color scale represents how often a uCAN equipped vehicle was present at a specific location. Dark blue (the background color) indicates that no vehicles in the data set have been at that location. Blue indicates that vehicles were rarely present at the location. Red indicates that vehicles were often present at the location.

In Figure 56, instead of the presence, the colours represent the average speed at every position. As one would expect, the average speeds are high at motorways, lower on urban roads and even more limited in residential areas.

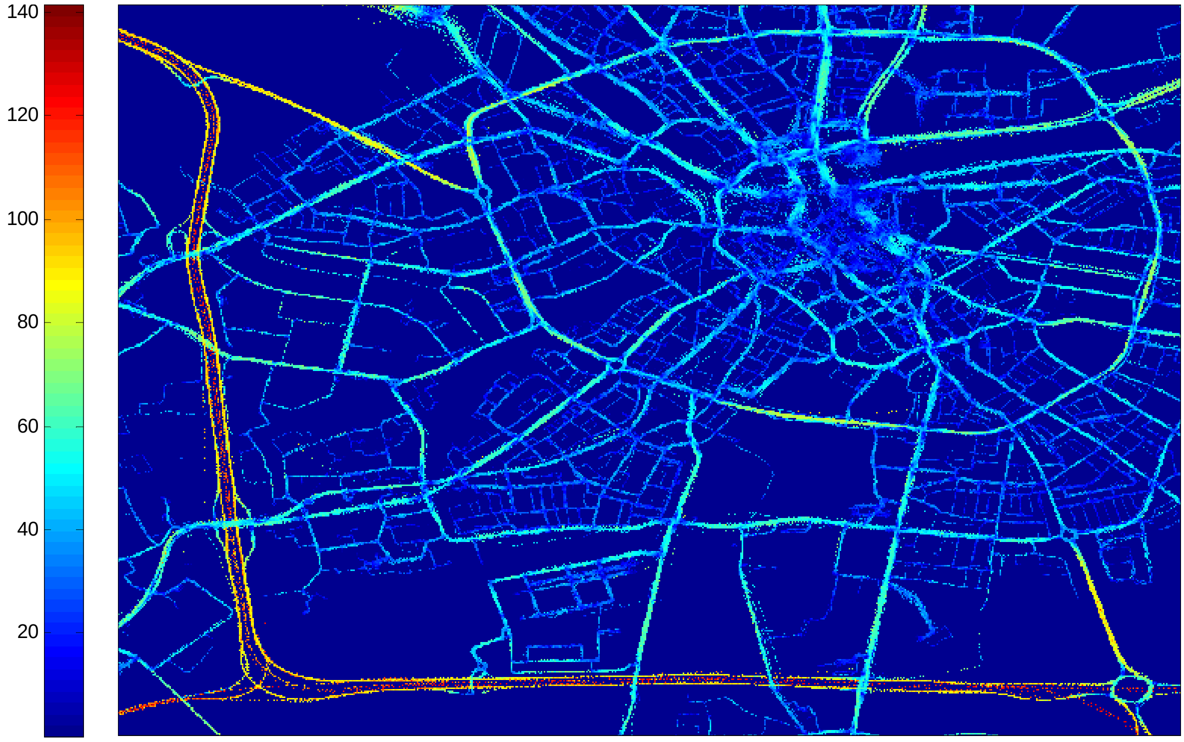

Figure 56: Average speed (km/h) in Eindhoven. The color scale represents the average speed of the uCAN equipped vehicles at a specific location. Dark blue (the background color) indicates that no vehicles in the data set have been at that location. Blue indicates that the average speed of the vehicles at the location is low. Red indicates that the average speed of the vehicles at the location is high.

Similarly, Figure 57 and Figure 58 show the usage of brakes and direction indicators respectively. The red colours in the figures represent locations where almost all drivers are using their brakes or direction indicators. The blue colours represent locations where brakes and direction indicators are sporadically or never used. As was the case in the previous figures, also these figures seem to be reliable. Direction indicators and brakes are used at the locations where this is expected, like at off ramps, before curves and at intersections.

Finally, Figure 59 shows the average steering position for every location. When closely observing the figure, we can see the steering position matches perfectly with the curves at the motorways, roundabouts, on- and off-ramps and straight road sections.

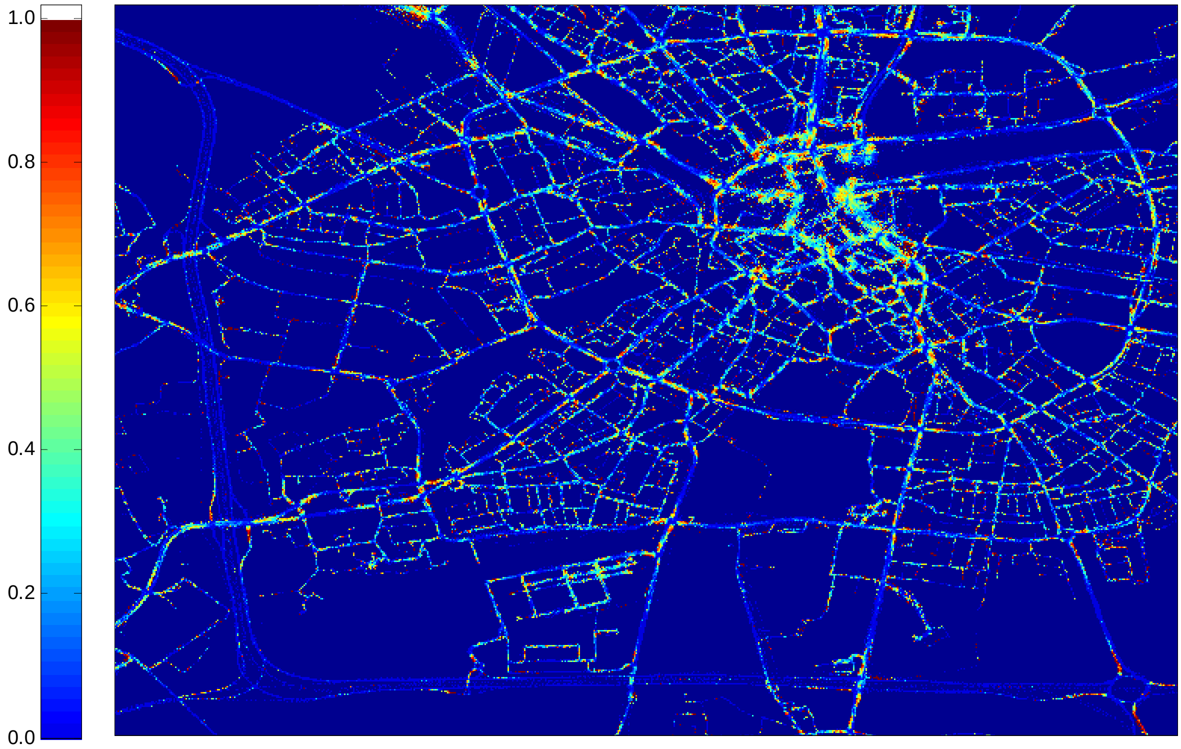

Figure 57: Brake usage in Eindhoven. The color scale represents which part of the uCAN equipped vehicles uses its brakes while being at a specific location. Dark blue (the background color) indicates that no vehicles in the data set have been at that location. Blue indicates that no vehicles use their brakes at the location. Red indicates that all vehicles use their brakes at the location.

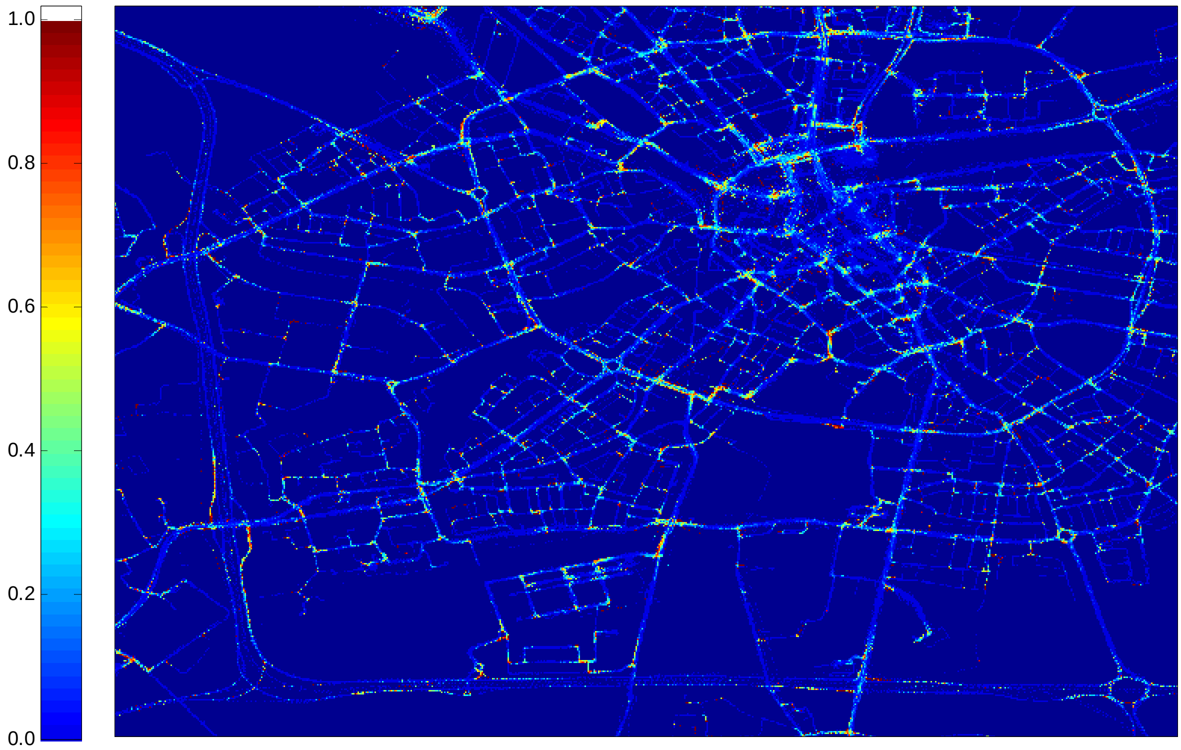

Figure 58: Direction indicator usage in Eindhoven. The color scale represents which part of the uCAN equipped vehicles uses its direction indicators while being at a specific location. Dark blue (the background color) indicates that no vehicles

in the data set have been at that location. Blue indicates that no vehicles use their direction indicators at the location. Red indicates that all vehicles use their direction indicators at the location.

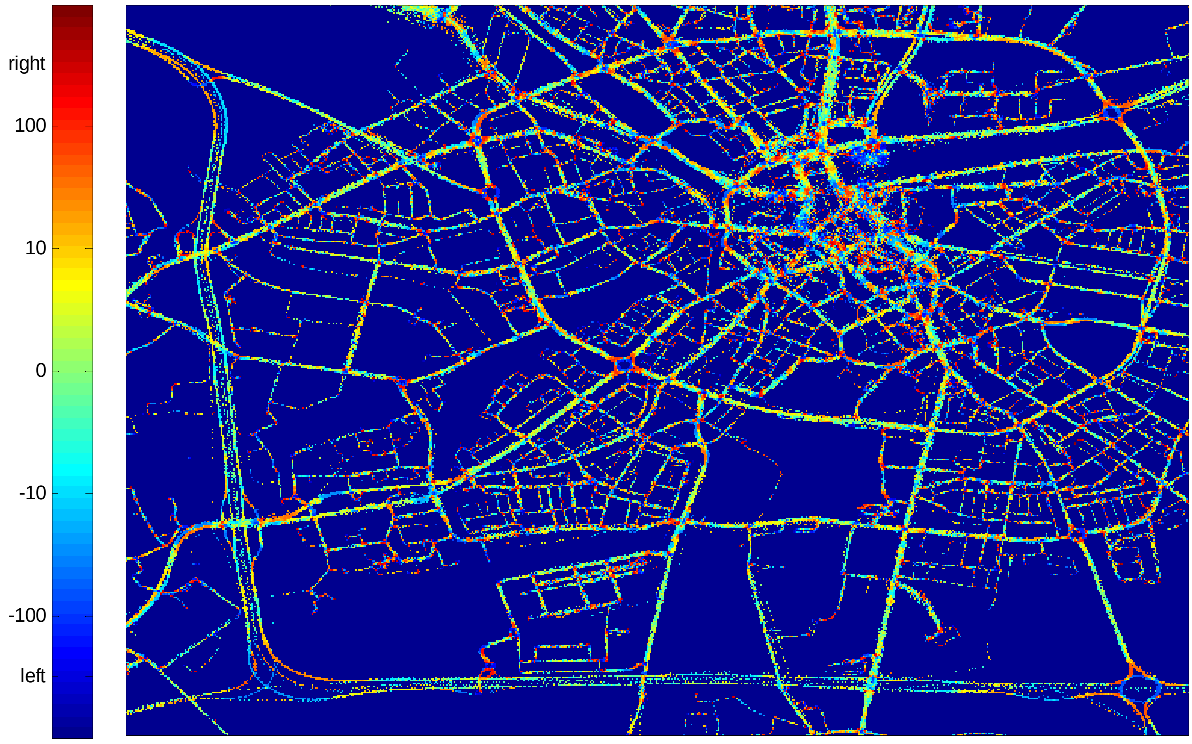

Figure 59: Steering position in Eindhoven. The color scale represents the average steering position of the uCAN equipped vehicles while being at a specific location. Dark blue (the background color) indicates that no vehicles in the data set have been at that location. Blue indicates that the vehicles on average are making a left turn at the location. Red indicates that the vehicles one average are making a left turn at the location.

1.3 Data quality

The previous paragraphs have shown us that the reliability and accuracy of the CAN bus parameters in the dataset seem to be very good on a large scale. In this paragraph, the reliability and accuracy of the most important CAN bus parameters will be discussed in closer detail. The parameters that will be investigated are GPS position and speed.

1.3.1 CAN bus speed

The distribution of the CAN bus speed is visualized in Figure 60. For the lion's share, 42% of the time, the vehicles are standing still. The figure also shows some outliers above 200 km/h. Especially the large amount of data points where the recorded speed is 420 km/h is noteworthy.

