

## Modelling driver behaviour in longitudinal vehicle-pedestrian scenarios

Analysis of driver comfort zone boundaries from naturalistic driving data and field tests

Master's thesis in Systems, Control and Mechatronics

ALEXANDER RASCH



MASTER'S THESIS IN SYSTEMS, CONTROL AND MECHATRONICS

# Modelling driver behaviour in longitudinal vehicle-pedestrian scenarios

Analysis of driver comfort zone boundaries from naturalistic driving data and field tests

ALEXANDER RASCH

Department of Mechanics and Maritime Sciences  
Division of Vehicle Safety  
CHALMERS UNIVERSITY OF TECHNOLOGY  
Göteborg, Sweden 2018

Modelling driver behaviour in longitudinal vehicle-pedestrian scenarios  
Analysis of driver comfort zone boundaries from naturalistic driving data and field tests  
ALEXANDER RASCH

© ALEXANDER RASCH, 2018

Master's thesis 2018:67  
Department of Mechanics and Maritime Sciences  
Division of Vehicle Safety  
Chalmers University of Technology  
SE-412 96 Göteborg  
Sweden  
Telephone: +46 (0)31-772 1000

Cover:

Visualisation of driver comfort and safety zone boundaries when passing a pedestrian with oncoming traffic present

Chalmers Reproservice  
Göteborg, Sweden 2018



Modelling driver behaviour in longitudinal vehicle-pedestrian scenarios  
Analysis of driver comfort zone boundaries from naturalistic driving data and field tests  
Master's thesis in Systems, Control and Mechatronics  
ALEXANDER RASCH  
Department of Mechanics and Maritime Sciences  
Division of Vehicle Safety  
Chalmers University of Technology

## ABSTRACT

In 2016, pedestrians accounted for the largest share within fatalities of vulnerable road users in traffic inside the European Union with a 21% fatality rate among all road users [1]. Active safety systems aim at avoiding crashes with pedestrians, however, there is a need to develop systems which are accepted by drivers, i.e. systems which are able to avoid collisions while keeping false alarm rate as low as possible. There has been a substantial amount of previous research with focus on driver interaction with pedestrians in crossing situations, but longitudinal scenarios have not gained the same amount of attention in research yet. This thesis aims at exploring the interaction between drivers and pedestrians in longitudinal scenarios, i.e. when drivers overtake or pass a pedestrian.

To allow an analysis of driver behaviour, two data sets of real world overtaking and passing manoeuvres were used. The first data set was extracted from the latest European naturalistic driving study, UDRIVE, and contained 77 overtaking events which occurred on rural roads in France. The second data set was collected within this thesis on a rural road in Sweden with a custom-made data logger. This second data set included 630 overtaking events which were collected by a pedestrian wearing the data logger. During the field test, experimental factors such as presence of oncoming traffic, walking direction of the pedestrian and lateral position of the pedestrian were varied.

Results from the UDRIVE data set indicate that, in presence of oncoming traffic, drivers start to steer away earlier from a collision path than when oncoming traffic is absent. From field test data, a significant difference in minimum clearance between driver and pedestrian is shown for the factors oncoming traffic and walking direction. The implication is that drivers are willing to give less space to pedestrians in presence of oncoming traffic as well as when pedestrians are walking towards them. Almost 50% of the drivers in the field tests conducted in Sweden and about 90% of the french drivers from the UDRIVE data set kept less than 1.5 m distance to the pedestrian, less than the minimum distance set by policies in other European countries [2]. Hence, infrastructure especially in rural roads should be designed to allow safe and comfortable collision avoidance among road-users, especially in combination with pedestrians. Results for distribution of time to collision when steering away go in accordance to the corresponding Euro NCAP scenario. Bayesian linear regression was applied to model the event describing metrics minimum clearance and time to collision at the moment of steering away. Results show that the models are able to predict posterior distributions of those metrics as well as estimates and uncertainty of changes under the influence of certain factor combinations.

Keywords: Vehicle safety, vulnerable road users, pedestrians, comfort zone boundaries, driver behaviour modelling, Bayesian regression model



## PREFACE

This master thesis was carried out within the European project “Drivers in Interaction with Vulnerable road users” (DIV) <sup>1</sup> which is led by Chalmers University of Technology, Autoliv AB, Veoneer AB and Toyota Motor Europe. The project aims at analysing driver behaviour in interaction with pedestrians and cyclists in intersection and overtaking scenarios with the aim to improve driver acceptance of active safety systems and evaluate new car assessment programmes. The work of this thesis was part of the work package 5 (DIV WP5), in which driver behaviour was to be analysed in the scenario of a driver overtaking a pedestrian.

## ACKNOWLEDGEMENTS

I would like to thank my colleague Gabriele Panero and my supervisor Christian-Nils Åkerberg Boda for a successful and fun cooperation and guidance in the DIV WP5 project. Furthermore, I thank my examiner at Chalmers, Marco Dozza, and my external supervisor at Autoliv AB, Nils Lübbe, for useful inputs and for having enabled all organisational parts of this thesis. Furthermore acknowledged are the other members of the DIV project for interesting discussions during the DIV meetings, Alessia Knauss, Tobias Aderum and Prateek Thalya from Veoneer AB as well as Tjark Kreuzinger and Pablo Puente Guillen from Toyota Motor Europe. Rustem Elezovic from Veoneer AB is acknowledged for all help during the field tests.

I would further like to thank the other members of the road-user vehicle interaction (RUVI) group at Chalmers for useful help and inspiration. Thank you to Jordanka Kovaceva, Jonas Bårgman, Erik Svanberg and Helena Gellerman for all help with UDRIVE. Thank you to András Bálint for the support in using the STRADA accident database and to Esko Lehtonen for help with the hardware for the field test data collection. I also want to thank Alberto Morando and Kevin Flannagan for useful inputs and help regarding Bayesian regression models. Oliver Lee is acknowledged for help with Python and the clustering of point clouds.

---

<sup>1</sup><http://divproject.eu/>



# Contents

<b>Abstract</b>	<b>i</b>
<b>Preface</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>iii</b>
<b>List of Figures</b>	<b>vii</b>
<b>List of Tables</b>	<b>xi</b>
<b>Nomenclature</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Objectives . . . . .	1
1.2 Hypotheses . . . . .	2
1.3 Outline . . . . .	2
<b>2 Background</b>	<b>5</b>
2.1 Pedestrian accident statistics . . . . .	5
2.1.1 World wide statistics . . . . .	5
2.1.2 The European perspective . . . . .	7
2.1.3 Sweden . . . . .	8
2.2 Road-user vehicle interaction . . . . .	10
2.2.1 Driver models . . . . .	10
2.2.2 Field of safe travel . . . . .	11
2.2.3 Comfort zone boundaries . . . . .	12
2.2.4 Overtaking manoeuvres . . . . .	14
2.2.5 Driver interaction with vulnerable road-users . . . . .	15
2.3 Intelligent safety systems . . . . .	16
2.3.1 Threat assessment . . . . .	16
2.3.2 System evaluation and assessment . . . . .	17
2.4 Naturalistic driving studies . . . . .	18
2.4.1 A brief overview . . . . .	18
2.4.2 UDRIVE . . . . .	19
<b>3 Methodology</b>	<b>21</b>
3.1 Overtaking manoeuvre terminology . . . . .	21
3.1.1 Definitions of phases . . . . .	21
3.1.2 Definitions of factors . . . . .	23
3.2 UDRIVE data . . . . .	25
3.2.1 Overtaking event extraction . . . . .	25
3.2.2 Extracting overtaking candidate segments . . . . .	27
3.2.3 Manual annotations . . . . .	29

3.2.4	Vehicle and pedestrian path reconstruction . . . . .	30
3.3	Pedestrian data logger . . . . .	32
3.3.1	Hardware development . . . . .	33
3.3.2	Logger software . . . . .	35
3.3.3	Data collection . . . . .	38
3.3.4	Data preparation and analysis . . . . .	40
3.4	Bayesian modelling of overtaking metrics . . . . .	46
3.4.1	Bayesian statistics . . . . .	46
3.4.2	Bayesian regression model . . . . .	47
3.4.3	Implementation in R . . . . .	48
<b>4</b>	<b>Results</b>	<b>51</b>
4.1	Results from UDRIVE data . . . . .	51
4.1.1	Qualitative observations . . . . .	51
4.1.2	Temporal results . . . . .	51
4.1.3	Spatial results . . . . .	53
4.1.4	Spatio-temporal results . . . . .	55
4.2	Results from field test data . . . . .	58
4.2.1	Spatial results . . . . .	58
4.2.2	Spatio-temporal results . . . . .	62
4.3	Driver modelling with Bayesian regression . . . . .	65
4.3.1	Minimum clearance model . . . . .	66
4.3.2	Time to collision model . . . . .	69
4.3.3	Speed model . . . . .	70
<b>5</b>	<b>Discussion</b>	<b>73</b>
5.1	Hypotheses review . . . . .	73
5.2	Comparison between UDRIVE and field test data results . . . . .	74
5.3	Comparison to existing studies . . . . .	75
5.3.1	Driver interaction with cyclists . . . . .	75
5.3.2	Driver interaction with pedestrians . . . . .	76
5.4	Impact on ADAS design and evaluation . . . . .	76
<b>6</b>	<b>Conclusion</b>	<b>77</b>
<b>7</b>	<b>Limitations and further work</b>	<b>79</b>
	<b>References</b>	<b>81</b>
<b>A</b>	<b>Pedestrian data logger</b>	<b>85</b>
A.1	Software . . . . .	85
A.2	Field test data overview . . . . .	87
<b>B</b>	<b>Bayesian regression models</b>	<b>89</b>
B.1	Minimum clearance model . . . . .	89
B.2	Time to collision model . . . . .	93
B.3	Vehicle speed model . . . . .	95

# List of Figures

1.1	Overview of this thesis, part I includes the data acquisition and is described in Methodology, part II and III are described in the Results and Conclusions part . . . . .	3
2.1	Road traffic deaths by type of road user by WHO region, image taken from [13]	5
2.2	Pedestrian fatalities number and rate in the USA, image adopted from [14] . .	6
2.3	Target pre-crash scenarios listed by NHTSA in USA, image adopted from [14]	7
2.4	Fatality distribution among different types of road-users for 2016 in EU countries, image taken from [1] . . . . .	8
2.5	Annual pedestrian fatality numbers extracted from STRADA and annual pedestrian injuries happening on right side of the road with a motorised vehicle involved . . . . .	9
2.6	Annual pedestrian fatality numbers extracted from STRADA and annual pedestrian injuries happening on right side of the road with a motorised vehicle involved . . . . .	10
2.7	FoST concept in a sample intersection with three approaching vehicles, green areas represent FoST, arrows indicate DVE trajectories . . . . .	11
2.8	Visualisation of safety and comfort zone approach, image taken from [20] . .	12
2.9	DVE trajectory in different pre-crash scenarios of a rear end collision, image taken from [20] . . . . .	13
2.10	Concept of overtaking phases, here illustrated with a vehicle overtaking a cyclist, phase definitions adopted from [6, 11] . . . . .	15
2.11	Concept visualisation of comfort zone boundaries and safety zone boundaries in a longitudinal vehicle-pedestrian scenario with oncoming traffic . . . . .	16
2.12	Euro NCAP scenario “Car-to-Pedestrian Longitudinal Adult” (CPLA) for different offsets, the 1 m distance is the acceleration distance of the pedestrian, the impact happens after the dummy pedestrian has walked another 10 m, images recreated from [37] . . . . .	18
2.13	UDRIVE naturalistic driving study, images taken from (a) [42] and (b) [43] . .	19
3.1	Start of phase 1 of the overtaking manoeuvre, driver recognises pedestrian, symbolised by red circles . . . . .	21
3.2	Start of phase 2 of the overtaking manoeuvre . . . . .	22
3.3	Start of phase 3 of the overtaking manoeuvre . . . . .	22
3.4	Start of phase 4 of the overtaking manoeuvre . . . . .	22
3.5	Start of phase 5 of the overtaking manoeuvre . . . . .	23
3.6	Overtaking vehicle (white) facing oncoming vehicle (red), the factor oncoming traffic is set to present if the oncoming vehicle is anywhere between 120 m away when the overtaking vehicle is starting to approach and 20 m away when the overtaking vehicle has returned . . . . .	24
3.7	Data structure in SALSA . . . . .	25
3.8	Interaction between global and local database in SALSA . . . . .	26
3.9	Pedestrian disappears from ME detection field . . . . .	27
3.10	Map of the 77 pedestrian overtaking events in France. The events mainly occurred in the south-eastern region. . . . .	28

3.11	Pie chart of driver ID distribution, each colour represents one driver ID. The three most frequently occurring drivers account for almost 50% of the events.	28
3.12	Visualisation of used point mass model for vehicle path extrapolation . . . . .	30
3.13	Visualisation of RANSAC straight line fit for pedestrian position extrapolation	32
3.14	Wiring sketch of the pedestrian data logger . . . . .	33
3.15	Photos of the 3D printed physical components of the pedestrian data logger .	35
3.16	Pedestrian data logger state machine . . . . .	36
3.17	Screen shot from <i>rviz</i> visualisation tool in ROS, showing the camera image (lower left corner) and the frame tree (right side) . . . . .	37
3.18	Screen shots of the PDL control web app . . . . .	38
3.19	Data collection site at Tuvevägen, GPS coordinates point A (57.7682, 11.9357) and point B (57.7594, 11.9390) . . . . .	39
3.20	Two lateral pedestrian positions tested, in case line the pedestrian is walking and curb . . . . .	39
3.21	Photos from data collection on 13th March, 2018 . . . . .	40
3.22	Flow chart of the data analysis process, yellow background indicates use of IMU data, blue background use of LiDAR data, red background symbolises the final output of the analysis . . . . .	41
3.23	Normalised histograms for IMU orientation in Euler angles, retrieved from Madgwick filter . . . . .	42
3.24	LiDAR point clusters corresponding to detections of an overtaking car at different phases of the overtaking event, recorded by the LiDAR mounted to the hip of the pedestrian. Approaching phase (1) produces a rather vertical line, while the steer away phase (2) causes an L-shape cluster. When passing (3), the cluster becomes a rather horizontal line, when steering back (4) an L-shape appears. At the end of the overtaking manoeuvre (5), the cluster appears as a vertical line . . . . .	44
3.25	Detected points in cluster when L-shape is present (red), blue points mark the extreme points of the cluster used for size and velocity estimation. The extremity points are the points with minimum/maximum x- and y-coordinate as well as the intersection point of the perpendicular line fits for front and side.	45
3.26	RANSAC line estimation for front (blue) and rear (red) detections for two different vehicles, inliers are marked in green . . . . .	46
3.27	Kruschke diagram, adopted from [54], of used BRM for MC . . . . .	49
4.1	Box plot diagrams showing the phases duration, outliers marked with red plus symbols, mean marked with white plus . . . . .	52
4.2	Empirical cumulative density function of time to collision for factor oncoming traffic and pedestrian direction . . . . .	53
4.3	Box plot diagrams showing the MAG and MC distributions depending on the factors oncoming traffic and pedestrian direction, outliers marked with red plus symbols, mean marked with white plus . . . . .	54
4.4	Empirical cumulative density function of minimum clearance for factor oncoming traffic and pedestrian direction, the red dashed line marks 1.5 m distance which is included in most of the European countries' policies as a minimum distance [2] . . . . .	55