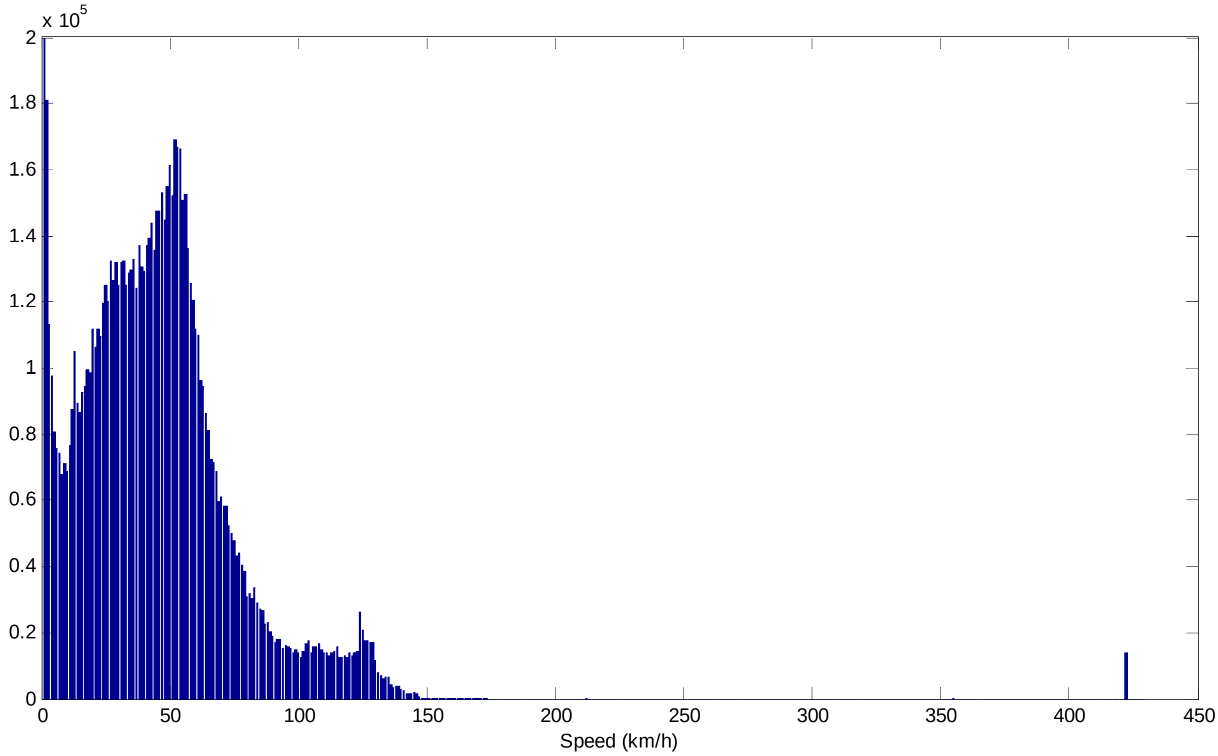

Figure 60: Amount of CAN bus speed occurrences in the dataset

These values seem to occur when the engine is started or turned off. At that moment, besides the anomalies in speed, also other CAN bus parameters show unrealistic values. Because values above 200 km/h are quite unrealistic, they are clearly faulty. But there could also be errors in the lower recorded speeds, which are not noticeable in the figure. To check for these errors, the difference between the speeds in subsequent seconds has been calculated. Because of the physical limitations of the cars (acceleration and deceleration) these values shouldn't be much higher than 6 m/s2 (22 km/h/s) or lower than -10m/s2 (-36km/h/s). The speed differences have been calculated for the whole dataset and have been visualized in Figure 61.

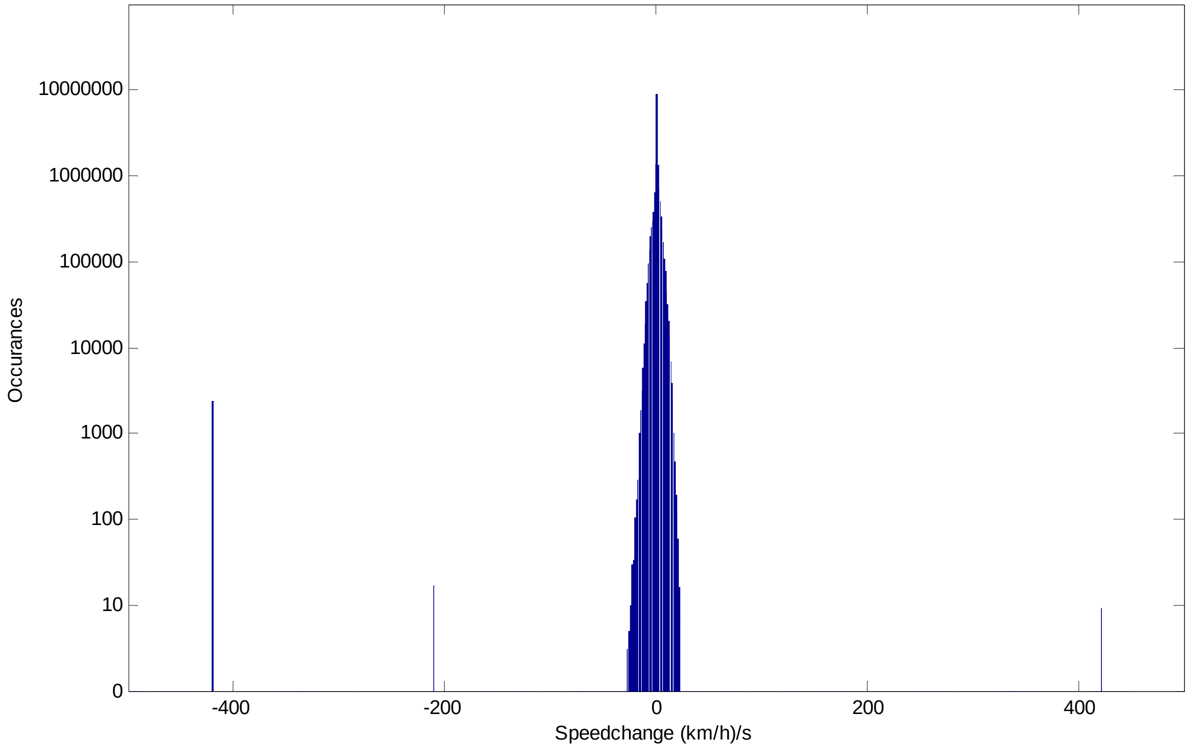

Figure 61: Occurrences of the change in CAN bus speed in subsequent seconds

Apart from the outliers that were already visible in Figure 60, all data points are relatively close to 0. When we zoom in, Figure 62 shows that all data points are between -28 km/h/s and 22 km/h/s, which are realistic values.

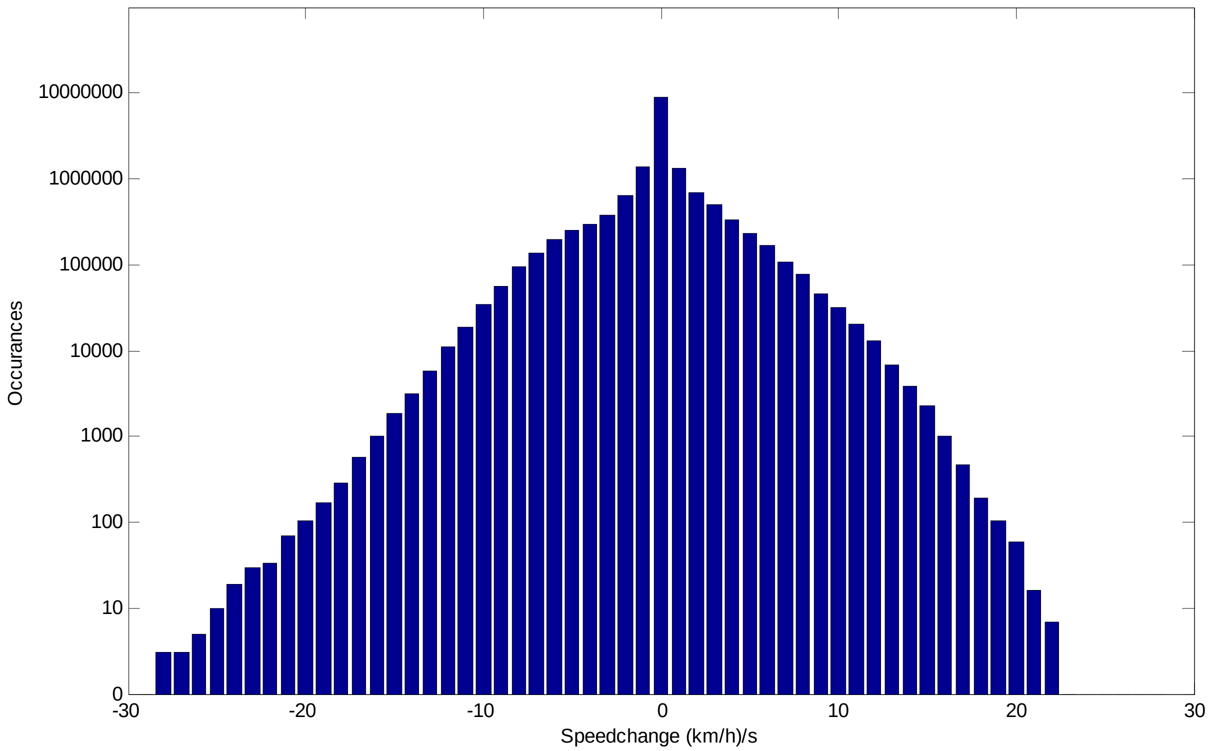

Figure 62: Occurrences of the change in CAN bus speed in subsequent seconds, without outliers

The CAN bus speed has also been checked by comparing it with the speeds that can be calculated from the GPS positions. Since the GPS positioning can be quite inaccurate in urban areas, as will be shown in the following paragraph, only data points from motorways have been taken into account in this comparison. Figure 63 shows the result of the comparison. On average, the GPS speed appears to be 6% smaller than the CAN bus speed, with a standard deviation of 2.57%. It is common knowledge, that vehicle speedometers usually overestimate the true vehicle speed. This offers an explanation for the 6% difference.

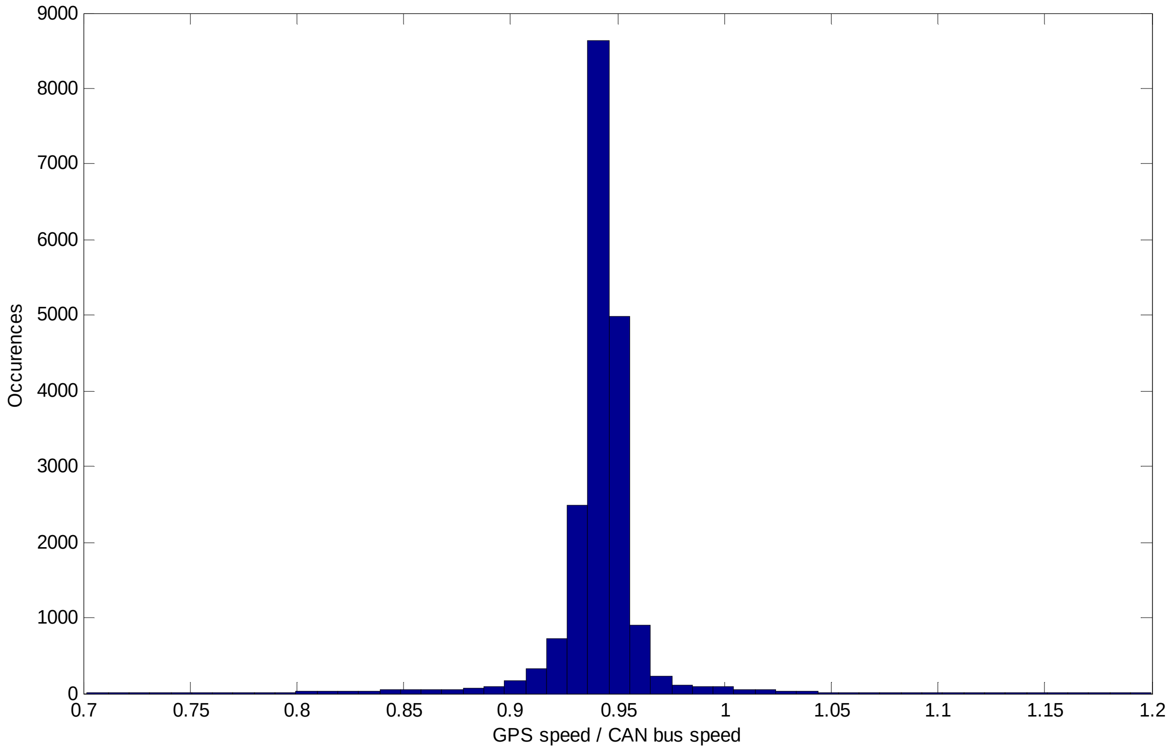

Figure 63: GPS speed / CAN bus speed - ratio

1.3.2 GPS position

Positional data acquired with GPS has an accuracy of a few meters. Objects like buildings or mountains can introduce an extra error. Modsching et al. [46] have found the average position error in urban areas ranges from 2 meters on an open square to 15 meters in wide streets with four story houses on both sides. This was already visible in some of the figures in the previous paragraph. In the centre of Eindhoven, where the buildings are quite tall and close to the streets, the positional data shows many anomalies. Figure 64 shows the recorded GPS positions near the Vestdijk in the centre of Eindhoven. A picture of the builtup alongside the Vestdijk can be found in Figure 65.