4.5	Box plot diagrams showing vehicle speed and longitudinal acceleration when being next to the pedestrian, including interaction of factors oncoming traffic and pedestrian direction . . . . .	56
4.6	Box plot diagrams showing the speed change during the four different overtaking phases, for accelerative and flying manoeuvres . . . . .	57
4.7	Scatter plots of time to collision (only flying overtaking manoeuvres) and minimum clearance over speed for accelerative and flying overtaking strategy . . . . .	58
4.8	Box plot diagrams for minimum clearance distribution in field test data, line and curb scenario in dependence of factors oncoming traffic and pedestrian direction. Means are marked with white plus symbols, outliers marked with red plus symbols. . . . .	59
4.9	Empirical cumulative density function of minimum clearance under influence of the interaction between the factors oncoming traffic and pedestrian direction, red dashed line symbolises minimum distance proposed by other European countries . . . . .	60
4.10	Box plot diagrams for the lateral position of the vehicle centre for walking on the line scenario, interaction between parameters oncoming traffic and pedestrian direction . . . . .	61
4.11	Box plot diagrams for the lateral position of the vehicle centre for walking on the line scenario, interaction between parameters oncoming traffic and pedestrian direction . . . . .	62
4.12	Speed of overtaking vehicles in walking on the line scenario with factors oncoming traffic and pedestrian direction, dashed line: speed limit . . . . .	63
4.13	Empirical cumulative density of speed under influence of interaction between factors oncoming traffic and pedestrian direction . . . . .	63
4.14	Minimum clearance over vehicle speed in walking on the line and curb scenario with factors oncoming traffic, vertical dashed line: speed limit, horizontal dashed line, 1.5 m limit set by other EU countries . . . . .	64
4.15	Empirical cumulative density of speed under influence of interaction between factors oncoming traffic and pedestrian direction . . . . .	65
4.16	Empirical cumulative density of speed under influence of interaction between factors oncoming traffic and pedestrian direction . . . . .	65
4.17	Empirical cumulative distribution functions of posterior predictive checks from UDRIVE data using 100 samples from the fitted model (blue lines) for minimum clearance from UDRIVE data, red line marks the measured data . . . . .	66
4.18	Model estimates of minimum clearance from UDRIVE data under different levels of the interaction between factors oncoming traffic, pedestrian direction and lateral position for field test data, fitted estimates are marked with black coloured bars. Fitted distributions are plotted with mean (circle), 66 % (thick line) and 95 % estimates (thin line). Predicted distributions are plotted with 50 %, 80 % and 95 % intervals. The red dashed line marks 1.5 m, the minimum distance set by law in other European countries. . . . .	67
4.19	Empirical cumulative distribution functions of posterior predictive checks using 100 samples from the fitted model (blue lines) for minimum clearance from field test data, red line marks the measured data . . . . .	68

4.20	Model estimates of minimum clearance from field test data under different levels of the interaction between factors oncoming traffic, pedestrian direction and lateral position for field test data, fitted estimates are marked with black coloured bars, predicted with blue colour scale . . . . .	68
4.21	Empirical cumulative distribution functions of posterior predictive checks using 100 samples from the fitted model (blue lines) for minimum clearance from field test data, red line marks the measured data . . . . .	69
4.22	Model estimates of time to collision under different levels of the interaction between factors oncoming traffic, pedestrian direction and lateral position for UDRIVE data, fitted estimates are marked with black coloured bars, predicted with blue colour scale, red dashed line marks the Euro NCAP 1.7 s TTC minimum limit for FCW trigger . . . . .	70
4.23	Empirical cumulative distribution functions of posterior predictive checks using 100 samples from the fitted model (blue lines) for vehicle speed from field test data, red line marks the measured data . . . . .	71
4.24	Model estimates of vehicle speed from field test data under different levels of the interaction between factors oncoming traffic, pedestrian direction and lateral position for field test data, fitted estimates are marked with black coloured bars, predicted with blue colour scale, red dashed line marks speed limit (70 km/h) . . . . .	71
A.1	Frame tree of pedestrian data logger, for each frame, the broadcasting node, the update rate, the time of the latest update and the buffer length is shown .	85
A.2	ROS nodes and topics overview of <i>div_datalogger</i> package . . . . .	86
B.1	MC model fit from field test data, including parameter posterior distributions (left) and MCMC chains (right), part 1 . . . . .	89
B.2	MC model fit from field test data, part 2 . . . . .	90
B.3	MC model fit from UDRIVE data, part 1 . . . . .	91
B.4	MC model fit from UDRIVE data, part 2 . . . . .	92
B.5	TTC model fit from UDRIVE data, part 1 . . . . .	93
B.6	TTC model fit from UDRIVE data, part 2 . . . . .	94
B.7	Speed model fit from field test data, part 1 . . . . .	95
B.8	Speed model fit from field test data, part 2 . . . . .	96

# List of Tables

3.1	Overview of the terminology used to describe factors which potentially influence an overtaking manoeuvre . . . . .	23
3.2	Overview of manually annotated attributes for identified overtaking manoeuvres in UDRIVE . . . . .	29
3.3	Overview of sensors connected to main computer . . . . .	34
3.4	Overview of the recorded topics by the PDL . . . . .	36
4.1	Overtaking phases duration for accelerating and flying manoeuvres, mean and standard deviation (std) . . . . .	52
4.2	Time to collision at steer away moment for accelerating and flying manoeuvres, number of observations (N), mean and standard deviation (std) . . . . .	53
4.3	Minimum clearance from UDRIVE events for accelerating and flying manoeuvres, number of observations (N), mean and standard deviation (std) . . . . .	55
A.1	Minimum clearance from field test overtaking manoeuvres for line scenario, including number of observations (N), mean and standard deviation (std) . .	87
A.2	Minimum clearance from field test overtaking manoeuvres for curb scenario, including number of observations (N), mean and standard deviation (std) . .	87



# Nomenclature

## Abbreviations

ADAS	Advanced Driver Assistance System
AEB	Automated Emergency Braking
AHRS	Attitude and Heading Reference System
CA	Collision Avoidance
CAN	Controller Area Network
CEESAR	Centre Européen d'Etude de Sécurité et d'Analyse des Risques
CI	Credible Interval
CPLA	Car-to-Pedestrian Longitudinal Adult
CZB	Comfort Zone Boundary
DAS	Data Acquisition System
DVE	Driver Vehicle Environment
ECDF	Empirical Cumulative Density Function
Euro NCAP	European New Car Assessment Program
FARS	Fatality Analysis Reporting System
FCW	Forward Collision Warning
FoST	Field of Safe Travel
FOT	Field Operational Test
GPIO	General Purpose Input/Output
GPS	Global Positioning System
I2C	Inter-Integrated Circuit
IMU	Inertial Measurement Unit
ISS	Intelligent Safety System
LiDAR	Light Detection and Ranging
MAG	Minimum approaching gap
MC	Minimum clearance between driver and pedestrian
ME	MobilEye
MIT-AVT	Massachusetts Institute of Technology - Advanced Vehicle Technology Consortium
NaN	Not a Number

NCAP	New Car Assessment Program
NDS	Naturalistic Driving Study
NHTSA	National Highway Traffic Safety Administration
PCAM	Pedestrian Crash Avoidance/Mitigation
PCL	Point Cloud Library
PCW	Pedestrian Collision Warning
PDL	Pedestrian data logger
PLA	Polyactic Acid
RANSAC	Random Sample Consensus
ROS	Robot Operating System
RPi	Raspberry Pi
RTC	Real Time Clock
SALSA	Smart Application for Large Scale Analysis
SPI	Serial Peripheral Interface
STRADA	Swedish Traffic Accident Data Acquisition
THW	Time headway
TTC	Time to collision
UDRIVE	European naturalistic Driving and Riding for Infrastructure & Vehicle safety and Environment
USA	United States of America
VEB	Vacuum Emergency Brake
VRU	Vulnerable Road User

### **Bayesian regression specific variables**

$\beta$	Fixed effect parameters
$\eta$	Predictor of model
$\mathbf{X}$	Fixed effect matrix
$\mathbf{y}$	Vector of model responses
$\mathbf{Z}$	Random effect matrix
$\sigma$	Standard deviation
$\theta$	Parameter of statistical distribution
$f$	Link function for predictor
$N$	Number of responses
$p$	Number of fixed effects
$q$	Number of random effects
$u$	Random effect parameters

$y_i$  i-th response of model

### Vehicle specific variables

$\dot{i}_{in,tmp}$	RANSAC indices of inliers (temporary)
$\dot{i}_{in}$	RANSAC indices of inliers
$p_{inter}$	Intersection point in LiDAR cluster
$p_{max-x}$	Point in LiDAR cluster with maximum x-coordinate
$p_{max-y}$	Point in LiDAR cluster with maximum y-coordinate
$p_{min-x}$	Point in LiDAR cluster with minimum x-coordinate
$p_{min-y}$	Point in LiDAR cluster with minimum y-coordinate
$p_{veh}$	2D position of the vehicle centre of gravity
$\hat{v}_{ped}^{GPS}$	Mean pedestrian speed from GPS velocity
$\mathcal{P}$	Set of points, point cloud
$\nu$	Number of freedoms in student t distribution
$\omega_{z,gyro}$	Gyroscope yaw rate
$\psi_{veh}$	Vehicle heading angle
$\tau$	RANSAC distance threshold
$b, m$	RANSAC line intercept, slope
$b_{tmp}, m_{tmp}$	RANSAC line intercept, slope (temporary)
$i_{iter}$	RANSAC iteration index
$N_{iter}$	Number of RANSAC iterations
$p_{in}$	RANSAC inlier ratio
$t_{i \rightarrow j}$	Time of change between phase $i$ and phase $j$
$v_{rel}$	Relative speed between vehicle and pedestrian
$v_{veh}$	Absolute speed of the vehicle centre of gravity
$W_{off}$	Lateral position offset between line and curb scenario
$W_{veh}$	Width of the vehicle
$x_{veh}, y_{veh}$	x-, y-position of the vehicle





# 1

## Introduction

European roads are considered the safest in the world where in 2016, 50 inhabitants out of one million died in traffic, while the global fatality rate was around 174 per one million. However, even though fatalities were reduced by 43% between 2001 and 2010, still over 25 000 people were killed in traffic in 2016 on roads in the European Union (EU). [1]

The term vulnerable road users (VRUs) comprises pedestrians, cyclists and motorcyclists who share roughly the same amount of traffic fatalities compared to car occupants. Pedestrians account for the largest share within VRU fatalities with an overall fatality rate of 21% among all road users. Most of the fatalities in the EU occur on rural roads (55%) where impact speeds are usually higher and sufficient infrastructure like sidewalks is often missing. [1]

Collision avoidance (CA) systems belong to the key features of modern advanced driver assistance systems (ADAS) or intelligent safety systems (ISS) to reduce fatalities caused by vehicles. CA technology can warn the driver when a collision with a VRU is imminent (forward collision warning, FCW) or intervene for example via automated emergency braking (AEB). [3, 4]

Recent research has adapted the concept of comfort zone boundaries (CZBs) to quantify driver comfort in overtaking and intersection scenarios with VRUs [5–8]. The concept of CZBs stems from the field of safe travel framework described by Gibson and Crooke in [9] which was adapted later on by Ljung-Aust and Engström in [10].

### 1.1 Objectives

Overtaking of cyclists has been objective to several previous studies using both naturalistic driving studies (NDS) and experiments [6, 11]. The used methodology from those studies is investigated in this thesis for application on overtaking or passing manoeuvres of pedestrians. This study aims at providing two different kinds of data to analyse overtaking manoeuvres from both driver as well as pedestrian perspective. Data acquisition includes a semi automated extraction process for the large scale European NDS called UDRIVE [12]. Field test data is recorded in a supervised walking scenario on a straight rural road with a pedestrian who is carrying a developed data logger. In the end, a driver behaviour model is the result of a comprehensive study of NDS and field test data.

The driver model shall contain important aspects of the overtaking manoeuvre such as spatial and temporal distance between vehicle and pedestrian as well as information about the trajectory. Such knowledge can be used as input to CA systems or new car assessment to mitigate fatalities including pedestrians in the future. Understanding and modelling the interaction between drivers and pedestrians in overtaking or passing scenarios is therefore the main research question of interest addressed in this thesis.

## 1.2 Hypotheses

To guide the analysis of the interaction between drivers and pedestrians, five main hypotheses are stated which are discussed throughout the thesis. The hypotheses concern the minimum distance between driver and pedestrian, referred to as minimum clearance (MC), as well as speed and steering behaviour of the driver.

1. When the pedestrian walks closer to the middle of the lane, minimum clearance decreases
2. When drivers and pedestrians face each other, minimum clearance decreases
3. When the pedestrian walks slightly out of the lane, meaning next to the lane marking or on the curb, drivers do not steer away
4. When drivers face oncoming traffic, minimum clearance decreases
5. When drivers face oncoming traffic, speed is reduced

## 1.3 Outline

The thesis is divided into three major parts, visualised in Figure 1.1.

- Data acquisition in form of two data sets, obtained from NDS and field tests on the road (I), described in the Methodology chapter
- CZB estimation from the two data sets (II), described in the Methodology and Results chapter
- CZB evaluation (III), including driver model creation from CZB metrics, described in the Results chapter and a review of the European new car assessment program (Euro NCAP) scenario for pedestrian AEB systems

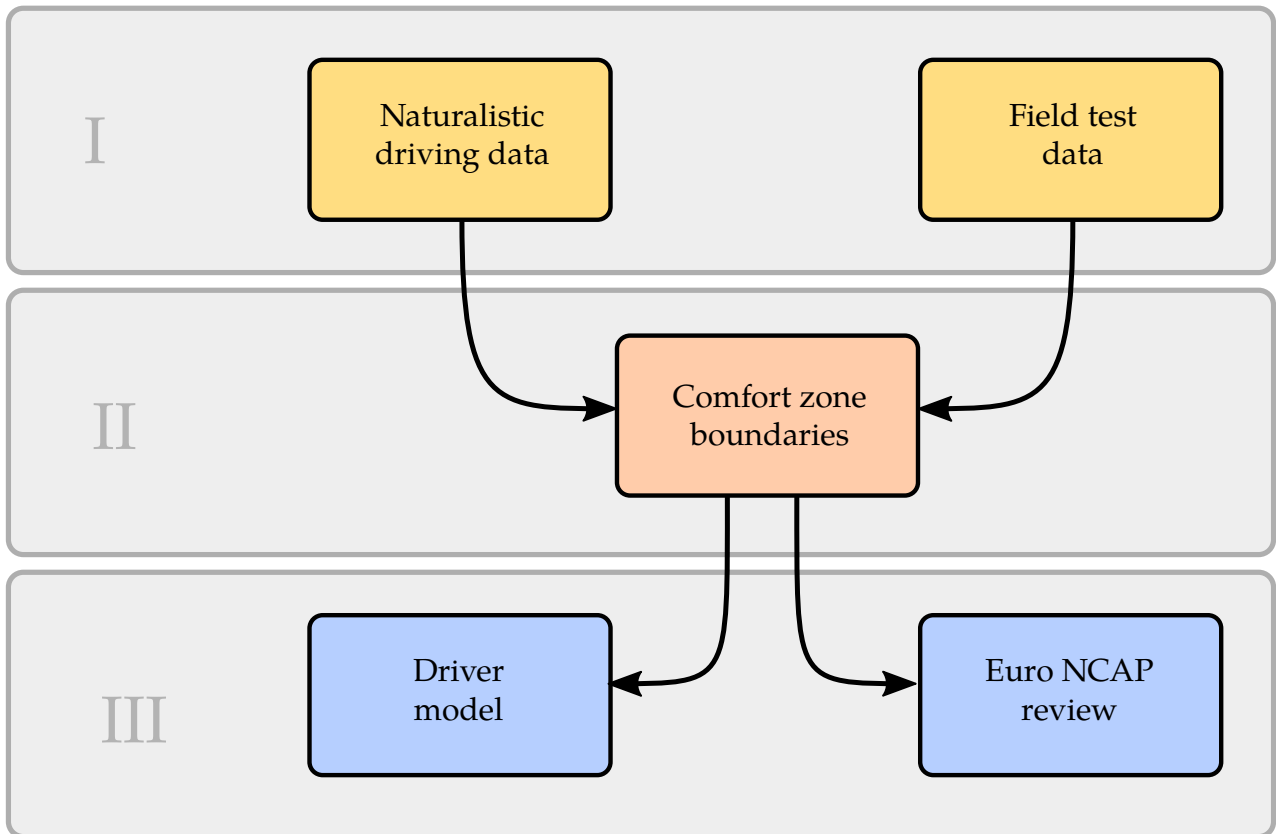


Figure 1.1: Overview of this thesis, part I includes the data acquisition and is described in Methodology, part II and III are described in the Results and Conclusions part



# 2

## Background

### 2.1 Pedestrian accident statistics

#### 2.1.1 World wide statistics

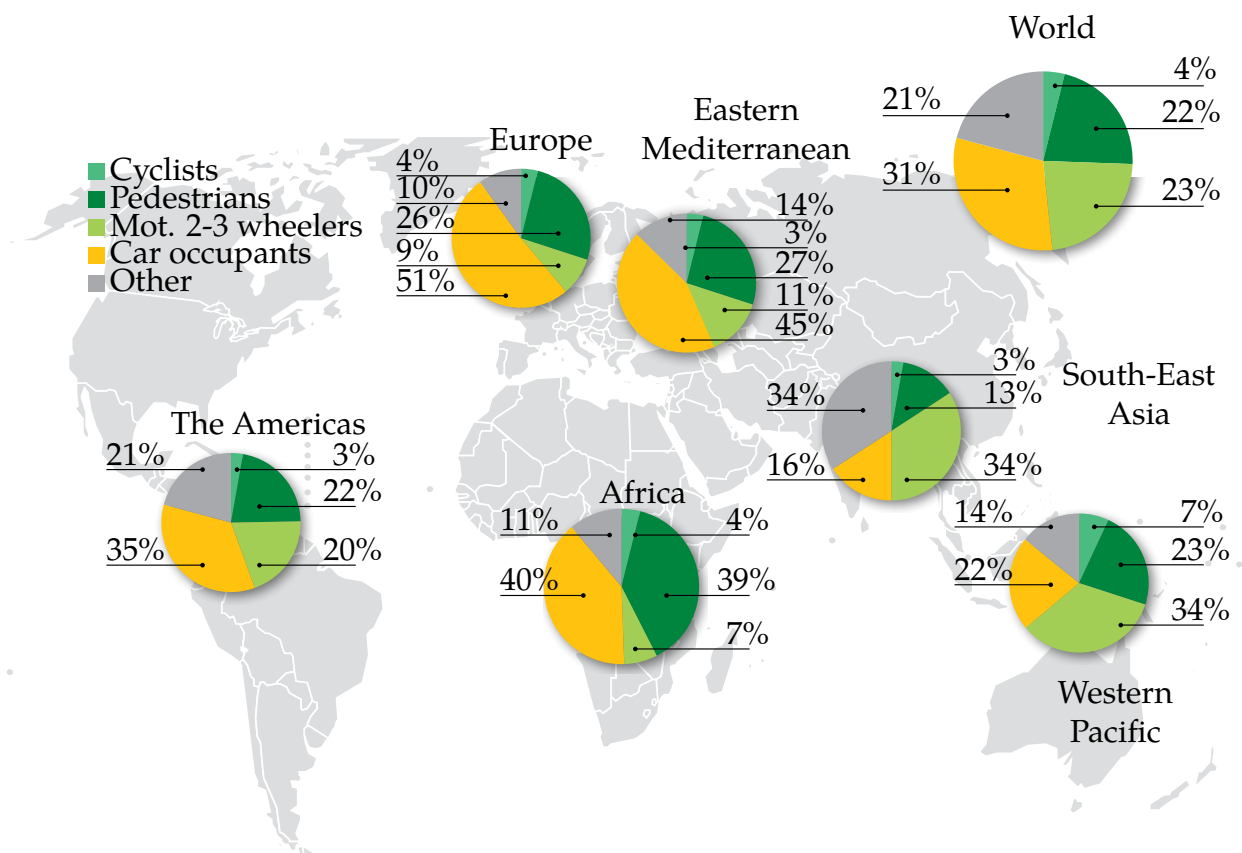


Figure 2.1: Road traffic deaths by type of road user by WHO region, image taken from [13]

The World Health Organisation (WHO) reports in their Global Status Report On Road Safety 2015 that almost half of all traffic fatalities involve VRUs. Pedestrians account for a share of 22% of all traffic fatalities yearly with 275,000 deaths a year. Figure 2.1 shows the global distributions of road deaths by type of transport. In the African region, the proportion of fatalities among pedestrians and cyclists is the worldwide highest with 43% of all road fatalities, whereas fatality rates for VRUs tend to be lower in the South-East Asia region. The study further describes the significant impact of vehicle speed on fatal accident outcome, where adult pedestrians have a 20% risk of dying if hit by a car at 50 km/h and 60% risk

if struck at 80 km/h. For urban areas speeds above 50 km/h are described as unsafe due to a high concentration of VRUs. The study claims that in order to improve road safety, infrastructure is a key mechanism and 91 countries have already implemented policies separating VRUs from traffic at higher speed. [13]

In the United States of America (USA), the National Highway Traffic Safety Administration (NHTSA) reports an increasing trend in the share of pedestrian fatalities over all fatalities within the last 10 years. While in 2010, 4290 pedestrians died (13% share) on roads in the USA, the number increased to 5987 (16% share) in 2016, the highest number since 1990. From 2015 to 2016 the percentage of pedestrian fatalities increased by 9%. In contrast, bicyclist deaths increased by 1.3 % to 840 fatalities, the highest number since 1991. Figure 2.2 shows an increasing trend for pedestrian fatalities from 2004 to 2013. In

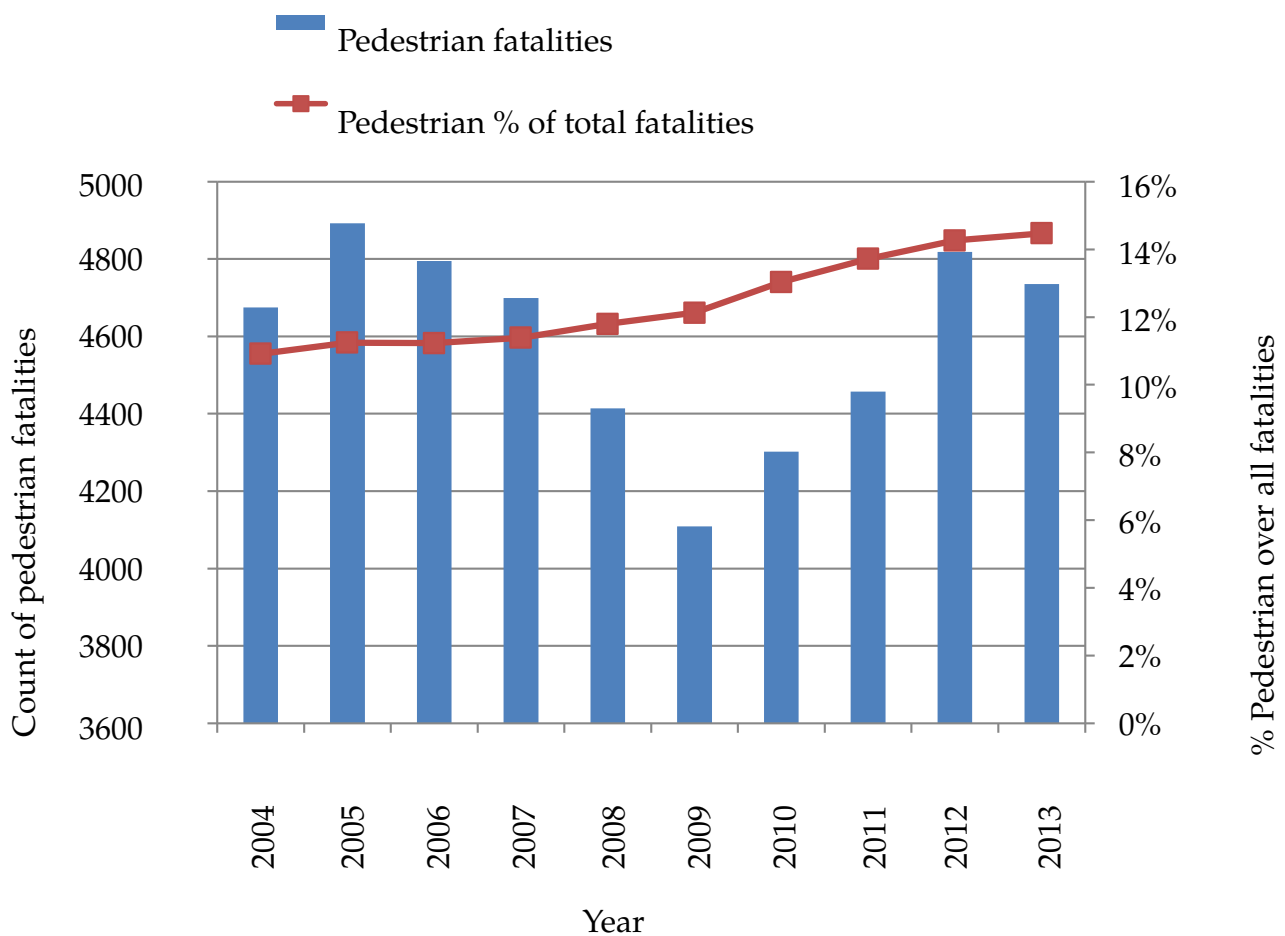


Figure 2.2: *Pedestrian fatalities number and rate in the USA, image adopted from [14]*

In Figure 2.3, the four target scenarios S1 to S4, listed by the NHTSA for pedestrian crash avoidance/mitigation (PCAM) systems are sketched where scenario S4 is marked with a red rectangle because S4 is the scenario which is subject of investigation in this thesis. Scenarios S1 to S4 account for 97% of all vehicle pedestrian crashes in the Fatality Analysis Reporting System (FARS), averaged over the years 2011 and 2012. Within the four scenarios, S1 and S4 are the most frequent ones, with a share of 64% and 28%, respectively. [14, 15]

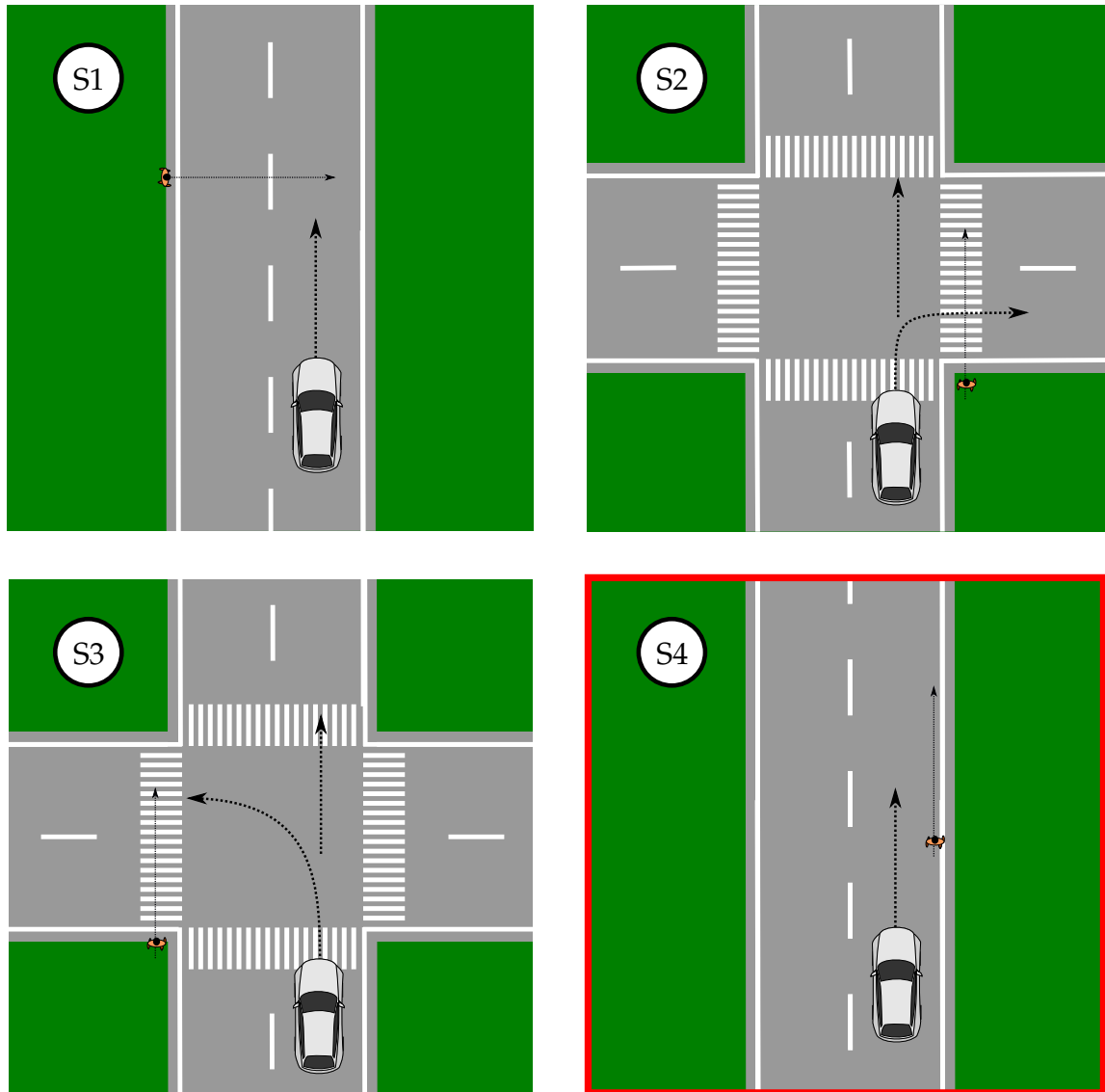


Figure 2.3: Target pre-crash scenarios listed by NHTSA in USA, image adopted from [14]

According to [16], most of the pedestrian fatalities occur in urban areas, however, on rural roads pedestrians are 2.3 times more likely to die in a traffic accident. In contrast to urban areas, vehicles drive at higher speeds on rural roads and separating infrastructures such as sidewalks, paths or trails are often missing. [16]

### 2.1.2 The European perspective

In the Traffic Safety Basic Facts 2016, published by the European Union (EU), the number of pedestrian deaths among all member states of the EU is estimated to 5621 in the year 2014, which translates to a percentage of 22% among all road fatalities, shown in Figure 2.4. Countries like Netherlands and Denmark show the lowest percentages with 11% and 12%, respectively. In contrast, fatality rates per million inhabitants are highest for Eastern EU states. In Lithuania, Romania, Latvia and Estonia more than 30% of road fatalities are pedestrians. The trend of the absolute number of pedestrians who died in traffic accidents in the EU decreased over the 2005-2014 decade, however, the share of pedestrian fatalities

within all road fatalities increased. The percentage of pedestrian fatalities peaks between 4 pm and midnight, as well as in winter. It is further stated that pedestrian fatality numbers per month are more variable than other road fatalities. [1]

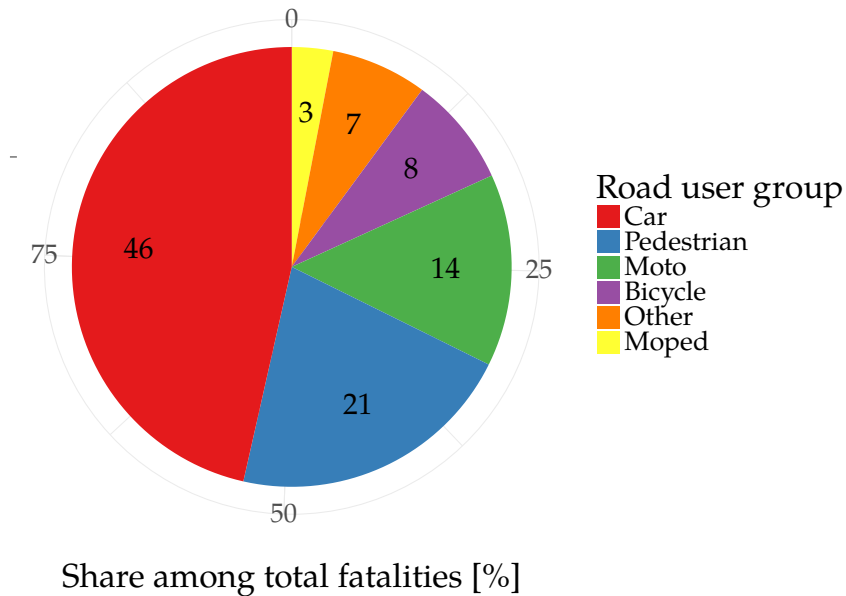


Figure 2.4: *Fatality distribution among different types of road-users for 2016 in EU countries, image taken from [1]*

The study also states significant differences among different demographic groups of pedestrians. Older pedestrians aged above 80 years are more than ten times more likely to die on roads than children. However, both age groups, elderly and children, account for the highest pedestrian fatalities. Furthermore, in almost all EU countries, there are more male pedestrians who died in traffic than female, in average two thirds of all pedestrian fatalities were male. [1]

The risk of fatality in darkness varies among countries where a maximum of 68% pedestrian fatality percentage is reported in Latvia while 34% in Bulgaria. A typical cause of pedestrian fatalities is driver distraction, resulting in wrong diagnosis, missed observations or inadequate planning. VRU's represent 63% of all road accident casualties and 56% of the casualties admitted to hospital. Typical injuries occur in lower extremities or involve the head. In accordance to the trends in other countries like the USA, pedestrian fatalities are more common in urban areas, 40%, than in rural areas, 11%. Crashes with VRU's in rural areas typically occur close to bus or tram stops where more frequent interactions between road-users are present. [1]

### 2.1.3 Sweden

The Swedish Transport Agency reports accident information in a database using an information system called "Swedish Traffic Accident Data Acquisition" (STRADA), which is used to manage reports about causation and injuries from both police and hospitals. Figure 2.5a shows the STRADA database police report results for pedestrian fatality numbers in Sweden over the decade from 2006 to 2016. It can be observed that pedestrian fatalities have not been



significantly reduced in Sweden over this time span. The ratio of pedestrian casualties over all traffic fatalities has even slightly been increased over this time span. The risk of death among all incidents where a VRU is involved into an accident, is on average almost three times higher for pedestrians than for cyclists. While for cyclists only 15% of all fatalities occurred during night time, 45% of all pedestrian fatalities happened during night time. In Figure 2.5b, pedestrian injury numbers are shown which happened on the right side of the road, from year 2003 to 2010, the scenario investigated in this thesis. It shows that most injuries happen in urban areas.