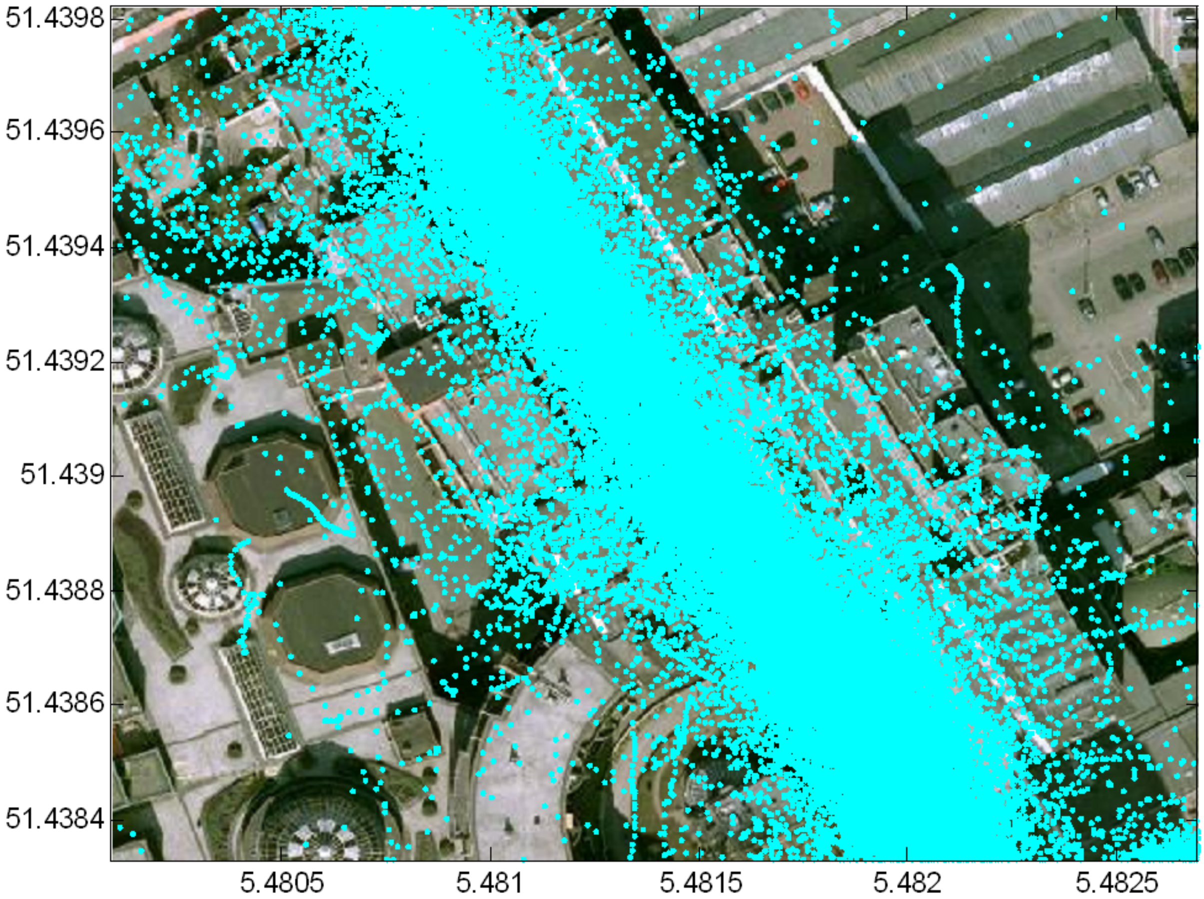

Figure 64: GPS positions at the Vestdijk, Eindhoven. The blue dots represent the recorded positions of the uCAN equipped vehicles in the data set.

Figure 65: Buildings at the Vestdijk disturbing the GPS signals [source: Google Streetview]

Also trees seem to disturb the GPS signal. The position of vehicles driving on the Oirschotsedijk in Eindhoven shows large deviations from the position of the actual road. In Figure 66 the recorded positions of the uCAN equipped vehicles are visualized. The soccer field gives an idea of the size of the error. Figure 66 shows a picture of the Oirschotsedijk with the tree canopy.

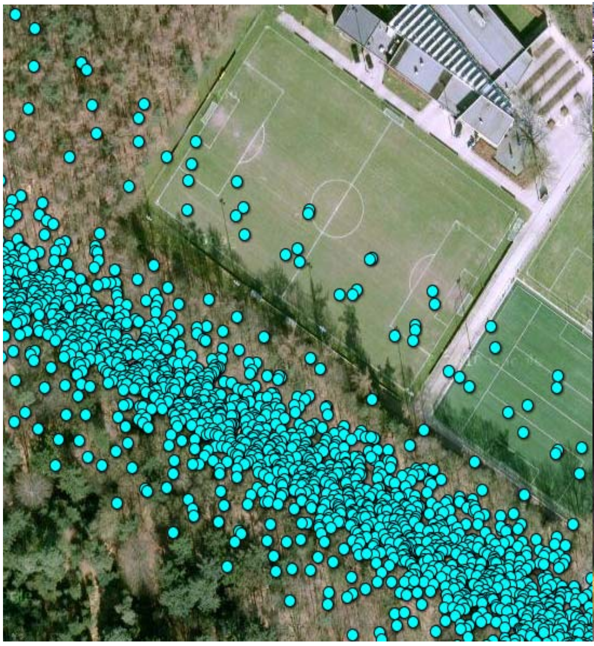

Figure 66: GPS positions at the Oirschotsedijk

Figure 67: Tree canopy disturbing the GPS signals

Most motorways are in open areas, so there the position error is relatively small, as can been seen in Figure 68. Although it is quite small, the position error is too large to accurately determine the lane or even the carriageway where the vehicle is driving.

Figure 68: GPS positioning on the A67 motorway. The blue dots represent the recorded positions of the uCAN equipped vehicles in the data set.

1.4 Conclusions

Although the uCAN equipped vehicles provide a vast amount of data, together they don't offer a high penetration rate. The frequency of the CAN bus parameters differs per vehicle and parameter, but are often around once per second. Also the available CAN bus parameters differ per vehicle.

Except for a few clear outliers which mostly occur when the vehicle engine is started or turned off, which can be easily filtered, the CAN bus data shows realistic values. When comparing some of the CAN bus parameters, they also show expected relationships. Also, when a single variable is measured in multiple ways, the results match. Not all parameters have been thoroughly examined. The accuracy of these parameters can't be guaranteed.

The GPS position in areas with builtup close to the road and below trees can show large errors from the actual position. The GPS position on motorways is quite accurate. The unit of the fuel consumption is unknown.

When using the data for analyses, there are some limitations that will have to be kept in mind. The data is only gathered by taxi drivers. It is quite probable their driving behaviour differs from other road users. The taxis are mainly present in the city centre of Eindhoven. Their usage of the motorways is relatively small. Also, the dataset only contains data from a limited amount of vehicle types.

Appendix F

Table 24: List of the locations of the loop detectors which were used in the study

Motorway	Direction	position (km)
A2	Li	124,195
A2	Li	124,795
A2	Li	125,295
A2	Li	125,795
A2	Li	127,705
A2	Li	128,195
A2	Li	128,845
A2	Li	129,480
A2	Li	130,195
A2	Li	130,895
A2	Li	132,605
A2	Li	133,395
A2	Li	134,940
A2	Li	135,495
A2	Li	136,695
A2	Li	136,950
A2	Li	137,295
A2	Li	170,960
A2	Li	171,251
A2	Li	171,590
A2	Li	171,935
A2	Li	172,635
A2	Li	172,990
A2	Li	173,335
A2	Li	173,690
A2	Li	174,350
A2	Li	174,390
A2	Li	174,735
A2	Li	175,330
A2	Li	175,350
A2	Li	177,720
A2	Li	178,950
A2	Re	124,205
A2	Re	124,805
A2	Re	125,305
A2	Re	125,805
A2	Re	127,715
A2	Re	128,205
A2	Re	128,854
A2	Re	129,480
A2	Re	130,205
A2	Re	130,905
A2	Re	134,945
A2	Re	135,505
A2	Re	136,105
A2	Re	136,705

Motorway	Direction	position (km)
A2	Re	137,305
A2	Re	171,945
A2	Re	172,270
A2	Re	172,295
A2	Re	172,645
A2	Re	172,990
A2	Re	173,345
A2	Re	173,690
A2	Re	174,390
A2	Re	174,450
A2	Re	174,745
A2	Re	175,340
A2	Re	175,450
A2	Re	177,720
A2	Re	178,105
A2	Re	178,500
A2	Re	178,950
A58	Li	15,450
A58	Li	15,695
A58	Li	17,660
A58	Li	18,355
A58	Li	20,780
A58	Li	21,545
A58	Li	22,625
A58	Li	22,950
A58	Li	23,295
A58	Li	23,955
A58	Li	24,435
A58	Li	25,145
A58	Li	25,745
A58	Li	26,345
A58	Li	26,895
A58	Li	27,535
A58	Li	28,145
A58	Li	28,695
A58	Re	17,450
A58	Re	17,690
A58	Re	18,365
A58	Re	20,790
A58	Re	21,555
A58	Re	22,105
A58	Re	22,635
A58	Re	23,305
A58	Re	23,965
A58	Re	24,445
A58	Re	25,155
A58	Re	25,755
A58	Re	26,355
A58	Re	26,905
A58	Re	27,545
A58	Re	28,155
A58	Re	28,705

Motorway	Direction	position (km)
A67	Li	28,710
A67	Li	29,345
A67	Li	29,400
A67	Li	29,795
A67	Li	30,150
A67	Li	30,495
A67	Li	30,850
A67	Li	31,208
A67	Li	31,510
A67	Li	31,795
A67	Li	32,100
A67	Li	32,395
A67	Li	32,735
A67	Li	33,450
A67	Li	33,480
A67	Li	33,895
A67	Li	34,290
A67	Li	36,195
A67	Li	36,465
A67	Re	24,695
A67	Re	25,122
A67	Re	28,710
A67	Re	29,150
A67	Re	29,345
A67	Re	29,805
A67	Re	30,150
A67	Re	30,505
A67	Re	30,850
A67	Re	31,213
A67	Re	31,510
A67	Re	31,805
A67	Re	32,100
A67	Re	32,405
A67	Re	32,735
A67	Re	33,480
A67	Re	33,550
A67	Re	33,905
A67	Re	34,290
A67	Re	34,695
A67	Re	35,480
A67	Re	35,800
A67	Re	36,205

Appendix G

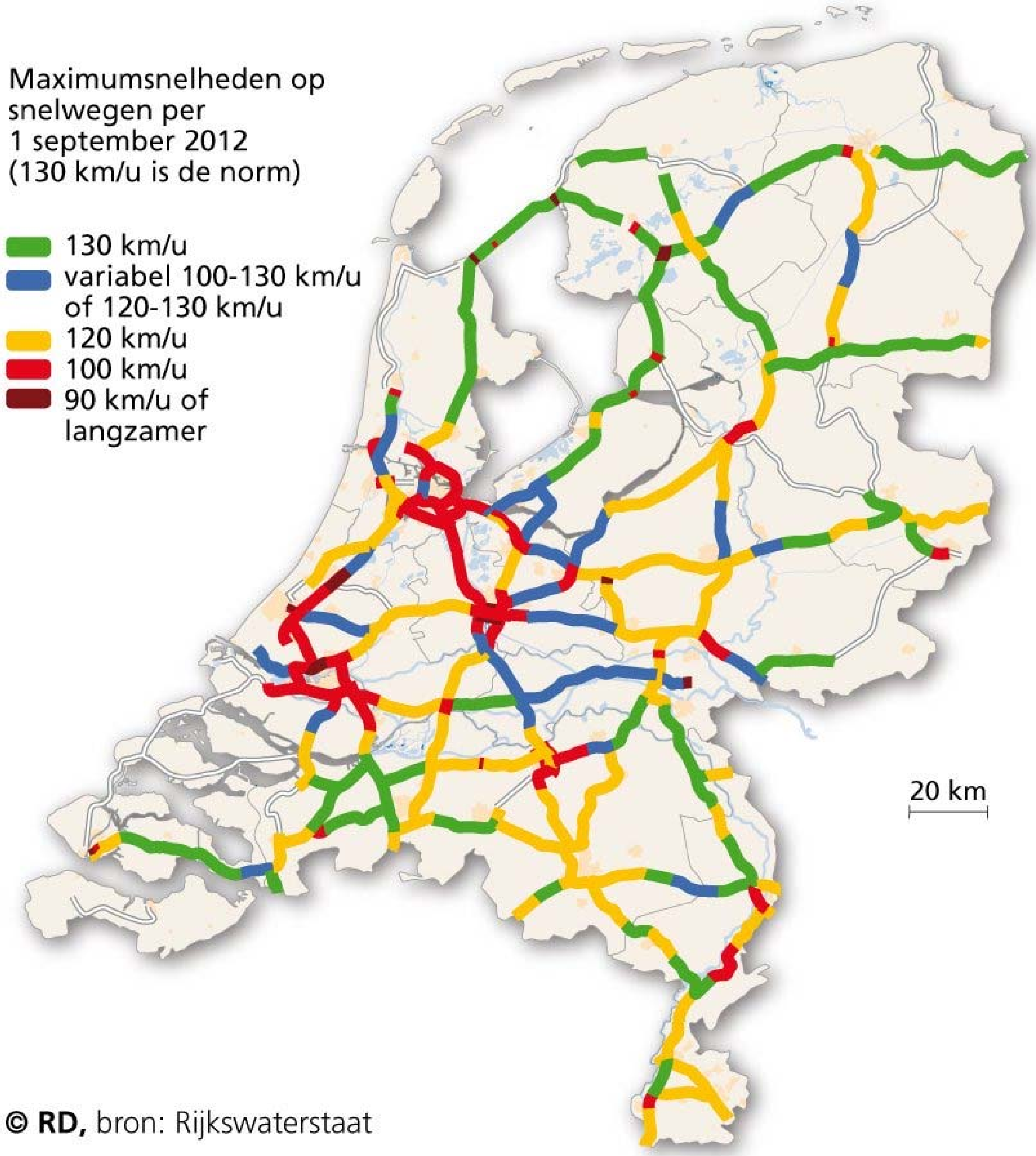

Figure 69: Speed limits on Dutch motorways [47]

Appendix H

Table 25: R values of Spearman correlation between LOS and CAN bus parameters, for 50m & 100m radii, with 1-6s CAN bus signal transmission periods