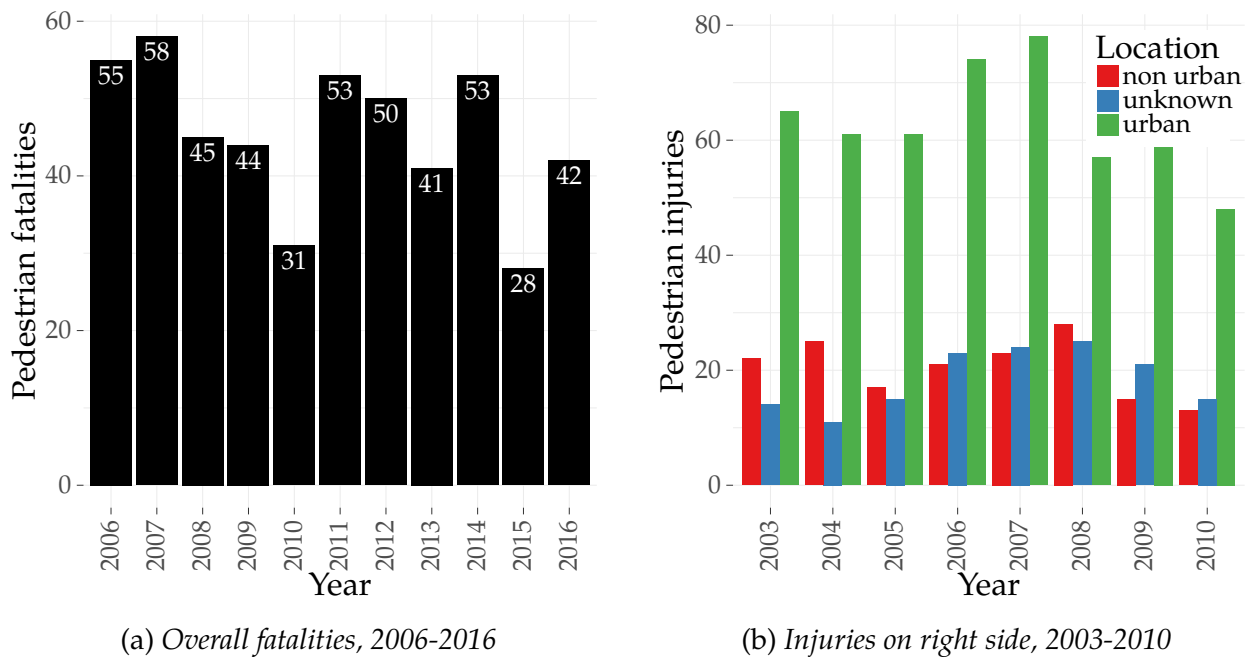
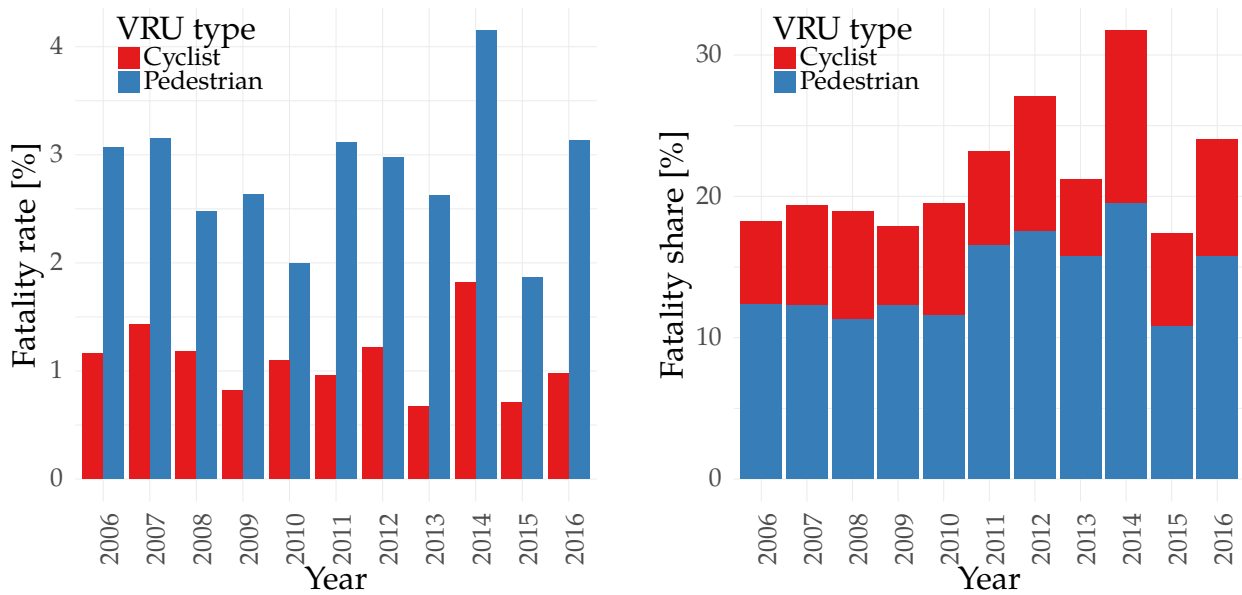


Figure 2.5: Annual pedestrian fatality numbers extracted from STRADA and annual pedestrian injuries happening on right side of the road with a motorised vehicle involved

In Figure 2.6, STRADA accident statistics are presented comparing different pedestrian and cyclist fatality rates. Figure 2.6a shows the ratio in percent of pedestrian and cyclist fatalities over the corresponding number involved in accidents with motorised traffic from 2006 to 2016. It can be seen that for pedestrians the likelihood of being killed in a traffic accident was on average nearly twice as much as for cyclists. Figure 2.6b shows the ratio of pedestrian and cyclist fatalities over the total number of traffic fatalities from 2006 to 2016. It can be observed that the share of cyclists and pedestrians has increased within all fatalities over the years. Furthermore, pedestrians account for a larger share on average than cyclists.



(a) Fatalities divided by involved in accident by VRU type, 2006-2016 (b) Fatalities divided by total fatalities by VRU type, 2006-2016

Figure 2.6: Annual pedestrian fatality numbers extracted from STRADA and annual pedestrian injuries happening on right side of the road with a motorised vehicle involved

The fatality rate for pedestrians in non-urban area at darkness was 19% on average, being 3% in urban area, which gives rise to a lack of pedestrian safety on rural roads. From 2006 to 2016, there has been a decreasing trend in pedestrians involved in accidents at darkness, being 534 in 2006 and 440 in 2016.

## 2.2 Road-user vehicle interaction

### 2.2.1 Driver models

While in passive safety, physical laws can guide the evaluation of safety systems, in active safety, system design and evaluation cannot solely rely on physical laws. Driver models aim at modelling the behaviour of drivers and have the potential to simulate driver behaviour in active safety system evaluation and enable a greater repeatability and economical benefit than test track experiments. [10, 17]

Driver models can be distinguished into two categories, descriptive and motivational models. In descriptive models, the aim is to describe the whole driving task or parts of it by which actions the driver has to take. Descriptive model can be further broken down into task models, adaptive control models and production models. While task models include sub-tasks which are hierarchically ordered, adaptive control models aim at modelling driving with environment adapting control systems which have inputs and outputs. A production model is a driver model in which the driving task is described with a set of formal rules, specifying how a driver should behave in a certain situation. [10, 18]

A limitation of descriptive driver models is the lack of the ability to predict driver behaviour and behavioural variation affected by driver motivation. Motivational models aim

at filling this gap by describing how drivers manage and control risk and task difficulty for example by making compromises between different decisions and outcomes. In this context, motivation is described as an influencing factor on the direction of behaviour, i.e. the goal or target selection, as well as the intensity and persistence of behaviour. However, motivational models have not been utilised or tested due to a lack of practicability. [10, 18]

### 2.2.2 Field of safe travel

One of the first approaches in modelling driver behaviour were made in 1938 by Gibson and Crooks, who describe the task of automobile driving as a type of locomotion through a field of space, similar to walking or running. In contrast to the latter, driving is described as travelling using a tool, the vehicle, which makes the locomotion more effective. The fundamental items in the driving task are the driver and the vehicle, together called the driver vehicle environment (DVE). The major contribution from the driver to the driving act is performed using visual perception in order to safely navigate the vehicle by means of a path, avoiding obstacles and reaching the destination. The responsibility of the driver within the DVE is further described as maintaining a field of safe travel (FoST). The FoST is the manoeuvring space in which the driver can safely bring the vehicle to a stop. [9]

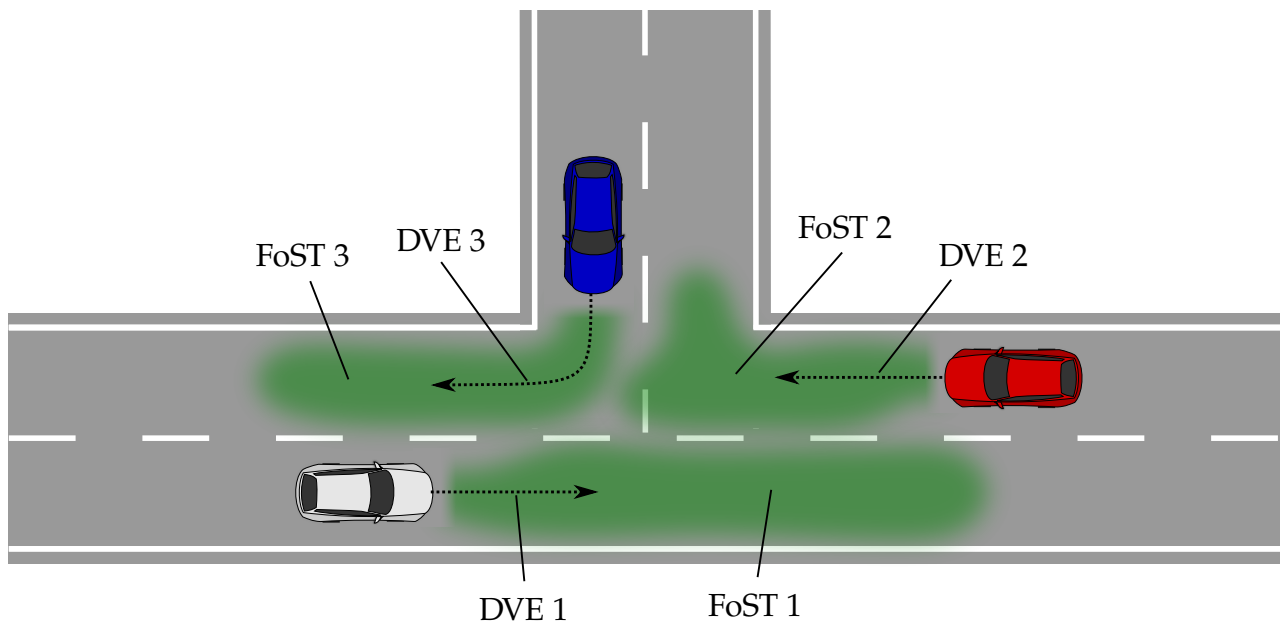


Figure 2.7: *FoST concept in a sample intersection with three approaching vehicles, green areas represent FoST, arrows indicate DVE trajectories*

The FoST idea is revised by Papakostopoulos et al. in [19], who distinguish between two perceptual entities and develops an extended FoST framework. Firstly, the spatial FoST for the vehicle, and secondly, the driver's mental perception of the ego vehicle dimensions and dynamics. The framework's basic concept is that the driver integrates both entities into the driving task to maintain the FoST. Using the framework, the driver's prioritisation between the two entities can be predicted in varying traffic situations. Driver perception of other

road users is included in the framework by means of the driver's ability to foresee their unpredictability in motion and hence keeping for example more space to a motor cycle or pedestrian than to another car. [19]

The FoST framework further describes the driving task as a holistic type of locomotion in which the driver's focus is not on separate sub tasks at a time, like lane keeping, but rather on a combination of visual-spatial perceptual processes. At all times during travel, the driver's construction of the FoST is product of the integration of all traffic elements. The construction depends on the quality of perceptual cues. According to the study, modern vehicles should therefore aim to assist the driver in constructing a more accurate FoST by giving visual cues such as displaying clearance lines between the ego-vehicle and other road-users. In the current context of automated driving, the interaction between driver and vehicle must therefore be improved with respect to transitions between assisted and automated driving as well interaction between manual and autonomous vehicles. [19]

### 2.2.3 Comfort zone boundaries

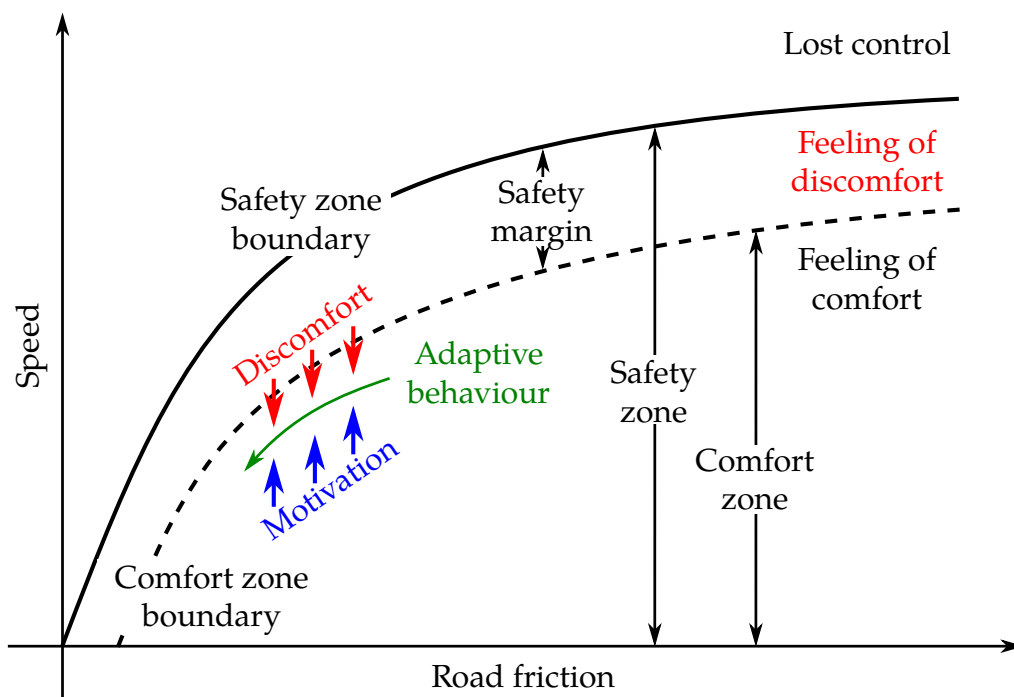


Figure 2.8: Visualisation of safety and comfort zone approach, image taken from [20]

Adapting to the FoST framework proposed in [9], Ljung-Aust and Engström propose a framework which divides the field of safe travel into safety and comfort zones. Figure 2.8 shows the key components of the concept. Safety zones represent the region of the DVE space in which the control of the vehicle is maintained at all times. The region outside of the DVE space represents states which result in crashes or other irreversible losses of control. The border which limits the safety zone is called safety zone boundary. Comfort zones are defined as regions in the DVE space in which drivers do not feel discomfort and therefore

prefer to reside in. The distinction between safety and comfort zones stems from research findings which state that drivers tend to avoid DVE states close to the safety zone boundary. The safety zone generally contains the comfort zone. The border which limits the two zones is referred to as comfort zone boundary, CZB. [17, 20]

Reasons for exceeding the safety zone can be found in adaption failures of the DVE system to a specific situation. For instance erroneous perception of the safety zone boundary can lead to such failures or the overestimation of the physical capabilities of the DVE. Wrong prediction of situation outcome over time and sudden unexpected events can be further causes which lead the DVE to exceed the safety zone boundary. [17]