CAN bus parameter	radius = 50m						radius = 100m					
	1s	2s	3s	4s	5s	6s	1s	2s	3s	4s	5s	6s
rpm_avg	-0,45	-0,45	-0,45	-0,45	-0,46	-0,47	-0,45	-0,45	-0,45	-0,44	-0,44	-0,45
rpm_std	0,30	0,27	0,22	0,19	0,20	0,19	0,32	0,30	0,29	0,25	0,22	0,21
rpm_max	-0,46	-0,45	-0,46	-0,46	-0,46	-0,47	-0,47	-0,46	-0,46	-0,46	-0,45	-0,45
rpm_min	-0,42	-0,43	-0,44	-0,45	-0,45	-0,46	-0,40	-0,40	-0,42	-0,42	-0,42	-0,43
speed_avg	-0,46	-0,46	-0,47	-0,47	-0,47	-0,49	-0,46	-0,46	-0,46	-0,46	-0,46	-0,46
speed_std	0,28	0,26	0,23	0,19	0,20	0,19	0,30	0,29	0,28	0,25	0,22	0,22
speed_max	-0,46	-0,46	-0,47	-0,47	-0,47	-0,49	-0,47	-0,46	-0,47	-0,46	-0,46	-0,46
speed_min	-0,45	-0,46	-0,47	-0,47	-0,47	-0,49	-0,45	-0,45	-0,45	-0,45	-0,46	-0,46
brake_avg	0,12	0,12	0,11	0,11	0,12	0,10	0,16	0,15	0,15	0,14	0,14	0,13
brake_std	0,13	0,12	0,11	0,11	0,12	0,11	0,15	0,15	0,15	0,13	0,14	0,15
brake_max	0,02	0,04	0,05	0,07	0,07	0,05	0,02	0,02	0,03	0,04	0,04	0,04
brake_min	0,13	0,12	0,11	0,11	0,12	0,10	0,16	0,15	0,15	0,14	0,14	0,13
steeringposition_avg	0,03	0,03	0,02	0,01	0,01	0,01	0,03	0,03	0,03	0,02	0,01	0,01
steeringposition_std	0,09	0,17	0,20	0,19	0,19	0,18	0,05	0,05	0,08	0,14	0,17	0,20
steeringposition_min	0,07	0,07	0,05	0,03	0,02	0,02	0,06	0,06	0,06	0,05	0,04	0,04
steeringposition_max	-0,03	-0,02	-0,02	-0,01	0,00	-0,01	-0,02	-0,02	-0,02	-0,02	-0,03	-0,02
steeringofcenter_avg	0,01	0,02	0,00	0,00	0,00	0,00	0,01	0,01	0,00	0,00	0,01	0,02
steeringofcenter_std	0,08	0,17	0,20	0,19	0,19	0,18	0,06	0,06	0,08	0,14	0,17	0,19
steeringofcenter_min	-0,05	-0,06	-0,06	-0,02	-0,02	-0,01	-0,05	-0,06	-0,07	-0,06	-0,04	-0,04
steeringofcenter_max	0,07	0,07	0,03	0,02	0,02	0,01	0,06	0,06	0,04	0,04	0,05	0,05
RDI_avg	-0,03	-0,03	-0,04	-0,04	-0,03	-0,04	-0,03	-0,03	-0,02	-0,03	-0,03	-0,02
RDI_std	-0,02	-0,01	0,01	0,02	NaN	NaN	-0,02	-0,02	0,00	-0,01	0,01	0,02
RDI_min	-0,02	-0,04	-0,04	-0,04	-0,03	-0,04	-0,03	-0,03	-0,04	-0,02	-0,04	-0,03
RDI_max	-0,03	-0,03	-0,04	-0,04	-0,03	-0,04	-0,03	-0,03	-0,02	-0,03	-0,03	-0,02
LDI_avg	0,01	0,00	0,00	-0,01	0,00	-0,01	-0,01	0,00	0,00	-0,01	0,00	0,00
LDI_std	0,00	0,01	0,03	0,01	NaN	NaN	-0,01	0,00	0,00	0,00	0,01	0,03
LDI_min	0,00	-0,01	-0,02	-0,02	0,00	-0,01	-0,01	-0,01	0,00	-0,02	0,00	-0,02
LDI_max	0,01	0,00	0,00	-0,01	0,00	-0,01	-0,01	0,00	0,00	-0,01	0,00	0,00
FrontFogLight_avg	-0,06	-0,06	-0,07	-0,07	-0,05	-0,07	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06
FrontFogLight_std	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
FrontFogLight_min	-0,06	-0,06	-0,07	-0,07	-0,05	-0,07	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06
FrontFogLight_max	-0,06	-0,06	-0,07	-0,07	-0,05	-0,07	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06
WiperFast_avg	-0,03	-0,03	-0,02	-0,01	0,00	0,00	-0,02	-0,03	-0,03	-0,02	-0,01	-0,02
WiperFast_std	-0,01	-0,01	0,03	0,04	0,03	0,04	-0,01	-0,01	-0,01	0,00	0,01	0,01
WiperFast_min	-0,03	-0,03	-0,03	-0,02	-0,01	-0,01	-0,02	-0,02	-0,03	-0,03	-0,03	-0,03
WiperFast_max	-0,03	-0,03	-0,02	-0,01	0,00	0,00	-0,02	-0,03	-0,03	-0,02	-0,01	-0,02

Table 26: R values of Spearman correlation between LOS and CAN bus parameters, for 200m & 400m radii, with 1-6s CAN bus signal transmission periods

CAN bus parameter	radius = 200m						radius = 400m					
	1s	2s	3s	4s	5s	6s	1s	2s	3s	4s	5s	6s
rpm_avg	-0,45	-0,45	-0,45	-0,45	-0,45	-0,45	-0,46	-0,46	-0,46	-0,46	-0,46	-0,46
rpm_std	0,33	0,33	0,32	0,31	0,30	0,28	0,34	0,34	0,33	0,32	0,32	0,32
rpm_max	-0,48	-0,48	-0,48	-0,47	-0,47	-0,47	-0,49	-0,49	-0,49	-0,49	-0,49	-0,49
rpm_min	-0,37	-0,38	-0,39	-0,39	-0,40	-0,42	-0,32	-0,33	-0,35	-0,35	-0,36	-0,38
speed_avg	-0,46	-0,46	-0,46	-0,46	-0,46	-0,47	-0,47	-0,47	-0,47	-0,47	-0,47	-0,47
speed_std	0,32	0,32	0,32	0,30	0,29	0,27	0,33	0,33	0,32	0,31	0,31	0,31
speed_max	-0,48	-0,48	-0,48	-0,47	-0,47	-0,47	-0,49	-0,49	-0,49	-0,49	-0,49	-0,49
speed_min	-0,44	-0,44	-0,44	-0,44	-0,44	-0,45	-0,42	-0,42	-0,43	-0,43	-0,43	-0,44
brake_avg	0,19	0,19	0,18	0,16	0,18	0,16	0,22	0,22	0,21	0,20	0,21	0,19
brake_std	0,19	0,19	0,18	0,16	0,18	0,16	0,22	0,22	0,21	0,19	0,21	0,19
brake_max	0,02	0,02	0,02	0,02	0,01	0,03	NaN	0,02	0,00	0,02	0,02	0,02
brake_min	0,19	0,19	0,18	0,17	0,18	0,16	0,22	0,22	0,21	0,20	0,21	0,19
steeringposition_avg	0,04	0,04	0,04	0,03	0,02	0,03	0,05	0,05	0,04	0,04	0,03	0,04
steeringposition_std	0,05	0,05	0,05	0,05	0,06	0,08	0,04	0,05	0,04	0,05	0,04	0,04
steeringposition_min	0,08	0,08	0,08	0,07	0,05	0,06	0,09	0,09	0,08	0,08	0,07	0,07
steeringposition_max	-0,01	-0,01	-0,01	-0,02	-0,03	-0,02	-0,02	0,00	0,00	0,00	-0,02	-0,01
steeringofcenter_avg	0,00	0,00	0,00	0,00	0,01	0,01	0,00	0,00	0,00	-0,01	0,00	0,01
steeringofcenter_std	0,07	0,07	0,06	0,06	0,05	0,09	0,06	0,06	0,06	0,05	0,04	0,05
steeringofcenter_min	-0,05	-0,07	-0,06	-0,06	-0,04	-0,06	-0,05	-0,06	-0,05	-0,06	-0,04	-0,05
steeringofcenter_max	0,08	0,07	0,07	0,06	0,05	0,07	0,08	0,07	0,07	0,06	0,06	0,07
RDI_avg	-0,04	-0,03	-0,02	-0,02	-0,03	-0,02	-0,04	-0,03	-0,02	-0,03	-0,03	-0,01
RDI_std	-0,03	-0,03	-0,02	-0,01	-0,02	0,00	-0,04	-0,03	-0,02	-0,03	-0,02	0,00
RDI_min	-0,02	-0,02	-0,01	-0,03	-0,03	-0,04	-0,02	-0,02	-0,02	-0,01	-0,02	-0,02
RDI_max	-0,04	-0,03	-0,02	-0,02	-0,03	-0,02	-0,04	-0,03	-0,02	-0,03	-0,03	-0,01
LDI_avg	-0,02	-0,01	0,00	-0,01	0,02	0,00	-0,02	-0,01	0,00	-0,01	0,01	-0,01
LDI_std	-0,02	-0,01	0,00	-0,01	0,01	0,00	-0,02	-0,01	0,00	-0,01	0,01	-0,01
LDI_min	NaN	0,01	NaN	0,02	0,02	0,00	NaN	NaN	NaN	NaN	NaN	-0,01
LDI_max	-0,01	-0,01	0,00	-0,01	0,02	0,00	-0,02	-0,01	0,00	-0,01	0,01	-0,01
FrontFogLight_avg	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06
FrontFogLight_std	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
FrontFogLight_min	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06
FrontFogLight_max	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06	-0,06
WiperFast_avg	-0,02	-0,02	-0,02	-0,02	-0,01	-0,02	-0,02	-0,01	-0,02	-0,01	-0,01	-0,02
WiperFast_std	-0,02	-0,01	-0,02	-0,01	0,01	-0,01	-0,01	-0,01	-0,02	-0,01	-0,01	-0,01
WiperFast_min	-0,01	-0,02	-0,02	-0,02	-0,03	-0,03	0,00	-0,01	-0,01	-0,02	-0,02	-0,02
WiperFast_max	-0,02	-0,02	-0,02	-0,02	-0,01	-0,02	-0,01	-0,01	-0,02	-0,01	-0,01	-0,02

Table 27: P values of Spearman correlation between LOS and CAN bus parameters, for 50m & 100m radii, with 1-6s CAN bus signal transmission periods

CAN bus parameter	radius = 50m						radius = 100m					
	1s	2s	3s	4s	5s	6s	1s	2s	3s	4s	5s	6s
rpm_avg	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
rpm_std	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
rpm_max	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
rpm_min	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
speed_avg	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
speed_std	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
speed_max	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
speed_min	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
brake_avg	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
brake_std	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
brake_max	0,06	0,01	0,00	0,00	0,00	0,00	0,11	0,05	0,01	0,00	0,00	0,00
brake_min	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
steeringposition_avg	0,03	0,02	0,22	0,49	0,42	0,68	0,02	0,02	0,03	0,12	0,36	0,37
steeringposition_std	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
steeringposition_min	0,00	0,00	0,00	0,08	0,12	0,30	0,00	0,00	0,00	0,00	0,00	0,00
steeringposition_max	0,03	0,06	0,09	0,42	0,85	0,69	0,17	0,22	0,19	0,08	0,02	0,07
steeringofcenter_avg	0,29	0,12	0,71	0,96	0,82	0,90	0,64	0,68	0,86	0,91	0,32	0,19
steeringofcenter_std	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
steeringofcenter_min	0,00	0,00	0,00	0,11	0,30	0,40	0,00	0,00	0,00	0,00	0,00	0,01
steeringofcenter_max	0,00	0,00	0,02	0,29	0,31	0,44	0,00	0,00	0,00	0,00	0,00	0,00
RDI_avg	0,01	0,01	0,01	0,01	0,12	0,04	0,04	0,02	0,10	0,04	0,02	0,11
RDI_std	0,09	0,36	0,38	0,08	NaN	NaN	0,12	0,16	0,80	0,24	0,41	0,11
RDI_min	0,07	0,00	0,00	0,00	0,12	0,04	0,03	0,01	0,00	0,06	0,00	0,02
RDI_max	0,01	0,01	0,01	0,01	0,12	0,04	0,04	0,02	0,10	0,04	0,02	0,11
LDI_avg	0,65	0,89	0,88	0,31	0,77	0,67	0,35	0,85	0,98	0,52	0,77	0,77
LDI_std	0,70	0,55	0,01	0,69	NaN	NaN	0,36	0,89	0,95	0,84	0,66	0,02
LDI_min	0,99	0,50	0,25	0,28	0,77	0,67	0,61	0,33	0,91	0,11	0,83	0,22
LDI_max	0,64	0,89	0,88	0,31	0,77	0,67	0,36	0,88	0,97	0,53	0,76	0,79
FrontFogLight_avg	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
FrontFogLight_std	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
FrontFogLight_min	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
FrontFogLight_max	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
WiperFast_avg	0,04	0,02	0,16	0,31	0,89	0,83	0,06	0,04	0,03	0,11	0,27	0,12
WiperFast_std	0,31	0,39	0,04	0,01	0,09	0,02	0,33	0,29	0,26	0,98	0,30	0,44
WiperFast_min	0,02	0,02	0,03	0,12	0,75	0,53	0,10	0,06	0,02	0,05	0,03	0,03
WiperFast_max	0,04	0,02	0,16	0,31	0,89	0,83	0,06	0,04	0,03	0,11	0,27	0,12

Table 28: P values of Spearman correlation between LOS and CAN bus parameters, for 200m & 400m radii, with 1-6s CAN bus signal transmission periods

CAN bus parameter	radius = 200m						radius = 400m					
	1s	2s	3s	4s	5s	6s	1s	2s	3s	4s	5s	6s
rpm_avg	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
rpm_std	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
rpm_max	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
rpm_min	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
speed_avg	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
speed_std	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
speed_max	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
speed_min	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
brake_avg	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
brake_std	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
brake_max	0,11	0,09	0,05	0,20	0,54	0,01	NaN	0,05	0,90	0,05	0,05	0,20
brake_min	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
steeringposition_avg	0,00	0,00	0,00	0,04	0,19	0,02	0,00	0,00	0,00	0,00	0,01	0,00
steeringposition_std	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
steeringposition_min	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
steeringposition_max	0,25	0,55	0,35	0,17	0,01	0,16	0,15	0,76	0,78	0,75	0,11	0,45
steeringofcenter_avg	0,90	0,89	0,78	0,93	0,61	0,30	0,89	0,91	0,80	0,66	0,76	0,37
steeringofcenter_std	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
steeringofcenter_min	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
steeringofcenter_max	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
RDI_avg	0,00	0,02	0,08	0,07	0,02	0,19	0,00	0,01	0,08	0,01	0,02	0,57
RDI_std	0,01	0,04	0,14	0,24	0,12	0,70	0,00	0,02	0,13	0,02	0,05	0,86
RDI_min	0,14	0,09	0,28	0,02	0,03	0,00	0,14	0,12	0,16	0,38	0,12	0,06
RDI_max	0,00	0,02	0,09	0,07	0,02	0,21	0,00	0,02	0,10	0,02	0,03	0,62
LDI_avg	0,22	0,39	0,81	0,49	0,22	0,82	0,06	0,27	0,71	0,46	0,61	0,56
LDI_std	0,22	0,38	0,81	0,40	0,32	0,80	0,06	0,28	0,75	0,48	0,60	0,62
LDI_min	NaN	0,62	NaN	0,20	0,09	0,91	NaN	NaN	NaN	NaN	NaN	0,47
LDI_max	0,24	0,41	0,78	0,51	0,21	0,84	0,07	0,32	0,82	0,52	0,56	0,61
FrontFogLight_avg	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
FrontFogLight_std	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
FrontFogLight_min	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
FrontFogLight_max	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
WiperFast_avg	0,07	0,16	0,07	0,16	0,42	0,19	0,22	0,32	0,10	0,27	0,27	0,18
WiperFast_std	0,17	0,41	0,22	0,39	0,66	0,51	0,38	0,56	0,12	0,52	0,38	0,42
WiperFast_min	0,34	0,20	0,16	0,06	0,02	0,04	0,75	0,42	0,59	0,11	0,22	0,18
WiperFast_max	0,07	0,16	0,07	0,16	0,42	0,19	0,23	0,33	0,10	0,27	0,27	0,18

Table 29: R values Pearson correlation between CAN bus parameters and speed, intensity and density, for 200m radius and 1s CAN bus signal transmission period.

CAN bus parameter	Speed	Intensity	Density
rpm_avg	0,62	-0,32	-0,47
rpm_std	-0,42	0,14	0,29
rpm_max	0,68	-0,33	-0,51
rpm_min	0,37	-0,26	-0,31
speed_avg	0,76	-0,29	-0,53
speed_std	-0,48	0,16	0,34
speed_max	0,77	-0,30	-0,55
speed_min	0,70	-0,29	-0,50
brake_avg	-0,33	0,07	0,22
brake_std	-0,30	0,10	0,22
brake_max	-0,03	0,01	0,02
brake_min	-0,30	0,11	0,23
steeringposition_avg	-0,08	-0,02	0,01
steeringposition_std	-0,13	0,04	0,08
steeringposition_min	-0,22	0,02	0,11
steeringposition_max	0,06	-0,04	-0,08
steeringofcenter_avg	-0,03	0,01	0,02
steeringofcenter_std	-0,12	0,05	0,09
steeringofcenter_min	0,04	-0,03	-0,04
steeringofcenter_max	-0,21	0,05	0,13
RDI_avg	0,02	-0,01	-0,02
RDI_std	0,03	-0,01	-0,02
RDI_min	0,01	0,00	0,00
RDI_max	0,02	-0,01	-0,02
LDI_avg	0,03	0,02	0,00
LDI_std	0,03	0,02	0,00
LDI_min	NaN	NaN	NaN
LDI_max	0,02	0,02	0,00
FrontFogLight_avg	0,08	-0,06	-0,06
FrontFogLight_std	NaN	NaN	NaN
FrontFogLight_min	0,08	-0,06	-0,06
FrontFogLight_max	0,08	-0,06	-0,06
WiperFast_avg	-0,10	-0,08	-0,04
WiperFast_std	-0,04	-0,06	-0,04
WiperFast_min	-0,10	-0,06	-0,02
WiperFast_max	-0,11	-0,08	-0,04

Table 30: P values Pearson correlation between CAN bus parameters and speed, intensity and density, for 200m radius and 1s CAN bus signal transmission period.

CAN bus parameter	Speed	Intensity	Density
rpm_avg	0,0000	0,0000	0,0000
rpm_std	0,0000	0,0000	0,0000
rpm_max	0,0000	0,0000	0,0000
rpm_min	0,0000	0,0000	0,0000
speed_avg	0,0000	0,0000	0,0000
speed_std	0,0000	0,0000	0,0000
speed_max	0,0000	0,0000	0,0000
speed_min	0,0000	0,0000	0,0000
brake_avg	0,0000	0,0000	0,0000
brake_std	0,0000	0,0000	0,0000
brake_max	0,0132	0,3206	0,0720
brake_min	0,0000	0,0000	0,0000
steeringposition_avg	0,0000	0,0640	0,6141
steeringposition_std	0,0000	0,0024	0,0000
steeringposition_min	0,0000	0,0823	0,0000
steeringposition_max	0,0000	0,0009	0,0000
steeringofcenter_avg	0,0085	0,2416	0,1899
steeringofcenter_std	0,0000	0,0002	0,0000
steeringofcenter_min	0,0010	0,0295	0,0004
steeringofcenter_max	0,0000	0,0000	0,0000
RDI_avg	0,0520	0,3860	0,1973
RDI_std	0,0391	0,5885	0,1436
RDI_min	0,4409	0,8157	0,8579
RDI_max	0,0459	0,6282	0,1872
LDI_avg	0,0065	0,0690	0,8314
LDI_std	0,0100	0,0858	0,9089
LDI_min	NaN	NaN	NaN
LDI_max	0,0514	0,0873	0,7438
FrontFogLight_avg	0,0000	0,0000	0,0000
FrontFogLight_std	NaN	NaN	NaN
FrontFogLight_min	0,0000	0,0000	0,0000
FrontFogLight_max	0,0000	0,0000	0,0000
WiperFast_avg	0,0000	0,0000	0,0041
WiperFast_std	0,0004	0,0000	0,0027
WiperFast_min	0,0000	0,0000	0,0878
WiperFast_max	0,0000	0,0000	0,0038

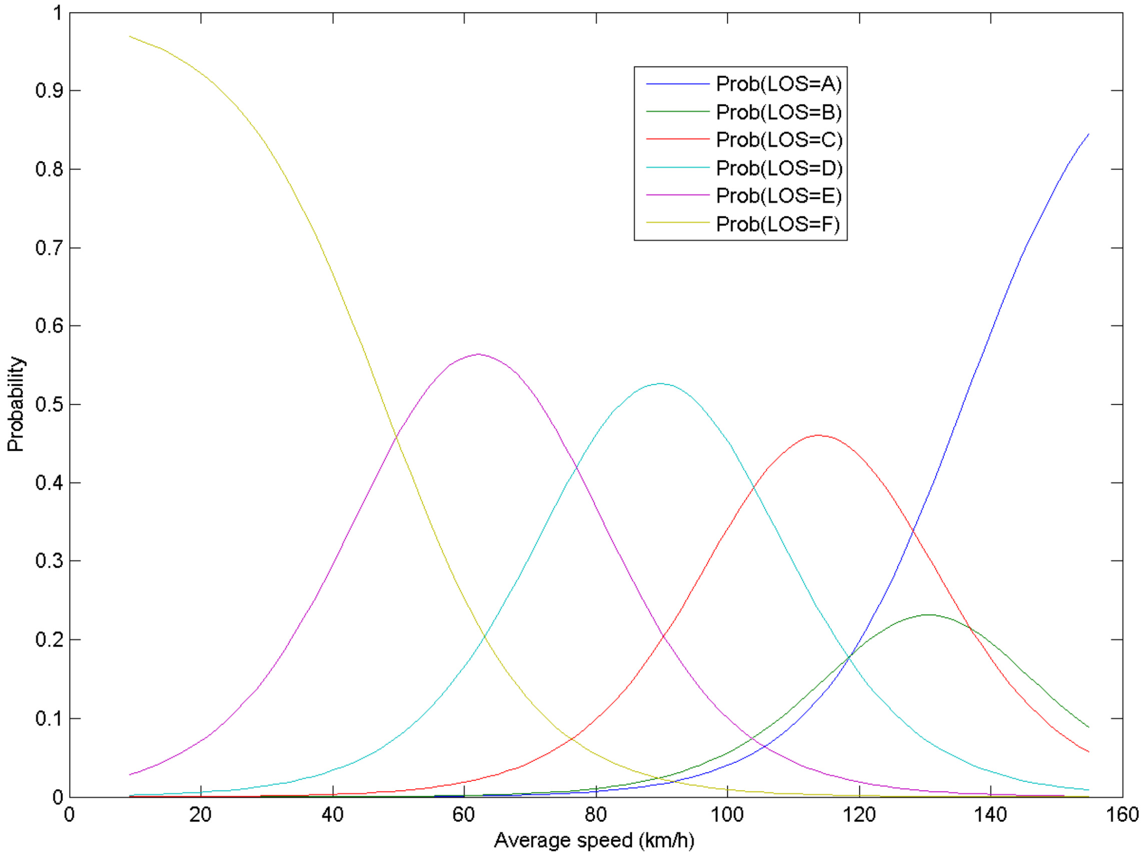

Figure 70: LOS probability functions (LOS ground truth 1)

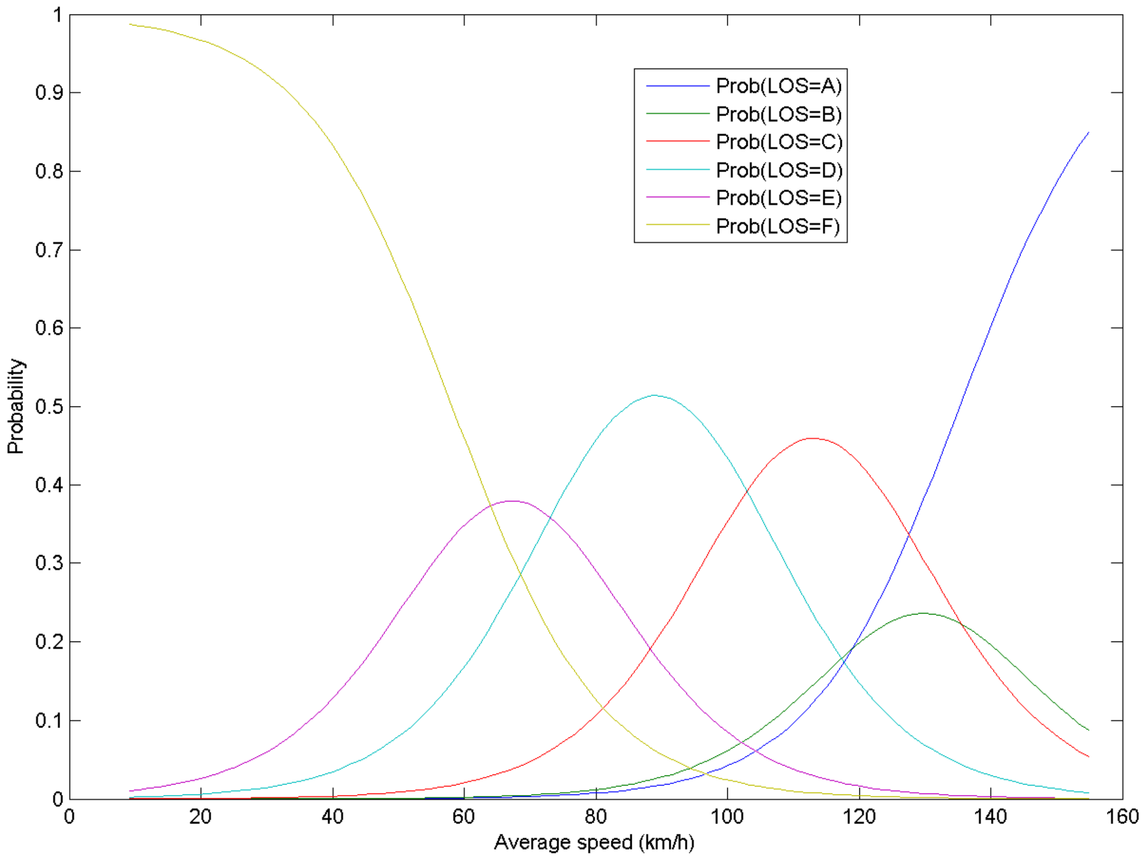

Figure 71: LOS probability functions (LOS ground truth 2)

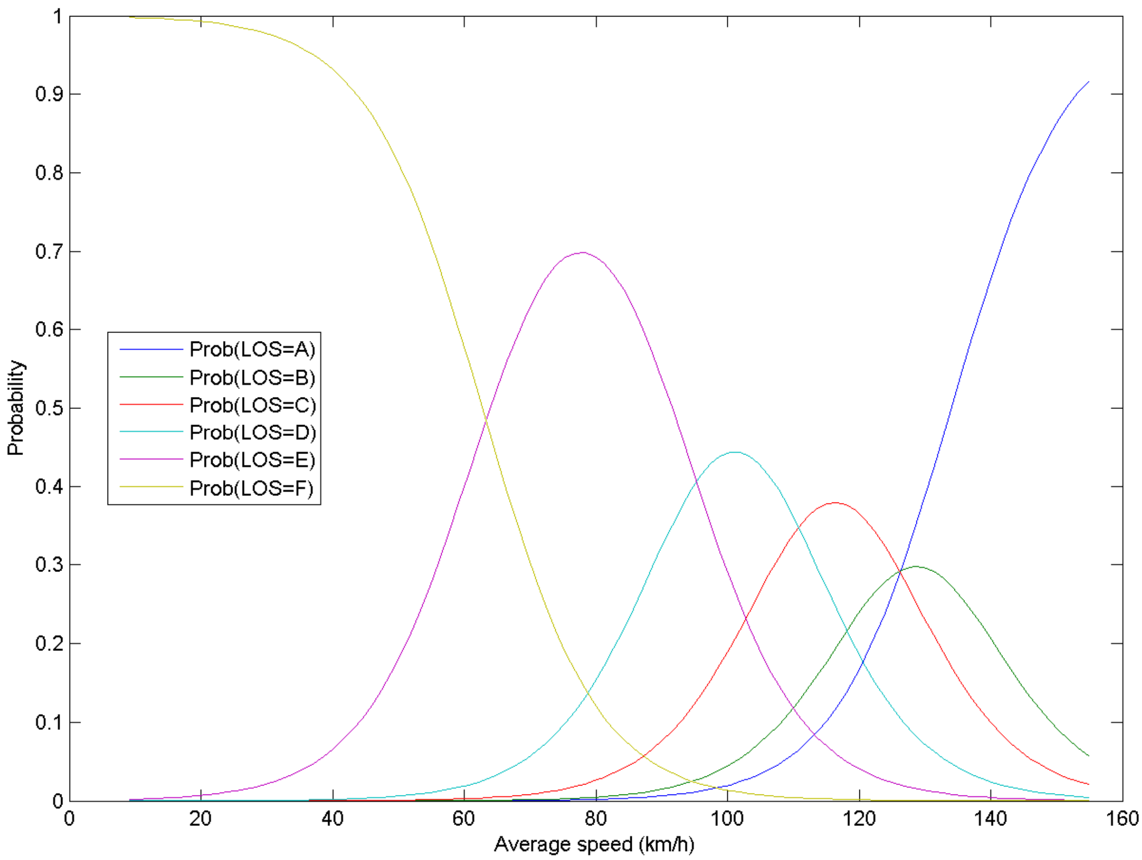

Figure 72: LOS probability functions (LOS ground truth 3)