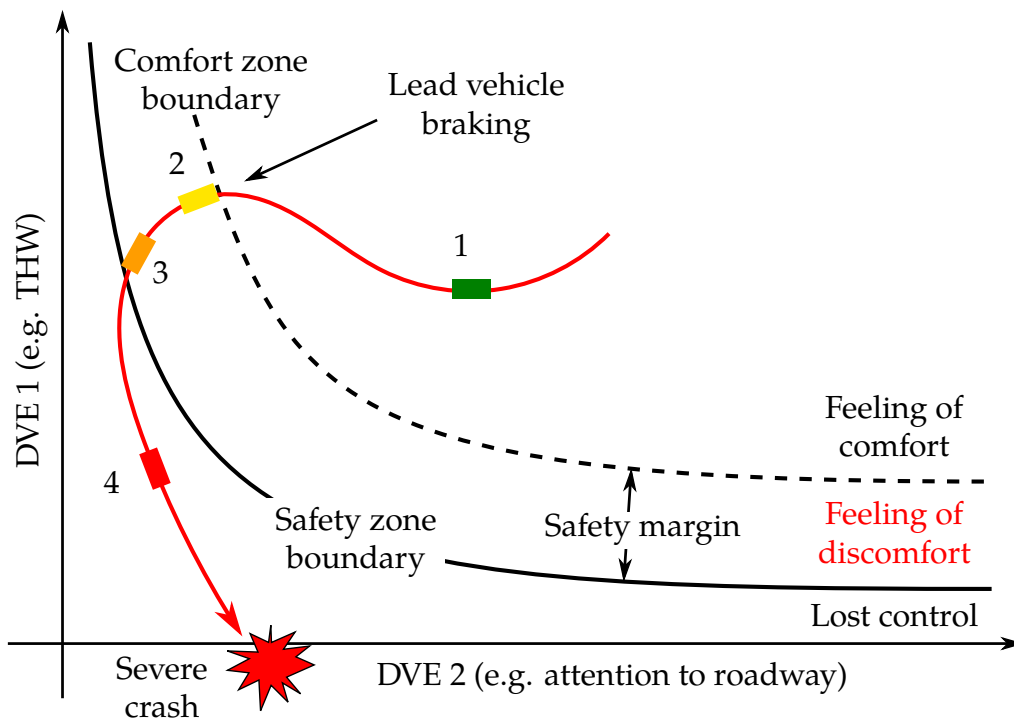


Figure 2.9: DVE trajectory in different pre-crash scenarios of a rear end collision, image taken from [20]

In Figure 2.9, pre-crash scenarios are shown which lead to a severe crash. In the initial situation (1), the DVE resides within the comfort zone. Due to lower attention to the roadway and the lead vehicle which starts to brake, the DVE is moved out of the comfort zone, hence the driver feels discomfort but is still able to control the DVE to avoid a crash (2). In situation (3), the DVE crosses the safety zone boundary, hence a crash becomes inevitable. Situation (4) marks the last pre-crash state outside of the safety zone where the outcome of the crash is dependent on impact type and skidding dynamics. [20]

## 2.2.4 Overtaking manoeuvres

### Definitions of “overtaking”

There exists a distinction between overtaking and passing in traffic [21]. Overtaking is generally defined as the process of moving past another vehicle or road user who is travelling in the same direction [5]. The German traffic law for example describes overtaking (§5 in [22]) as a special case of passing (§6 in [22]), in which one traffic participant passes another traffic participant who is moving in front on the same lane and in the same direction. It further clarifies that an overtaking is an intentional process, hence steering out to the left to gain an overview of the traffic situation is not an overtaking, whereas steering out at higher speed is an overtaking. The overtaking is completed as soon as the overtaking vehicle steers back in front of the overtaken vehicle or when it has travelled far enough in longitudinal direction that a safe steering back into the original lane is possible. The law states that overtaking is only allowed when the relative speed to the overtaken vehicle is significantly high, at least 10 km/h, in order to avoid obstruction of the other traffic participants. An overtaking is further forbidden if the oncoming traffic is obstructed or the visibility conditions not sufficiently good. [22] With regard to pedestrians, the overtaking definition still holds. Pedestrians who are walking on the lane in the same direction as the driver, are overtaken. If the pedestrian stops walking or evades off the lane while the vehicle is approaching, the manoeuvre definition changes from overtaking to passing. If the pedestrian is walking towards the vehicle, hence facing the traffic, an overtaking is not present due to different directions of travel. [22] In [21], the overtaking of pedestrians and cyclists is described as particularly dangerous because of the high relative speed. In case of oncoming traffic, drivers are forced to slow down the vehicle well in advance or brake hard enough to avoid a collision. In this thesis, the term overtaking is used in a more general manner, lumping together classical overtaking manoeuvres where the pedestrian is walking in the same direction as the vehicle and in the same lane, with cases where pedestrians are walking outside of the lane but close to it or facing the traffic. In contrast to most other European countries which have introduced a 1.5 m minimum distance to VRUs, there is no such law established yet in Sweden [2, 23]. The law only states that there has to be an “adequate” distance towards the overtaken road user [24].

### Definitions of overtaking phases

Overtaking manoeuvres can be performed using three main strategies. In a *flying* overtaking manoeuvre, a driver overtakes another vehicle with a roughly constant speed throughout the whole manoeuvre. An *accelerative* manoeuvre involves a speed change of the overtaking vehicle, typically a deceleration before reaching the lead vehicle, possibly a following sequence and an acceleration when the driver starts to steer away to the adjacent lane. The third strategy, *piggy backing*, defines a driver who performs the overtaking manoeuvre by following the lead vehicle close by, more or less imitating its behaviour. *Piggy backing* overtaking manoeuvres can involve an arbitrary long sequence of vehicles. [6]

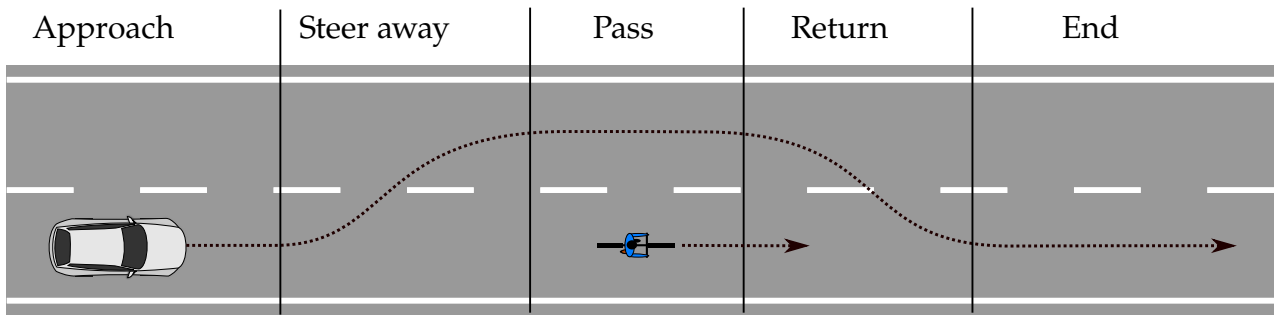


Figure 2.10: Concept of overtaking phases, here illustrated with a vehicle overtaking a cyclist, phase definitions adopted from [6, 11]

In research, vehicle overtaking manoeuvres are generally described using multiple phases. There exist three, four and five phase approaches. In [6], the manoeuvre of a driver overtaking a cyclist is defined with a four phase approach. Figure 2.10 visualises the four phase approach used in [2, 6, 11]. The first phase is the approaching phase up to the time when the vehicle starts to steer away to enter the second phase. The third phase, called passing phase, begins when the vehicle has reached a sufficient lateral distance to the cyclist and ends when it starts to return to the original lane. The process of returning is the fourth and last phase, which ends when the vehicle is back in its original lane. [6]

### 2.2.5 Driver interaction with vulnerable road-users

In previous studies, driver interaction with VRU's has been investigated in lateral and longitudinal scenarios involving mainly cyclists and pedestrians. Most of the studies made use of data from simulator experiments and test track experiments and typically involved a set of participants who executed the same set of tasks. [8, 25, 26]

Longitudinal interaction has been the subject of studies including drivers and cyclists. Studies have shown that for overtaking manoeuvres of drivers, the main influencing factor on CZBs and manoeuvre strategy is the presence of oncoming traffic [5–7, 27]. Lateral interaction has been investigated in [8], in which a crossing scenario with a cyclist and a driver was characterised from test track and simulator studies. It was found that the main influencing factor on the driver's behaviour is the time of visibility of the cyclist in the scenario [8].

Up to date, research on driver interaction with pedestrian solely focuses on lateral conflict scenarios where pedestrians cross a road while a vehicle is approaching, e.g. in a crossing scenario [25, 26]. Distinction in driver behaviour is generally made based on the moment of brake onset, hence the first time the driver applied a brake to avoid colliding with the pedestrian. Driver behaviour was found to be affected by pedestrian walking speed [26]. Other studies emphasise the importance of eye contact between driver and pedestrian which resulted in a safer braking behaviour. Furthermore, speed was found to be reduced significantly with eye contact being present. As a conclusion, the study emphasises the importance of focus of both driver and pedestrian and suggests warning pedestrians to not use smart phones while walking, for example [28].

The aim of this thesis is to investigate longitudinal conflict scenarios between drivers and pedestrians. In Figure 2.11, the concept of CZBs applied to this scenario is sketched.

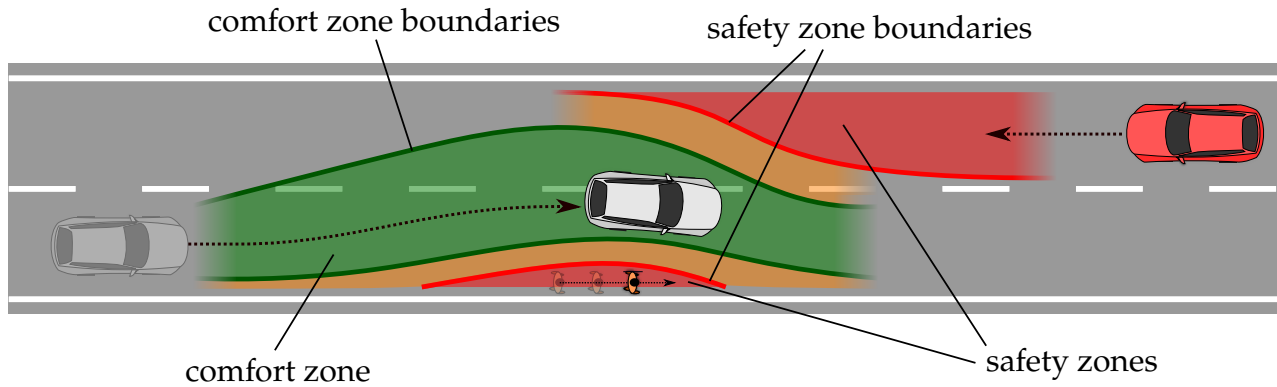


Figure 2.11: Concept visualisation of comfort zone boundaries and safety zone boundaries in a longitudinal vehicle-pedestrian scenario with oncoming traffic

## 2.3 Intelligent safety systems

From the FoST framework perspective, intelligent safety systems represent crash prevention mechanisms which aim at helping the driver by adapting to DVE changes. The aim is to ensure that the DVE system does not leave the safety zone or in case of an inevitable crash to mitigate the collision. [10, 17]

The most commonly used ISS for pedestrian protection on the market is currently the AEB which decelerates a vehicle to a full stop to avoid a crash and FCW which provides an alert to the driver to warn of an impending collision. The effectiveness of AEB in speed reduction is enhanced by improvements in the reliability of the activation algorithm, the earliness of activation and strength of deceleration. Modern approaches include the use of vacuum emergency braking (VEB) in which a vacuum pressure mechanism shoots a rubber plate towards the ground. The applied forces enhance both friction and normal force, resulting in a maximum deceleration of  $16 \text{ m/s}^2$ . [29]

### 2.3.1 Threat assessment

To decide if a traffic situation poses a threat to the ego vehicle, a threat assessment algorithm is implemented in an ISS. In [3], a driver model based threat assessment algorithm is developed to decide if a vehicle should either brake or steer in case of an imminent collision with a polygon-shaped object. The CA algorithm assumes that drivers usually steer with a constant steering wheel angle rate and at some time switch to a constant steering wheel angle input. Analogously for braking, drivers are assumed to utilise a linear acceleration profile followed by a constant acceleration as input. The algorithm predicts the ego vehicle's motion with a vehicle dynamics model and selects the proper steering or braking input to guide the ego vehicle off from the collision path. [3, 30]

In threat assessment for FCW systems as well as in the quantification of CZBs, the most common metric is time to collision (TTC). TTC is for example the time that elapses until a vehicle which is driving behind a lead vehicle at a faster speed collides with the lead vehicle, under the assumption that both vehicles maintain a constant speed and stay in the same lane. TTC is, as written in equation (2.1), defined as the quotient of the distance between



two road-users  $d_{\text{rel}}$  and the rate of change in distance, or relative speed  $v_{\text{rel}}$  [31].

$$\text{TTC} = \frac{d_{\text{rel}}}{v_{\text{rel}}} \quad (2.1)$$

TTC is further described as a more suitable metric to quantify CZBs than a spatial metric because it is less affected by vehicle speed [26, 28, 31]. An extension to TTC is ETTC (extended time to collision) which also takes the relative acceleration into account. In [31], ETTC is described as an improved metric compared to TTC due to a lower variance among different drivers [31].

Time headway (THW) is the time that it takes for a vehicle to pass a certain absolute distance with a constant speed. For example, the time which elapses until a vehicle travelling at a constant speed reaches a fixed point on the road, or the time which elapses until the front of a car reaches the front of the lead vehicle with the lead vehicle speed being equal to zero. THW is defined in (2.2) and is equal to the TTC with the approaching vehicle speed being  $v_{\text{veh}}$  and the lead vehicle being static. [32]

$$\text{THW} = \frac{d_{\text{rel}}}{v_{\text{veh}}} \quad (2.2)$$

Current approaches for ADAS TA make use of the interaction between vehicles and VRUs. In [33], a TA algorithm is developed which incorporates a walking model for pedestrians in a control strategy to improve CA.

### 2.3.2 System evaluation and assessment

ISS are evaluated with respect to performance and safety benefits. Methods of evaluation comprise the new car assessment programs (NCAPs), Hardware- and software-in-the-loop tests, field operational tests (FOTs) and test track experiments and simulation tests as well as analysis of crash data. FOTs are NDS in which system performance is analysed by means of system activation and intervention in critical situations in real traffic. FOTs are typically done by research organisations. One example for FOT evaluation is the EuroFOT project in which AEB interventions in everyday driving were analysed. [17, 34]

NCAPs such as the European NCAP (Euro NCAP) are carried out by independent organisations which assign ratings in form of points to new cars depending on their ISS performance [34]. Euro NCAP evaluates passive safety systems for pedestrian protection since 1997 while active ISS were introduced in 2016 [29, 35, 36]. The latest Euro NCAP test protocol for AEB systems with VRU interaction (version 2.0.2, released in November 2017, valid from January 2018) includes both lateral and longitudinal scenarios which involve cyclists and pedestrians. Pedestrians are distinguished among age groups. One of the scenarios is Car-to-Pedestrian Longitudinal Adult (CPLA) in which a forward travelling car approaches an adult pedestrian who is walking in the same direction inside the car's lane and who is hit by the car if the AEB does not intervene or an evasive steering manoeuvre is performed after an FCW. The scenario is sketched in Figure 2.13. The scenario is divided into one scenario in which the pedestrian is struck at a lateral offset of 25% of the car's width (CPLA-25, Figure 2.12a) and one scenario with 50% offset (CPLA-50, Figure 2.12b) [37].

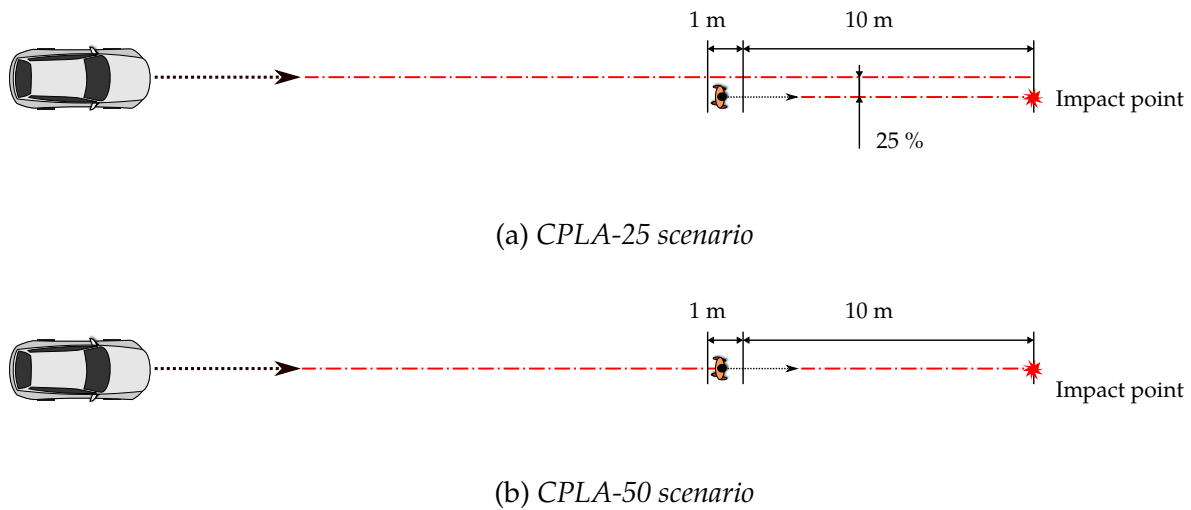


Figure 2.12: Euro NCAP scenario “Car-to-Pedestrian Longitudinal Adult” (CPLA) for different offsets, the 1 m distance is the acceleration distance of the pedestrian, the impact happens after the dummy pedestrian has walked another 10 m, images recreated from [37]

The assessment methodology is explained in the latest Euro NCAP assessment protocol for pedestrian protection (version 9.0.2, released in November 2017, valid from January 2018). CPLA-25 tests are performed for FCW’s at car speeds ranging from 50 to 80 km/h and a pedestrian speed of 5 km/h, under day and night conditions. CPLA-50 tests for AEB interventions under the same conditions as CPLA-25 but with car speeds from 20 to 80 km/h. Euro NCAP assigns a maximum score of 6 points for pedestrian AEB systems combining both lateral and longitudinal scenarios, including 3 points for day and 3 points for night scenarios. For CPLA-25 and CPLA-50 a total sub score of 30 points is achievable in which different speeds yield different amount of points for instance. The sub scores are normalised and multiplied with 3 points according to the respective condition, day or night. Sub scores are only awarded if the FCW system of the car triggers a warning at latest 1.7 s TTC. [38]

## 2.4 Naturalistic driving studies

### 2.4.1 A brief overview

Naturalistic driving studies (NDS) were introduced in the late nineties of the 20th century and represent a form of data collection in which driving data is systematically collected from a set of participants during their everyday trips over a longer period of time. The aim is to not constrain the data collection by a strict experimental protocol in order to gain insights into driver’s natural behaviour in traffic. NDS typically record vehicle internal and external sensor data which ranges from video and audio data to telemetry or even questionnaire data. Hence, vehicles which are used in NDS are equipped with sensors, cameras and a data logger to temporarily store the data until the data is extracted for analysis. Gained insights into driver behaviour as well as interaction with other road users and conflict situations can be an important input of knowledge to reduce road fatalities and reduce the environmental impact of transportation. [39, 40]

Recent NDS are researching the application of deep learning techniques to analyse driver behaviour. The MIT-AVT (Massachusetts Institute of Technology, Advanced Vehicle Technology Consortium) NDS for example, conducted within 21 months from 2016 to 2017 investigated driver behaviour and interaction with automation. The study recorded and evaluated camera, IMU, GPS, and CAN data from the vehicles at all times with a total number of driven kilometres around 440,000. Driver behaviour patterns were extracted with deep learning based computer vision techniques to estimate the driving scene, the driver body pose and the driver state. [40, 41]

## 2.4.2 UDRIVE

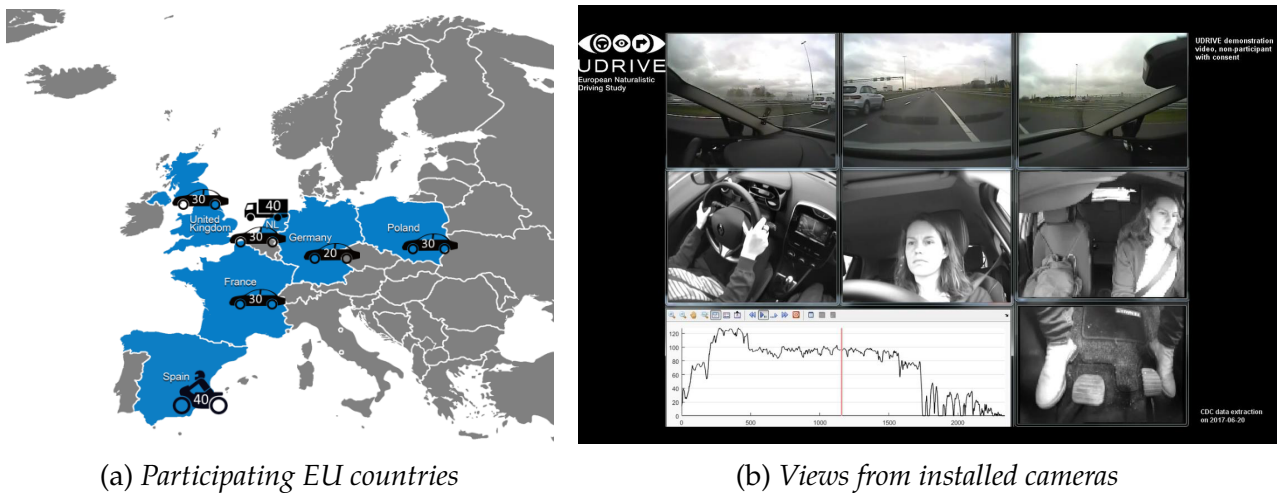


Figure 2.13: UDRIVE naturalistic driving study, images taken from (a) [42] and (b) [43]

UDRIVE is the acronym for “European naturalistic Driving and Riding for Infrastructure and Vehicle safety and Environment”. It is the latest completed European NDS issued by the European Commission and finished in 2017 with the aim to investigate aspects of sustainability in traffic on European roads. Sustainability includes driving ecology in terms of driven distance and actuator usage as well as traffic safety by means of driver behaviour and interaction with other road-users. UDRIVE shall furthermore help at understanding differences in driver behaviour among different countries and cultures in Europe [44]. The study is led by 19 European research institutions which have organised the conduction of the study in six different countries, United Kingdom, Netherlands, France, Spain, Germany and Poland (Figure 2.13a). The project budget compounds 10.6 million Euro out of which 8 million are covered by EU funding. The data set contains 38,157 h of driving data from 192 car drivers yielding several million of driven kilometres. All used cars in the study are from a set of three different Renault compact car types. The cars are equipped with a data acquisition system (DAS) which logs inertial data from a Phidgets IMU (inertial measurement unit), CAN (controller area network) data, GPS (global positioning system) data as well as steering wheel angle. Video cameras provide views of the road, the steering wheel, the driver’s face, the cabin and the pedals (Figure 2.13b). Furthermore, a MobilEye (ME) camera provides positions of obstacles around the car obtained from image recognition. The data is hosted at SAFER in Gothenburg, Sweden, and can be accessed through the partner organisations.

Results from UDRIVE are summarised and published in form of project deliverables. [39, 45]

The deliverable D44.1 of the UDRIVE project includes driver interaction with VRUs. Therein, the interaction between drivers and pedestrians is analysed for possible conflict situations and driver behaviour in presence of pedestrians on the road. Using the pedestrian collision warnings (PCWs) triggered by the ME system, the timing of the alert is compared to the time when the driver takes action, e.g. by braking. Benefits of warning the driver are verified by the fact that drivers brake harshly after the PCW onset. The study estimates a benefit of an early alert by a PCW mainly for situations involving high speed. [12]

In deliverable 51.1, driver interaction with VRUs is investigated from near-crashes in the UDRIVE data. In the results, the study claims that most of the near-crashes would not have occurred if the car driver and the VRU were physically separated. The study further states that drivers become more aware of other potential pedestrians when pedestrians are already present in the scene. Results also show that the more dangerous interactions or conflicts are associated to higher car speeds which require harsher braking. [45]

# 3

## Methodology

### 3.1 Overtaking manoeuvre terminology

This section describes the terminology used within this thesis to describe an overtaking manoeuvre. The terminology includes the definitions of the overtaking phases as well as the definitions of the factors which may influence the driver behaviour.

#### 3.1.1 Definitions of phases

The definitions of the overtaking phases follow the definitions of the phases for bicycle overtaking manoeuvres defined in [6]. Figure 3.1 shows the start of the overtaking manoeuvre at time  $t_{0 \rightarrow 1}$ , the start of phase 1, when the driver recognises the pedestrian and starts the approaching phase. In this case, the pedestrian is walking in the lane and in the same direction as the car is travelling which is the proper overtaking scenario according to [21, 22]. However, within this thesis, also other cases where the pedestrian is walking either outside of the lane or opposite to the traffic, i.e. passing events, will be referred to as overtaking out of convenience.



Figure 3.1: Start of phase 1 of the overtaking manoeuvre, driver recognises pedestrian, symbolised by red circles

Figure 3.2 shows the start of the second phase of the overtaking manoeuvre, called steer away phase, occurring at time  $t_{1 \rightarrow 2}$ , when the driver starts to steer the vehicle away from the collision course with the pedestrian. In accordance with [2], the minimum Euclidean distance between vehicle and pedestrian at the time of steer away is used as the metric minimum approaching gap (MAG). Hence, MAG is the Euclidean distance between vehicle and pedestrian at time  $t_{1 \rightarrow 2}$ .

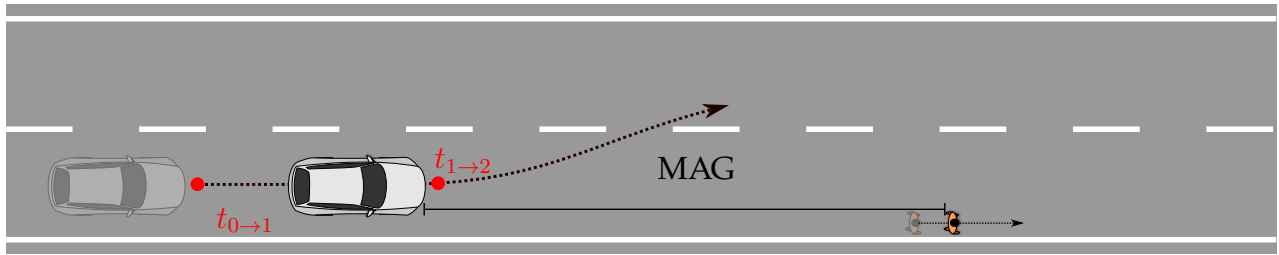


Figure 3.2: Start of phase 2 of the overtaking manoeuvre

Phase 3 of the overtaking manoeuvre starts when the vehicle reaches a longitudinal distance of 3 m to the pedestrian which marks the moment when the car is approximately parallel to the pedestrian path. This phase is called passing phase and starts at time  $t_{2 \rightarrow 3}$ . The definition leans to the convention used in [2, 6, 23] and will in the following chapters be investigated for the applicability on pedestrian overtaking manoeuvres.

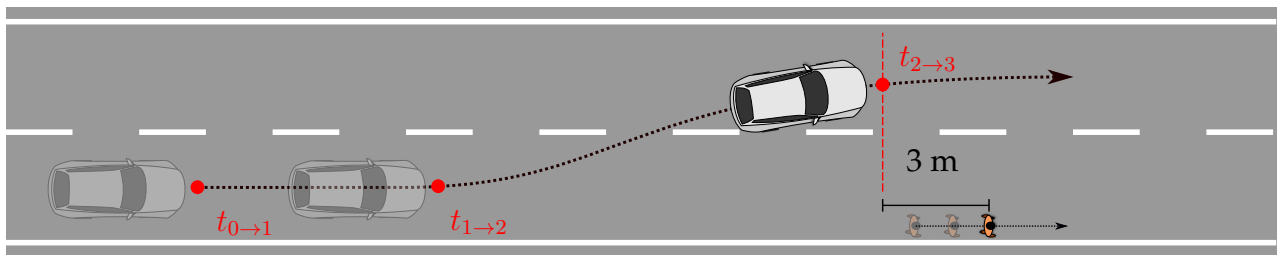


Figure 3.3: Start of phase 3 of the overtaking manoeuvre

Figure 3.4 shows the end of the passing phase 3 m in front of the pedestrian at time  $t_{3 \rightarrow 4}$  in longitudinal direction which marks the start of phase 4 when the driver steers back to the original path. Phase 4 is called return phase. The minimum clearance, MC, between vehicle and pedestrian marks the minimum Euclidean distance between the closest point of the vehicle to the pedestrian throughout the whole overtaking manoeuvre.

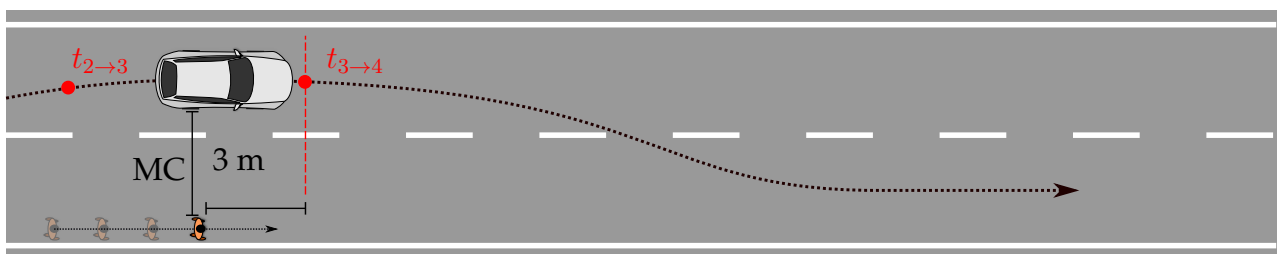


Figure 3.4: Start of phase 4 of the overtaking manoeuvre

The end of the return phase is reached at time  $t_{4 \rightarrow 5}$ , as soon as the driver has steered the vehicle back to its original path. This time marks the end of the overtaking manoeuvre.

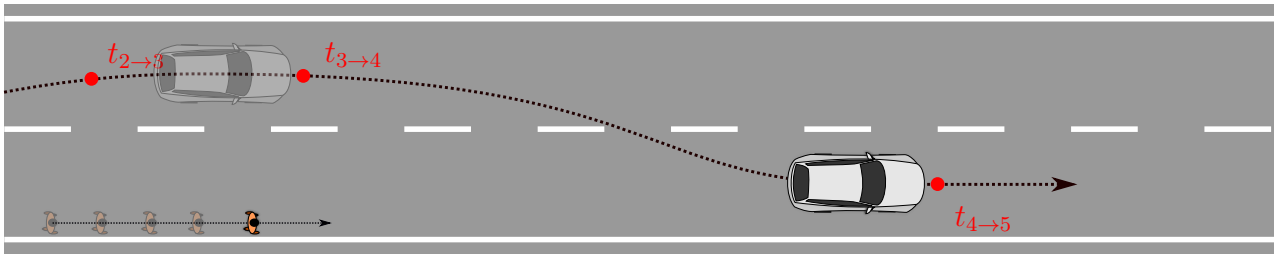


Figure 3.5: Start of phase 5 of the overtaking manoeuvre

### 3.1.2 Definitions of factors

Table 3.1 shows the factors which are used in this thesis to distinguish among different types of overtaking manoeuvres. Some of the factors are solely used for one of the two data sets, UDRIVE and field tests, which is marked in the table.

Factor	Type	Values	UDRIVE	Field tests
Oncoming traffic	Categorical	absent, present	used	used
Pedestrian direction	Categorical	same, opposite	used	used
Multiple pedestrians overtaken	Boolean	true, false	used	not used
Overtaking strategy	Categorical	accelerative, flying, piggybacking	used	not used
Pedestrian lateral position	Categorical	curb, line	used	used
Overtaking vehicle type	Categorical	short, medium, long	not used	used

Table 3.1: Overview of the terminology used to describe factors which potentially influence an overtaking manoeuvre

#### Oncoming traffic present

The factor oncoming traffic is set to “present” if an oncoming vehicle is approaching in the adjacent lane within a distance of 120 m from the position of the ego vehicle at  $t_{0 \rightarrow 1}$  or within a distance of 20 m from the position at  $t_{4 \rightarrow 5}$ . Figure 3.6 shows the two conditions which are adapted from [6].