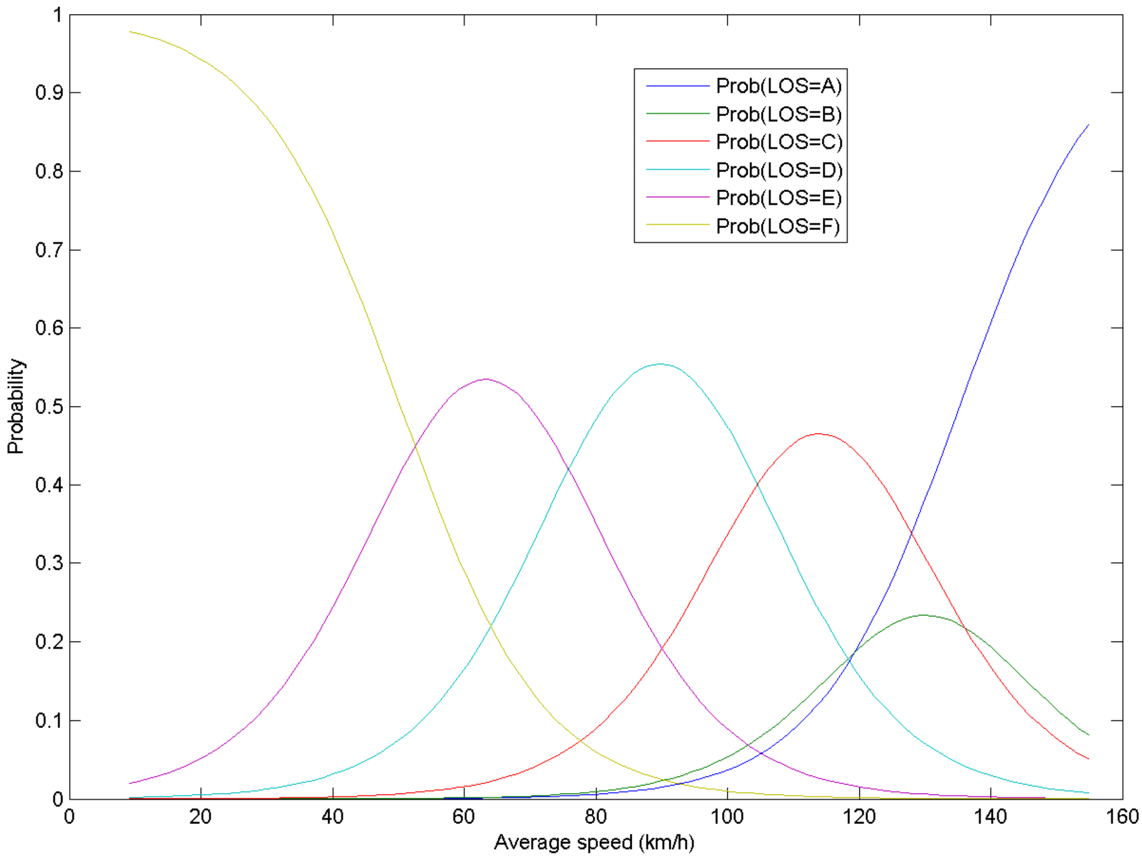

Figure 73: LOS probability functions (LOS ground truth 4)