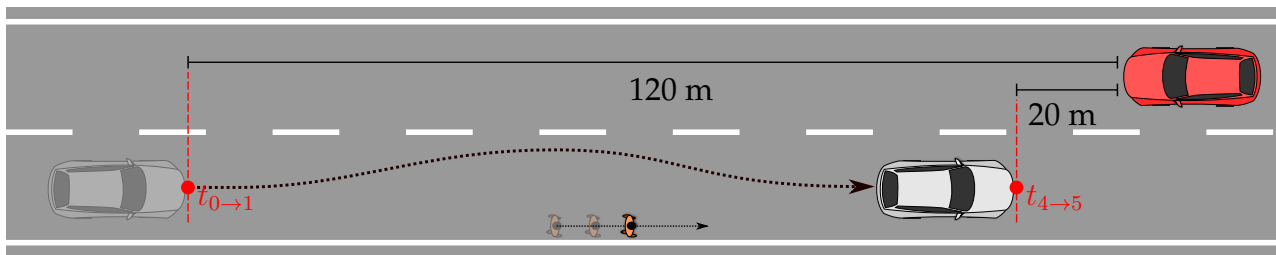


Figure 3.6: *Overtaking vehicle (white) facing oncoming vehicle (red), the factor oncoming traffic is set to present if the oncoming vehicle is anywhere between 120 m away when the overtaking vehicle is starting to approach and 20 m away when the overtaking vehicle has returned*

### Pedestrian facing driver

The factor pedestrian direction is set to “opposite” if the pedestrian is walking in opposite direction to the overtaking car, i.e. facing the traffic in the nearest lane. Otherwise it is set to “same”.

### Multiple pedestrians overtaken

The multiple pedestrians overtaken factor is set to true in case the driver overtakes more than one pedestrian during a manoeuvre.

### Overtaking strategy

The factor overtaking strategy is categorical and set to “accelerating” if the absolute speed reduction during the approaching phase is more than 10 km/h. This definition is in accordance with [11] and [2], in which a clear speed reduction is set as indicator for an accelerative overtaking manoeuvre. If the speed reduction condition is not met, a manoeuvre is classified as “flying”. “Piggybacking” is defined if the time headway (THW) of the vehicle with respect to the lead vehicle is less than or equal to 3 s. This refers to the definition in [6] in which a THW of 2 s to the lead vehicle is used to detect piggybacking manoeuvres.

### Pedestrian lateral position

Pedestrian lateral position is a categorical factor specifying at which lateral offset the pedestrian is walking on the road. The value “curb” means the pedestrian is walking close to or on the road edge, “line” means the pedestrian is walking on the lane marking.

### Overtaking vehicle type

The factor overtaking vehicle type, is categorical with possible values “short”, “medium”, “long” and describes the length of the vehicle which is overtaking. “Short” is defined as a length less than 5 m, “medium” is defined as between 5 and 10 m, and “long” is defined for vehicles which are longer than 10 m.



## 3.2 UDRIVE data

The following section presents the procedure which is used to extract, annotate and analyse events from the UDRIVE data set in which drivers overtake or pass pedestrians which are either in their lane, walking on the lane marking line or in case of a rural road without lane markings walking on the curb.

### 3.2.1 Overtaking event extraction

The major part of data screening, querying and extraction for the UDRIVE data is performed using a tool called “Smart Application for Large Scale Analysis” (SALSA). The tool is developed in MATLAB by the french non profit organisation CEESAR (“Centre Européen d’Etudes de Sécurité et d’Analyse des Risques”) and installed on the analysis computers which have access to the UDRIVE database.

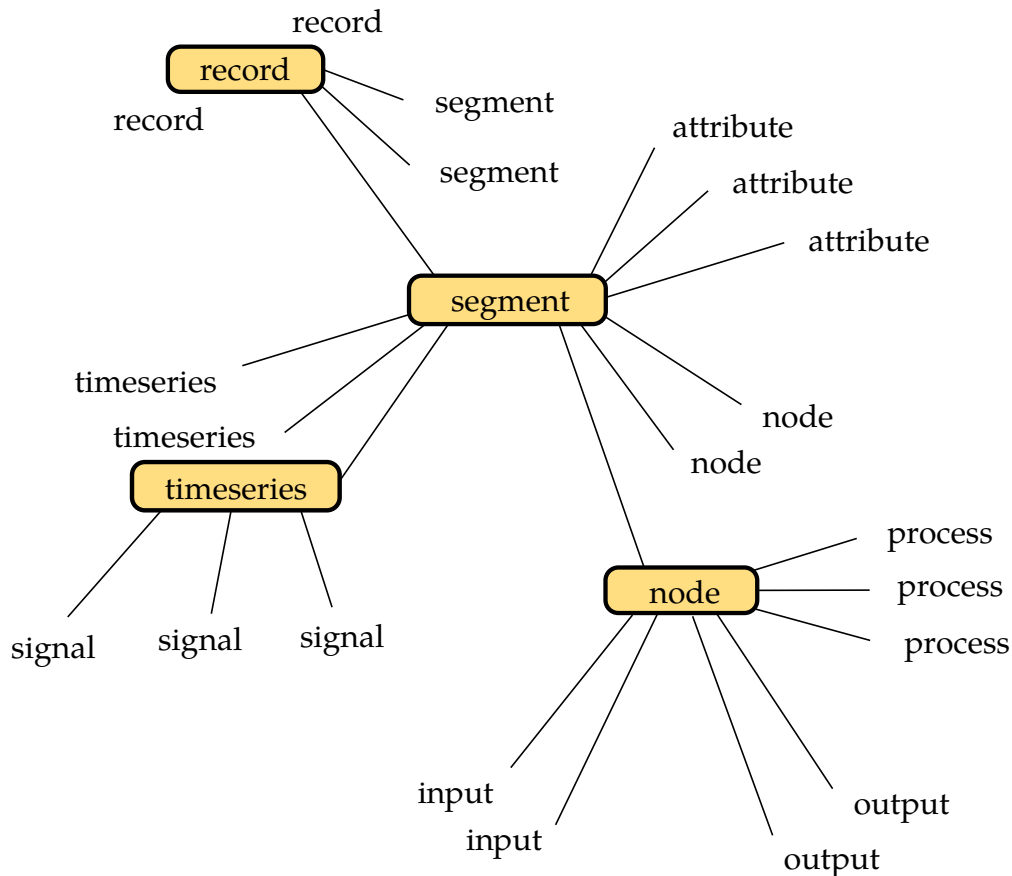


Figure 3.7: *Data structure in SALSA*

In SALSA, users can define a certain hierarchical data structure in order to process data, shown in Figure 3.7. A record is the top layer of the structure and can have various time segments. A segment is typically the starting point of a data extraction procedure. Segments are generated from a user specified signal like a state change signal which yields the start and end time of the segment. A segment can contain multiple attributes, time series and nodes.

Attributes are single parameters which can be set automatically or by manual annotation. Time series contain signals which can be raw signal data or processed signal data from other segments. Under nodes, users can write MATLAB scripts, so called processes, which are algorithms that run over the segment, taking into account specified input and output attributes. Each node can have multiple inputs and outputs which can be used by processes.

One purpose of the tool is data querying and visualisation, exemplified in Figure 3.8. Users can work in “Test mode” to develop algorithms to search for specific events and test them on small sub sets of the database, a so called “Batch” process. During a “Push”, the algorithms developed on the local database can be integrated into the global database. The algorithms are then run on all data in the global database. In “Production mode”, the global database can be used [12] to annotate events, i.e. mark different times in an event or set different factors. In both modes, data from cameras and sensors can be visualised and queried.

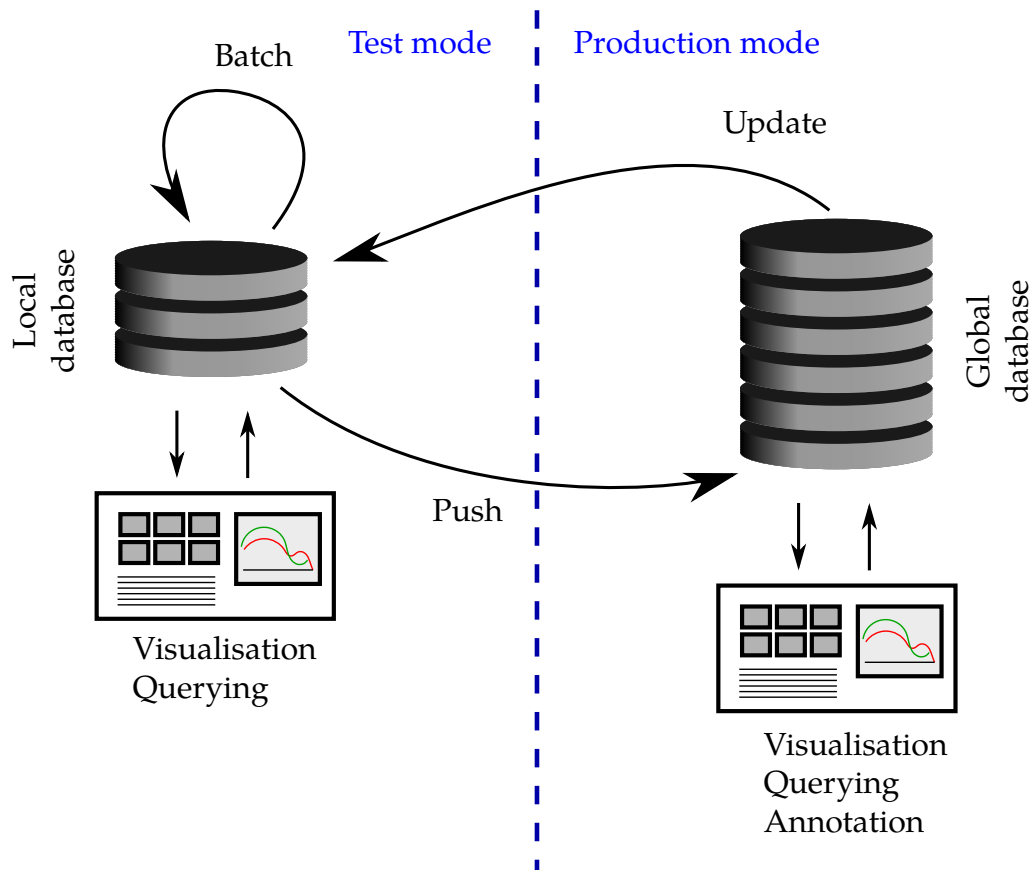


Figure 3.8: Interaction between global and local database in SALSA

SALSA is a collaborative tool in which multiple partner organisations of UDRIVE develop algorithms simultaneously. In order to prevent high computational load and to synchronise segments among UDRIVE partners, the administration “pushes” all temporary changes to the common database on a regular base. As soon as segments are pushed, they can be accessed by all partner organisations and can be applied to the whole UDRIVE data set. A push is usually scheduled weekly by the administration and can take up to three days to be

processed.

### 3.2.2 Extracting overtaking candidate segments

The segment developed in this thesis is based on previous work from [11], in which cyclist overtaking manoeuvres were extracted from UDRIVE. The segment is generated from a state change signal which yields the start and end time of a segment when the ME camera loses track of a VRU within 50 m in longitudinal direction if the ego vehicle travels faster than 20 km/h. The start and end time of the segment was set to the time at which the VRU disappears from the ME view  $\pm 10$  s, respectively, to capture the whole possible overtaking process. Figure 3.9 illustrates the moment when the VRU, in this case a pedestrian, disappears from the ME detection field.

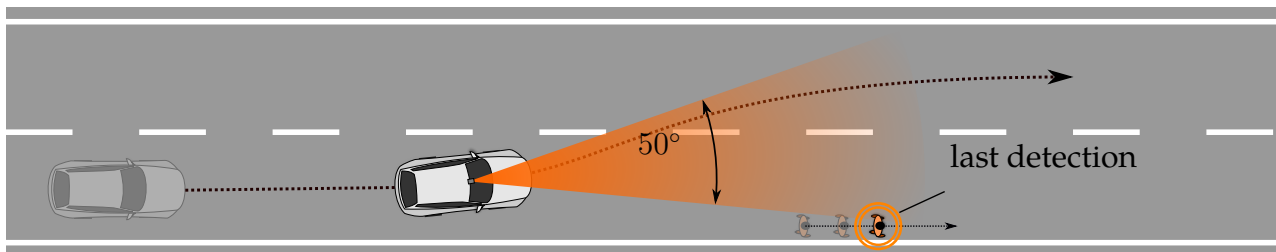


Figure 3.9: *Pedestrian disappears from ME detection field*

After this step, 960282 matching segments were found in the database. Another data quality filtering step was then done to exclude segments which included less than 15 ME detections or contained less than 68% of useful variables, hence more than 32% NaN (not a number) values. To further classify a segment to be an overtaking manoeuvre, a condition on the peak lateral acceleration was used, based on [11]. Therein, an overtaking candidate is a segment which passes the quality check and shows a pattern in the lateral acceleration signal. If the lateral distance to the pedestrian, obtained from ME data 6 time instances around the peak values of the lateral acceleration, is greater than the mean lateral position of the pedestrian, a segment is classified to be an overtaking candidate. Another condition which was used to exclude candidates is based on the steering wheel angle, filtering out events where steering wheel angles were greater than  $180^\circ$ , to avoid events where pedestrians are standing still at the corner of an intersection while the driver turns around the corner.

Misclassifications of pedestrians by ME occurred frequently. One common case was a road sign which was classified as a pedestrians. The other common case was a cyclist which was located far away and looked similar to a pedestrian from behind until the car approached closer and the type of detection changed to cyclist. The latter case was treated with a filter which classifies a pedestrian event only if the last instance of the ME detection is a pedestrian. After this step, there were 5315 pedestrian overtaking events left. From those events, 117 events were manually selected as proper longitudinal events for analysis, i.e. there was no misclassification, the event occurred in a rural area and the road was roughly straight. Out of the 117 events, 77 occurred in France, 27 in United Kingdom, 5 in Germany, 4 in Poland and another 4 in the Netherlands. Because of the small number of events in other countries and the limited time frame of this thesis, only events from France were analysed. Figure 3.10 shows a map of the events found in France, which occurred mostly on rural roads in the

south-eastern parts around Lyon and Avignon.

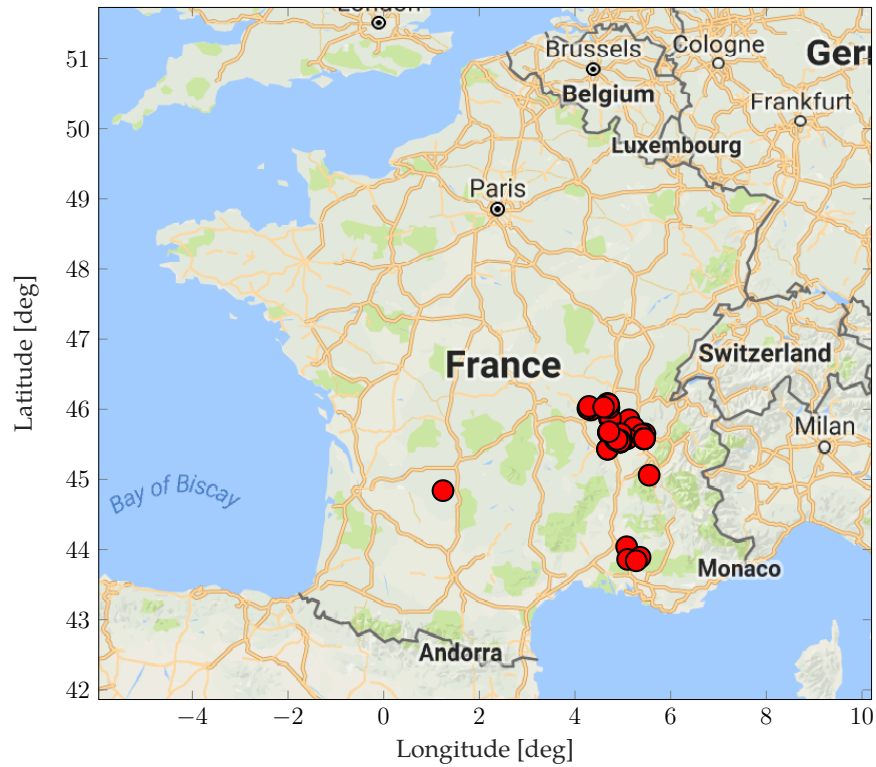


Figure 3.10: Map of the 77 pedestrian overtaking events in France. The events mainly occurred in the south-eastern region.

Repetitions of drivers occur within the used data set of overtaking events in France. Figure 3.11 shows the driver ID distribution. Out of the 77 events, almost half is represented by three individual drivers.