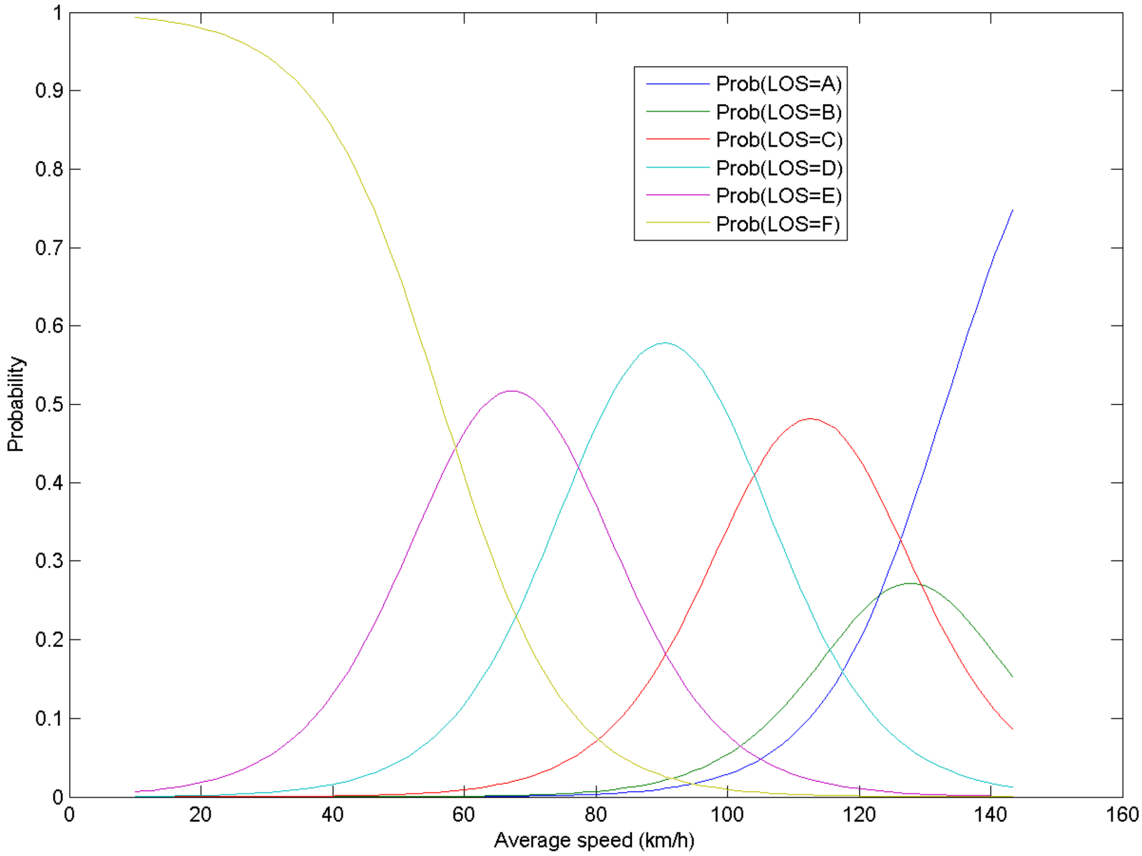

Figure 74: LOS probability functions for 1% penetration

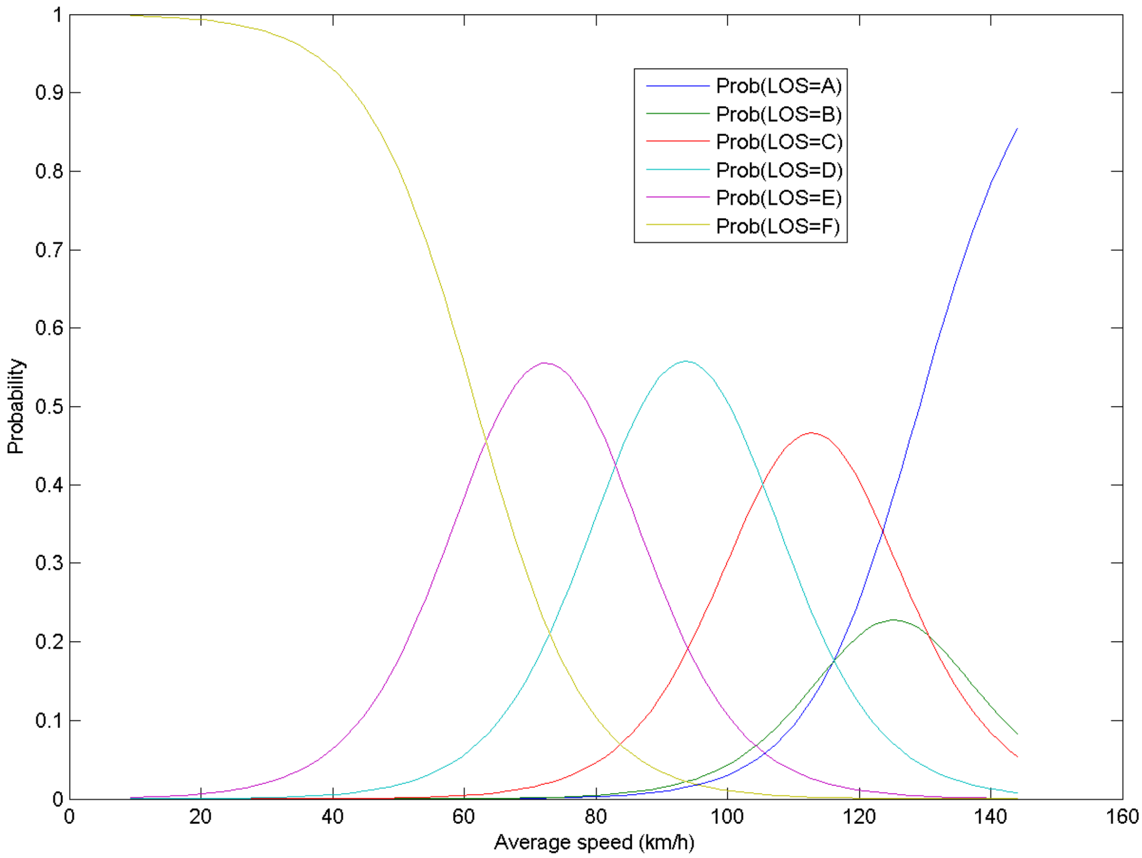

Figure 75: LOS probability functions for 2% penetration

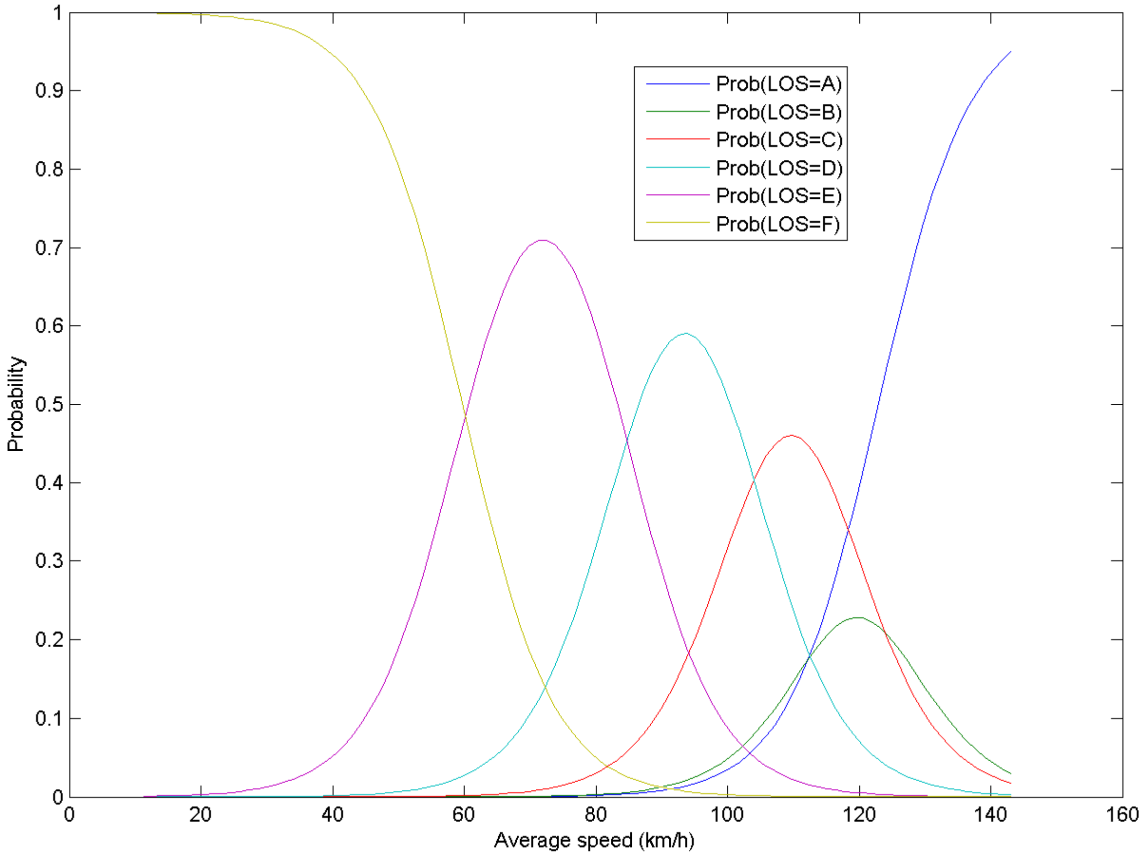

Figure 76: LOS probability functions for 5% penetration

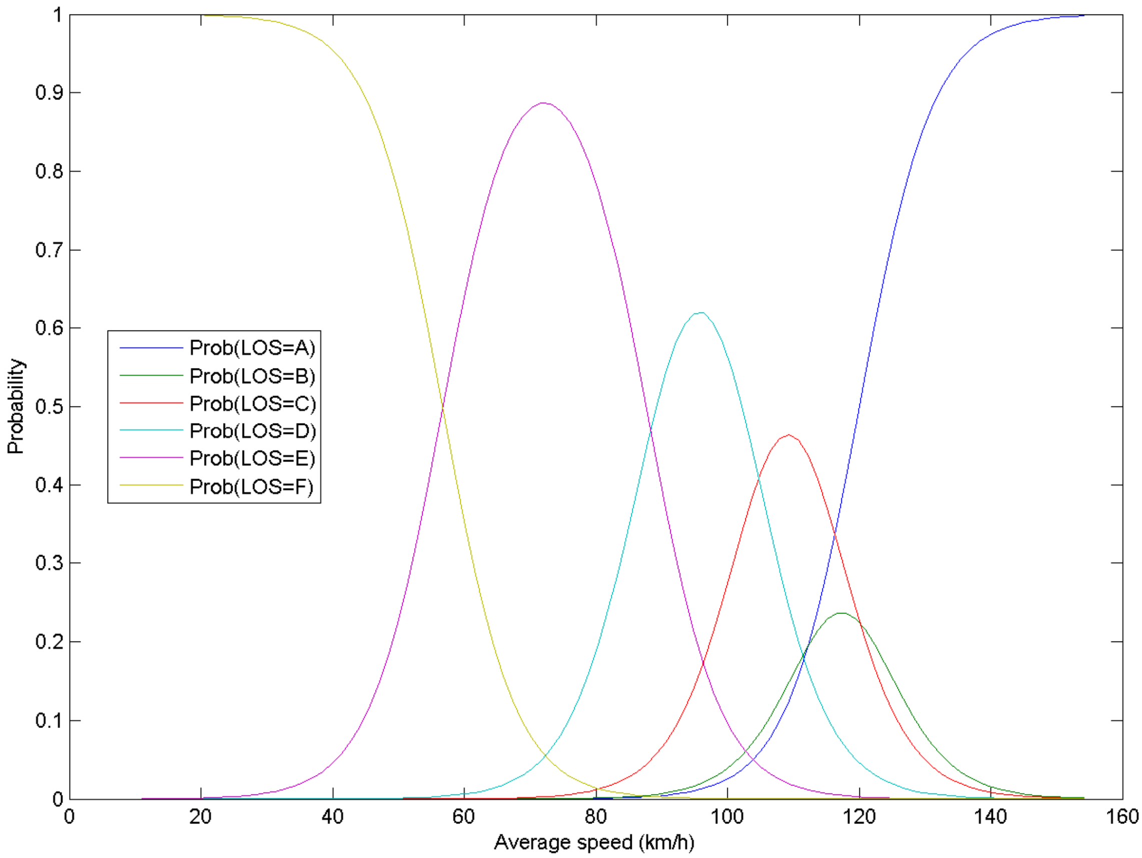

Figure 77: LOS probability functions for 10% penetration

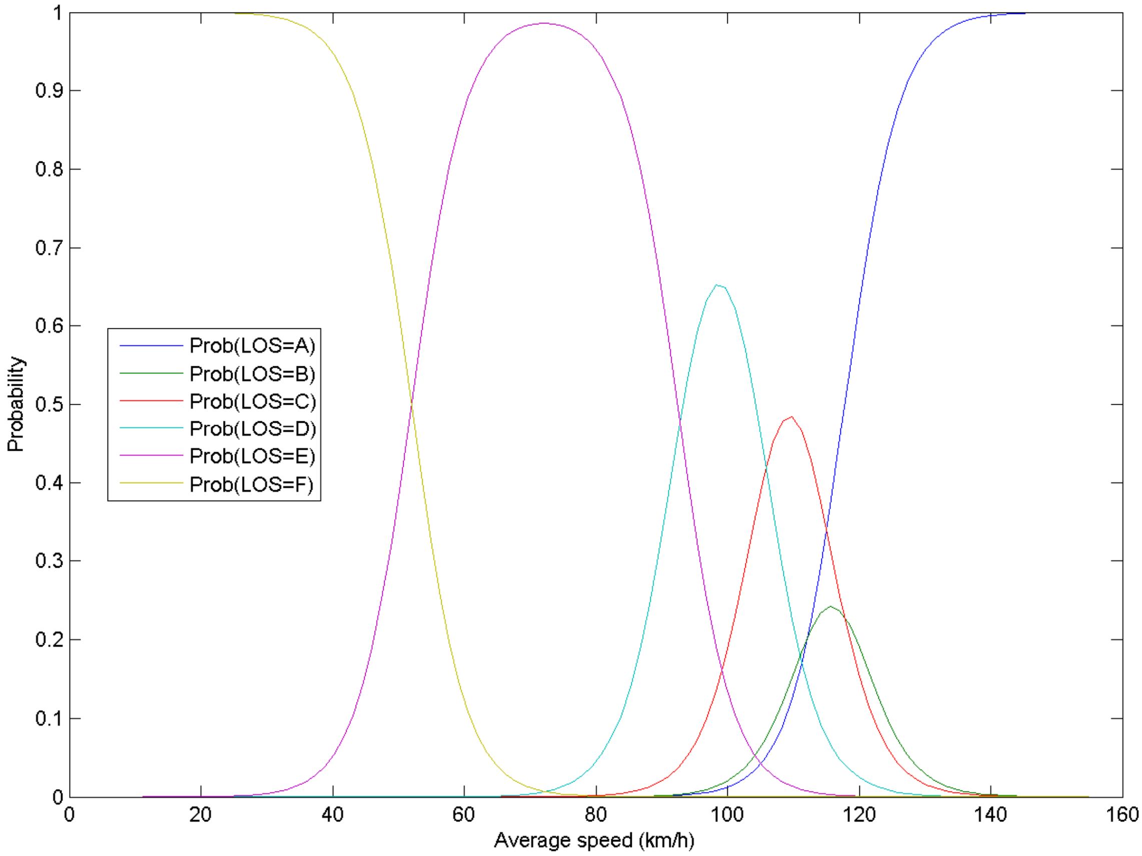

Figure 78: LOS probability functions for 20% penetration

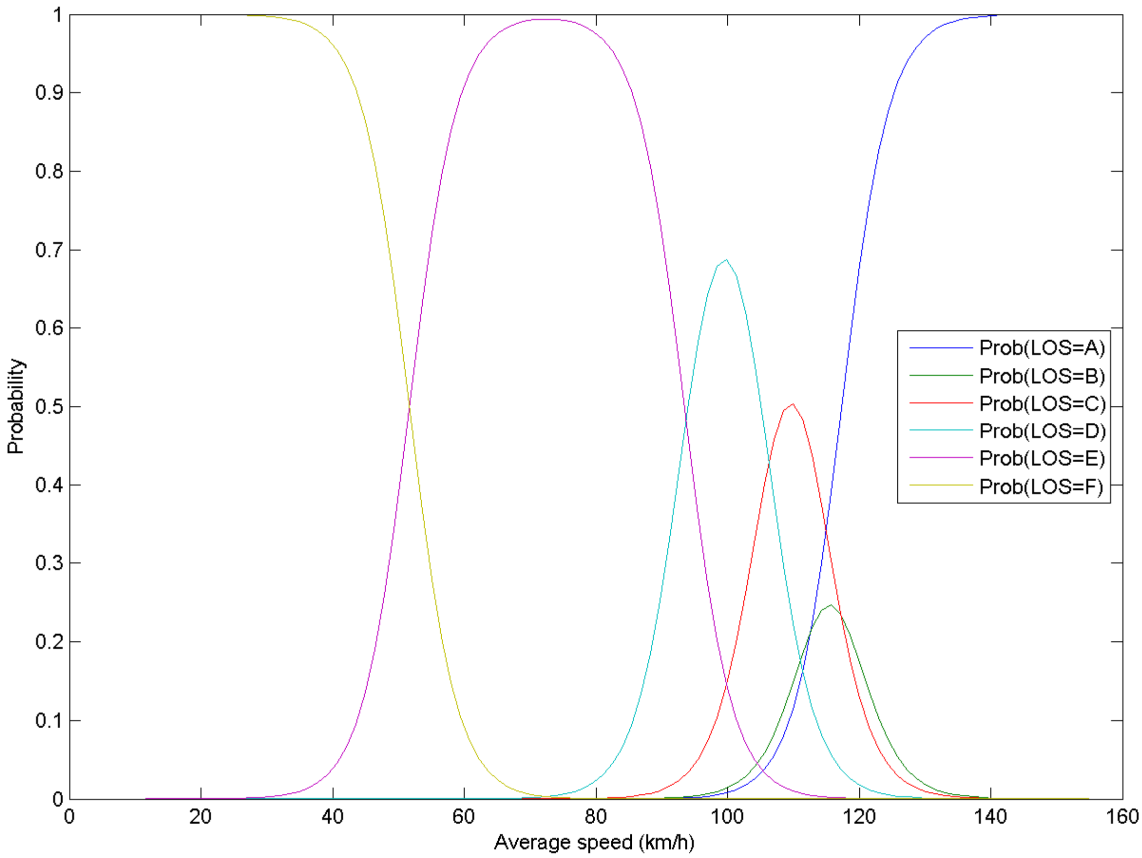

Figure 79: LOS probability functions for 30% penetration

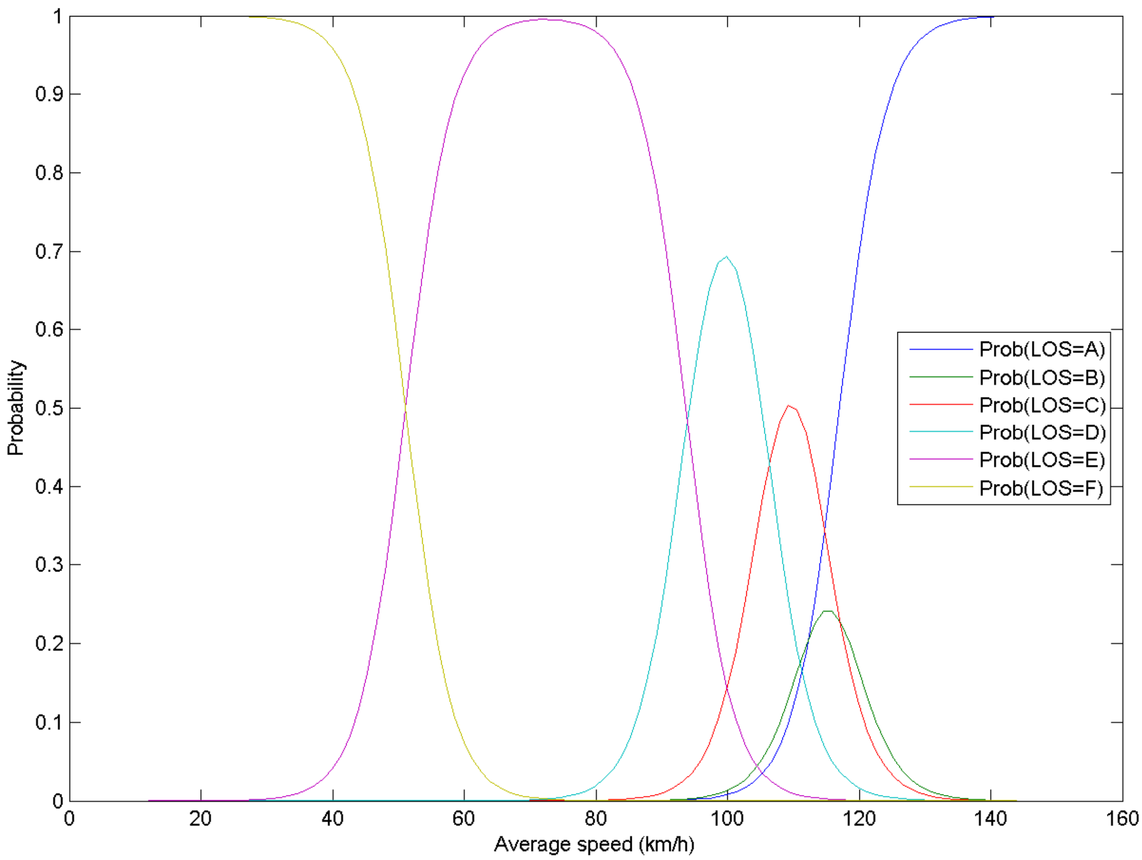

Figure 80: LOS probability functions for 40% penetration

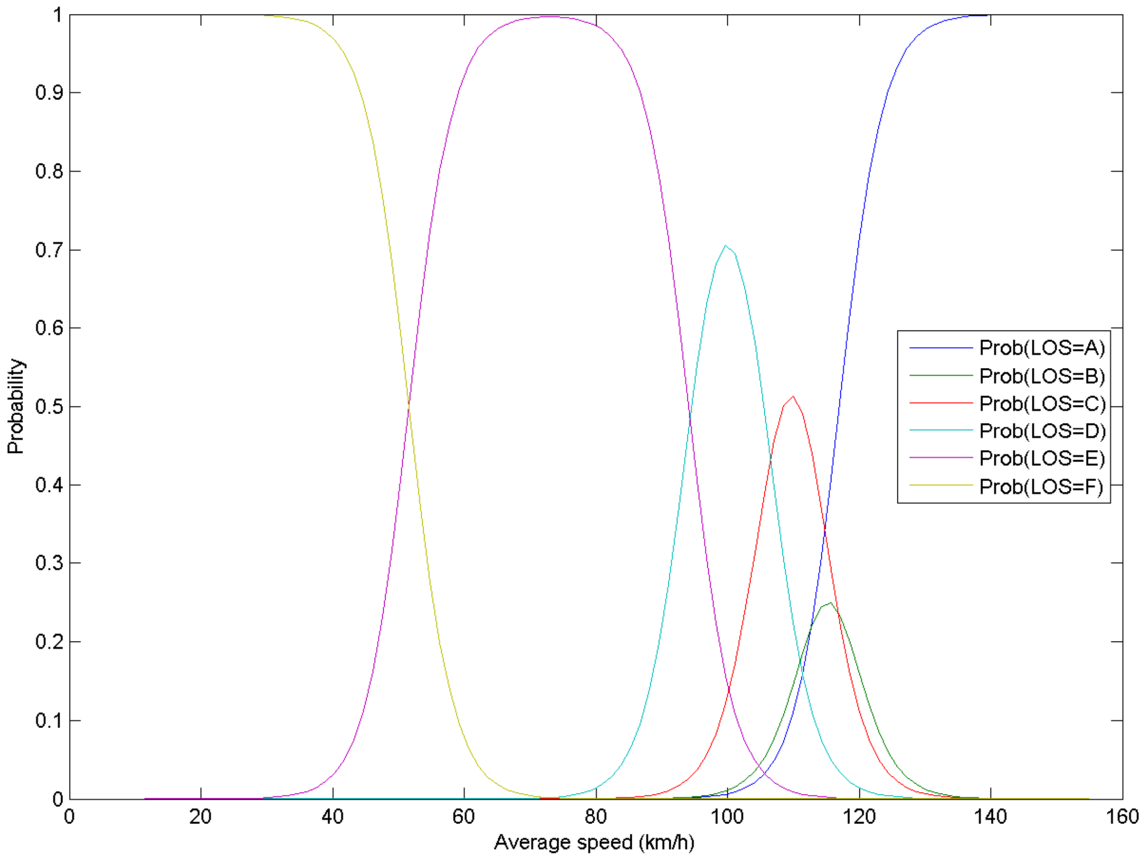

Figure 81: LOS probability functions for 50% penetration

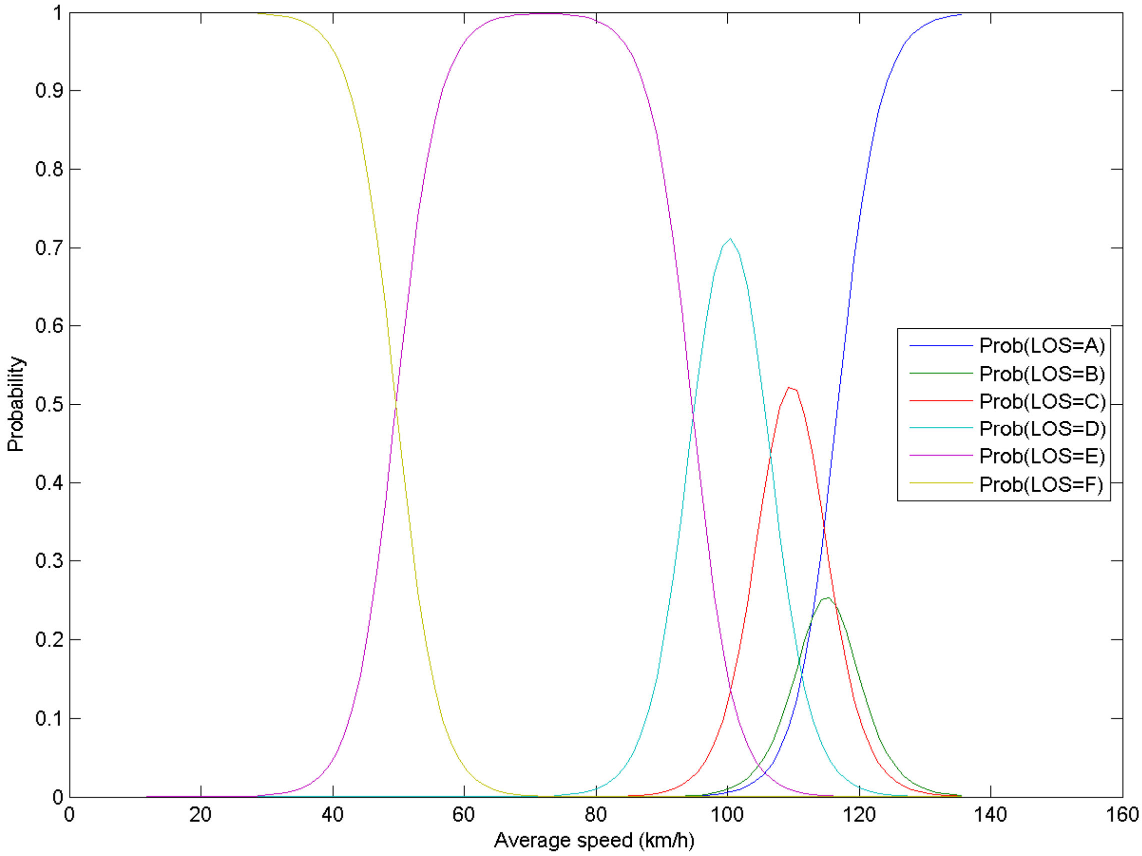

Figure 82: LOS probability functions for 60% penetration

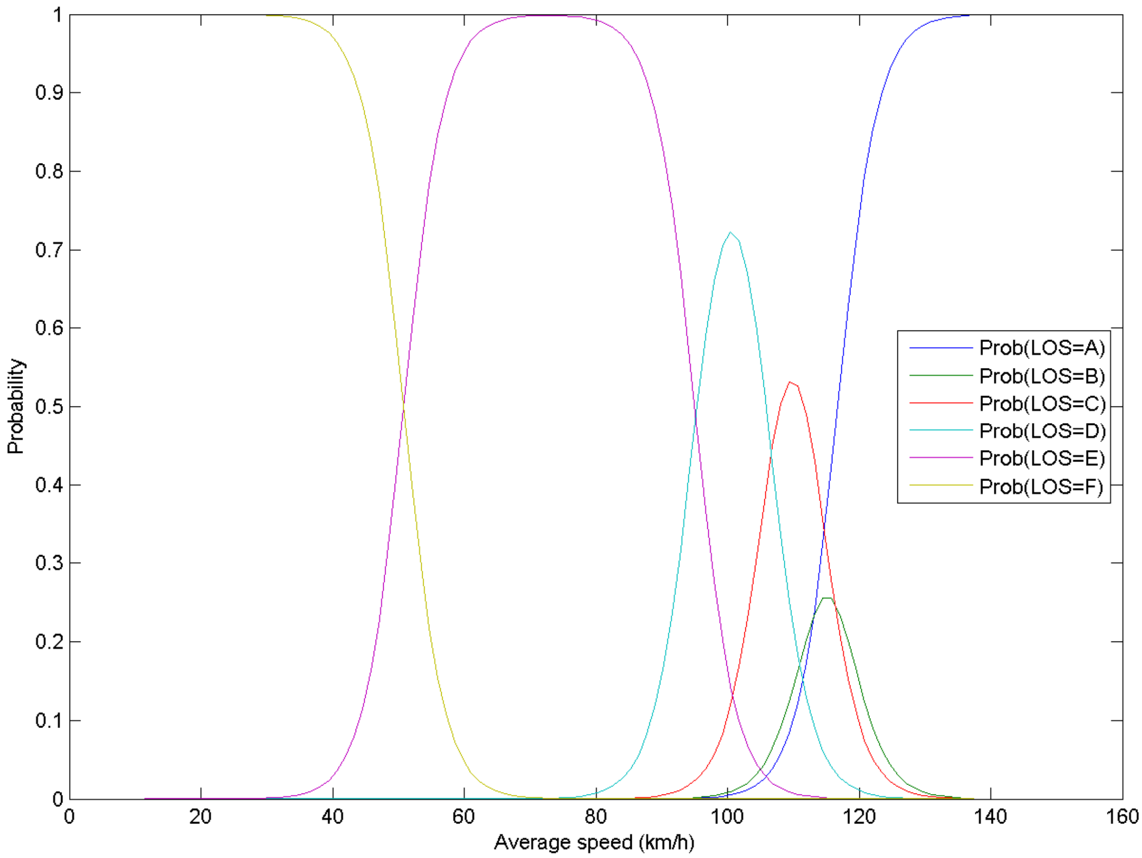

Figure 83: LOS probability functions for 70% penetration

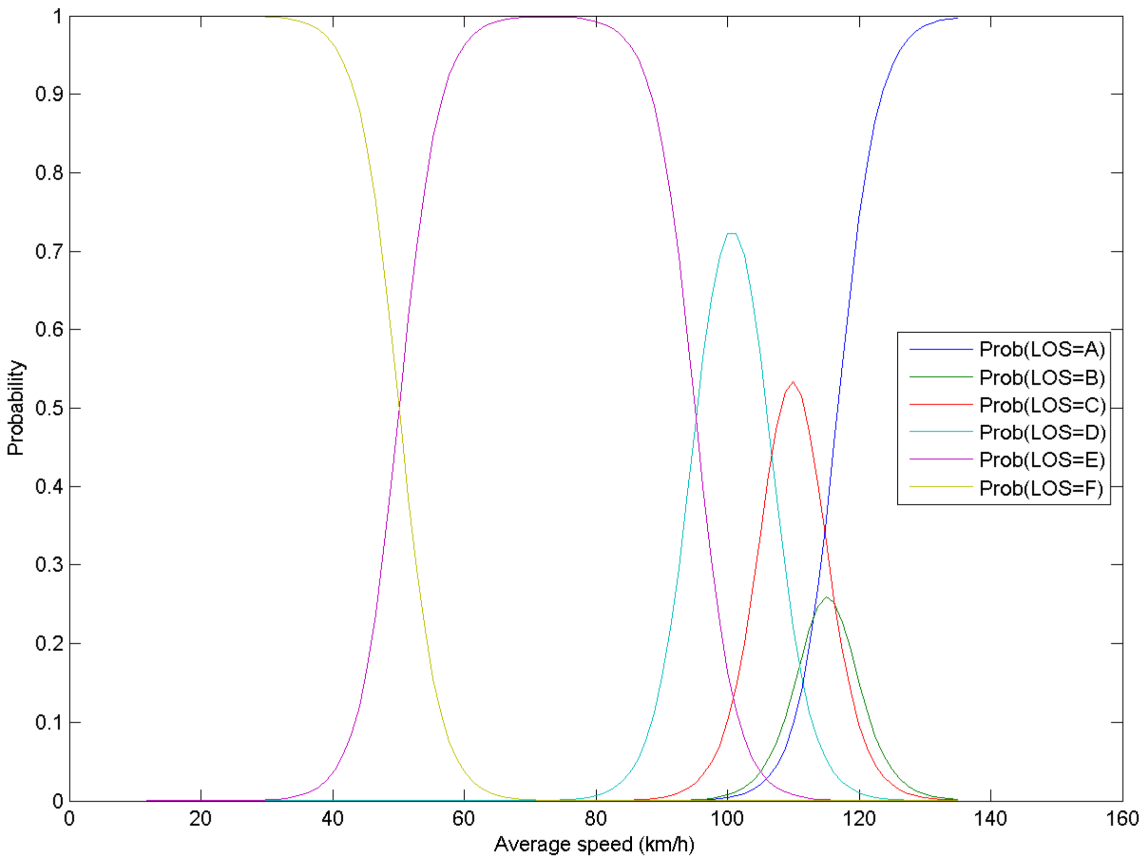

Figure 84: LOS probability functions for 80% penetration

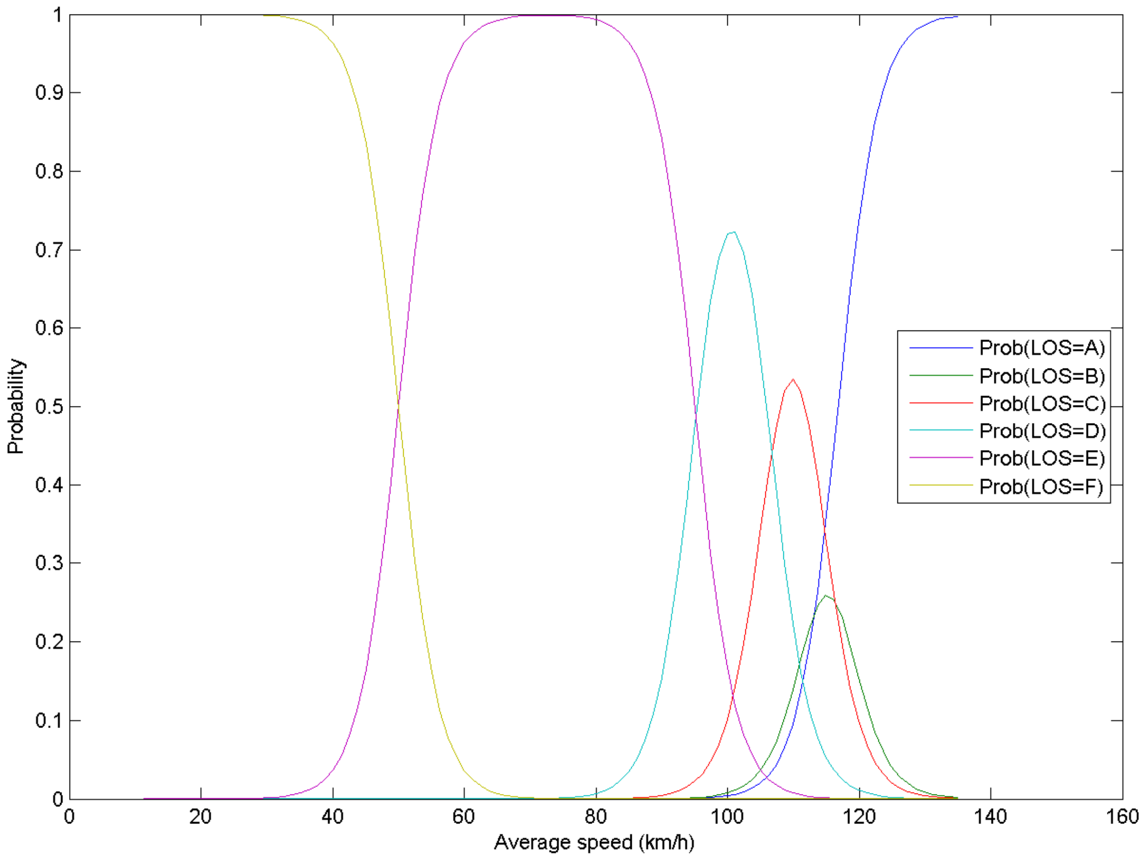

Figure 85: LOS probability functions for 90% penetration

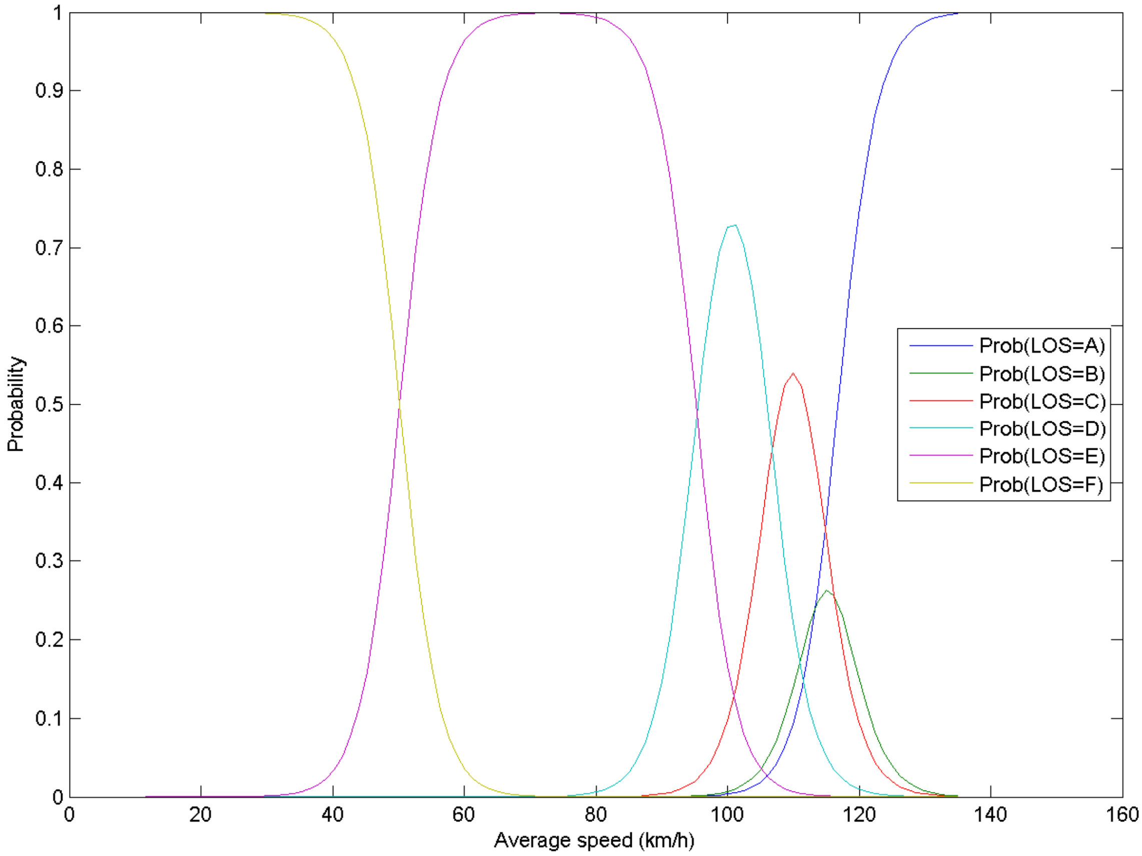

Figure 86: LOS probability functions for 100% penetration

Appendix I

In this section, the digital appendices are mentioned, which are enclosed with the report:

Appendix I.1: MATLAB code

Appendix I.2: LOS estimation accuracy loop detectors with assumed accuracy