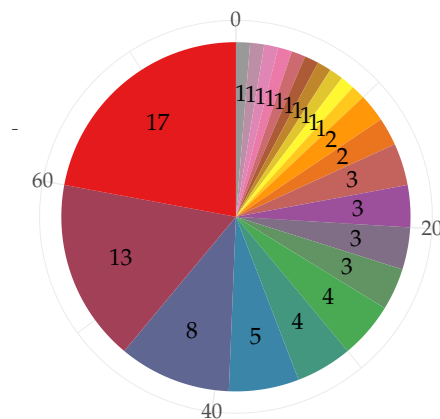


Figure 3.11: Pie chart of driver ID distribution, each colour represents one driver ID. The three most frequently occurring drivers account for almost 50% of the events.

### 3.2.3 Manual annotations

To gain values for the mentioned parameters, an annotation panel was developed in SALSA which could be used by the annotator, i.e. a person, to annotate factors of an overtaking manoeuvre manually and store this annotation in the database. Table 3.2 gives an overview of the annotated attributes. The speed of the pedestrian was annotated in a categorical manner. If the video feed showed the pedestrian standing still, the speed was set to 0. If the pedestrian was walking the speed was set to 5 km/h and if it was running to 9 km/h. The overtaking type was set to piggy backing when the front camera video feed revealed that the driver was clearly following a lead vehicle during the overtaking manoeuvre. Manoeuvres were annotated accelerative when there was a clear decrease in speed, i.e. 10 km/h, visible in the CAN speed data during the approaching phase.

The start times of the different overtaking phases were annotated with the help of the video feed, mainly from the front and the driver facing camera. The approaching phase start was set to the time when the pedestrian appeared in the video feed of the front facing camera, in accordance to [11]. The steer away phase start was identified from the steering wheel angle signal and video feed of the steering wheel facing camera. The start was set to the time when the driver started to rotate the steering wheel or when a change in steering wheel angle could be recognised from the CAN data. In contrast to [11] and the definition in sub section 3.1.1, the passing phase was annotated when the driver had steered the vehicle to a parallel course next to the pedestrian. This moment was identified from a pause in steering wheel movement by the driver, indicating that the driver had reached the maximum lateral deviation. The start of the return phase was then annotated accordingly when the driver started to steer back from the parallel course. Hence, in events where there was no clear time period where the driver was keeping the vehicle parallel to the pedestrian, the start time of phase 3 and 4 were set equal. The end of the return phase was set to the time when the driver had returned to the original lane.

Attribute	Signals used for decision
Oncoming traffic	Video (front) and ME
Pedestrian direction	Video (front)
Pedestrian speed	Video (front, side)
Multiple VRUs present	Video (front, side)
Pedestrian lateral position	Video (front, side)
Overtaking type	Video (front, side), speed
Overtaking phase 1 start	Video (front)
Overtaking phase 2 start	Video (steering wheel), steering wheel angle
Overtaking phase 3 start	Video (steering wheel), steering wheel angle
Overtaking phase 4 start	Video (front, steering wheel), steering wheel angle

Table 3.2: Overview of manually annotated attributes for identified overtaking manoeuvres in UDRIVE

### 3.2.4 Vehicle and pedestrian path reconstruction

#### Vehicle path estimation

Since the ME detections of the pedestrian usually occurred before the vehicle had passed the pedestrian, the vehicle path as well as the pedestrian path needed to be extrapolated using the known positions. The vehicle path trajectory was reconstructed from an integration of the yaw rate given from the Phidgets gyroscope,  $\omega_{\text{gyro}}$ , and using a trapezoidal integration. The integration was started as late as possible, hence from the time  $t_{1 \rightarrow 2}$  when the vehicle started to steer away, in order to reduce the deviation from the actual trajectory by the gyroscope drift effect. The side slip angle of the car was further assumed to be close to zero during the overtaking event since it was observed that drivers almost never steered away with a sudden or large steering wheel input. The dynamical model hence used was a point mass model, visualised in Figure 3.12.

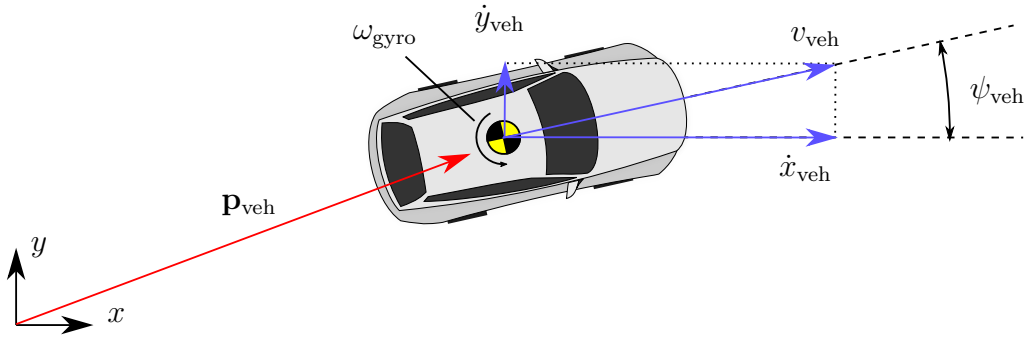


Figure 3.12: Visualisation of used point mass model for vehicle path extrapolation

The path position of the vehicle,  $\mathbf{p}_{\text{veh}}$  could then be estimated from the point mass dynamical model by

$$\dot{\mathbf{p}}_{\text{veh}}(t) = \begin{bmatrix} \dot{x}_{\text{veh}}(t) \\ \dot{y}_{\text{veh}}(t) \end{bmatrix} \approx \begin{bmatrix} v_{\text{veh}}(t) \cos \psi_{\text{veh}}(t) \\ v_{\text{veh}}(t) \sin \psi_{\text{veh}}(t) \end{bmatrix} \quad (3.1)$$

$$\Rightarrow \mathbf{p}_{\text{veh}}(t) = \int_{t_{1 \rightarrow 2}}^{t_{3 \rightarrow 4}} \begin{bmatrix} v_{\text{veh}}(t) \cos \psi_{\text{veh}}(t) \\ v_{\text{veh}}(t) \sin \psi_{\text{veh}}(t) \end{bmatrix} dt. \quad (3.2)$$

Discrete trapezoidal integration of the dynamical model yielded the vehicle heading angle

$$\psi_{\text{veh}}(t_{k+1}) \approx \frac{t_{k+1} - t_k}{2} \left[ \dot{\psi}_{\text{veh}}(t_k) + \dot{\psi}_{\text{veh}}(t_{k+1}) \right] \quad (3.3)$$

$$= \frac{1}{2 f_{\text{gyro}}} [\omega_{\text{gyro}}(t_k) + \omega_{\text{gyro}}(t_{k+1})], \quad (3.4)$$

with  $t_k, t_{k+1} \in [t_{1 \rightarrow 2}, t_{3 \rightarrow 4}]$  and  $\psi_{\text{veh}}(t_0 = t_{1 \rightarrow 2}) = 0$ .  $f_{\text{gyro}} = 1/(t_{k+1} - t_k) = 30$  Hz is the sample rate of the Phidgets gyroscope.

The position of the vehicle was then estimated by another discrete trapezoidal integration of the dynamical model in (3.1).

## Pedestrian path estimation

Due to error in the position estimate from the ME camera, the trajectory of the pedestrian deviated unreasonably much in  $x$  and  $y$  direction. However, the path of the pedestrian was usually still visible from the points. In order to obtain a satisfactory trajectory, the path was assumed to be a straight line which was found to be the most common case observed in the UDRIVE video data. In order to filter out outliers, a RANSAC algorithm was applied on the positions.

RANSAC stands for random sample consensus and is an iterative method to find models in data and return the corresponding inliers of the original data, developed by Fischler et al. in 1980 [46]. Algorithm 1 shows RANSAC for 2D line fitting. In each iteration, the algorithm chooses two random points from the original set of points  $\mathcal{P}$ , fits a line through these points and determines intercept  $b$  and slope  $m$ . By comparing the Euclidean distance of all points in  $\mathcal{P}$  to the line with a distance threshold  $\tau$ , the algorithm determines all inliers which are within this distance. The number of inliers is the objective to be optimised and thus kept throughout the loops to ensure a steady increase. After the predefined number of iterations have passed, the line parameters intercept and slope of the best line fit are output together with the corresponding inliers. [47]

**Input** : 2D point set  $\mathcal{P} = \mathbf{x}_i = (x_i, y_i)$ , threshold  $\tau$ , maximum number of iterations  $N_{\text{iter}}$ , inlier ratio  $p_{\text{in}}$   
**Output** : Line intercept  $b$  and slope  $m$ , inliers indices  $\mathbf{i}_{\text{in}}$   
**Initialise**:  $\mathbf{i}_{\text{in}} \leftarrow \emptyset$ ,  $i_{\text{iter}} \leftarrow 0$   
**while**  $i_{\text{iter}} < N_{\text{iter}}$  **do**  
    Pick two random points  $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{P}$   
    Fit a line through  $\mathbf{x}_1$  and  $\mathbf{x}_2$   
     $b_{\text{tmp}}, m_{\text{tmp}} \leftarrow$  intercept, slope of the fitted line  
     $\mathbf{i}_{\text{in,tmp}} \leftarrow$  indices of points from  $\mathcal{P}$  within a distance of  $\tau$  from the fitted line  
    **if**  $|\mathbf{i}_{\text{in,tmp}}| \geq p_{\text{in}} |\mathcal{P}|$  &  $|\mathbf{i}_{\text{in,tmp}}| \geq |\mathbf{i}_{\text{in}}|$  **then**  
         $b, m \leftarrow b_{\text{tmp}}, m_{\text{tmp}}$   
         $\mathbf{i}_{\text{in}} \leftarrow \mathbf{i}_{\text{in,tmp}}$   
    **end**  
     $i_{\text{iter}} \leftarrow i_{\text{iter}} + 1$   
**end**

**Algorithm 1:** RANSAC outliers removal algorithm for 2D line search

Figure 3.13 depicts the straight line fit concept for the extrapolation of the ME detections. For the position detections of the pedestrian, RANSAC was applied with a distance threshold of  $\tau = 0.5$  m, an inlier ratio of  $p_{\text{in}} = 0.2$  and a maximum number of iterations of  $N_{\text{iter}} = 300$ . For all overtaking events the RANSAC model was verified manually for consistence with the video feed. In case of significant deviation, the line fit was done between the mean of all ME pedestrian positions and another manually selected point.

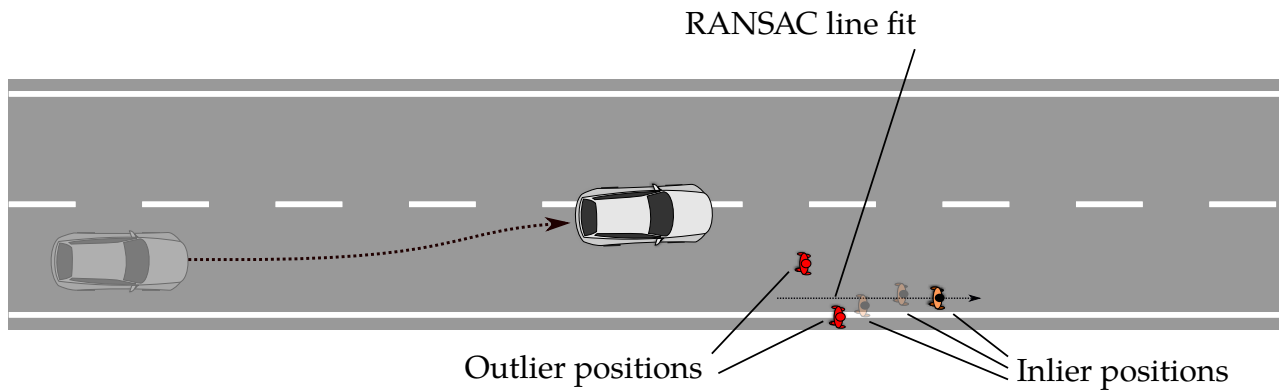


Figure 3.13: *Visualisation of RANSAC straight line fit for pedestrian position extrapolation*

The position of the pedestrian was extrapolated using the slope of the fitted line and the speed of the pedestrian. The point through which the trajectory line is going, was set to the mean of the last five ME detections which were assumed to be the more accurate detections since they happened closer to the ME camera. The speed of the pedestrian was determined from either the annotated speed or a speed measure calculated from the relative speed determined by ME, developed in [11]. The speed determined by ME was chosen if it did not deviate more than  $\pm 2$  km/h from the annotated speed, otherwise the annotated speed was chosen.

### 3.3 Pedestrian data logger

To analyse vehicle overtaking manoeuvres from the pedestrian perspective, a measurement setup was developed which can be worn by a person to conduct measurements on the road. The setup allows measuring distances to surrounding obstacles, inertial data as well as GPS data. During two days of data collection on a public road, data from 630 overtaking events was collected under the aim of investigating the formulated hypotheses.



### 3.3.1 Hardware development

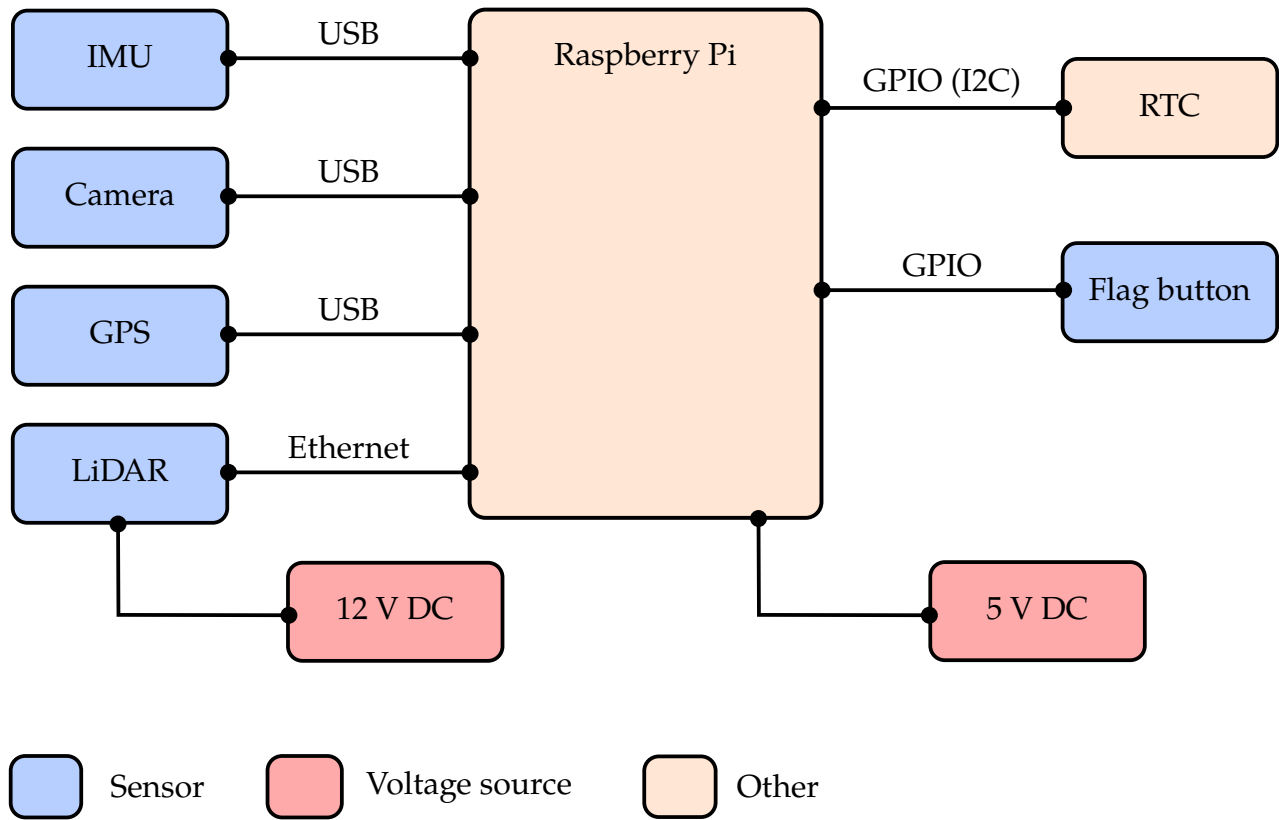


Figure 3.14: Wiring sketch of the pedestrian data logger

The components of the pedestrian data logger (PDL) are depicted in Figure 3.14. Its main computing unit is a Raspberry Pi (RPi), model 3 B which is running the operating system Ubuntu 16.04.4 LTS. The RPi houses a Quad Core 1.2GHz Broadcom BCM2837 64 bit CPU and a 1 GB RAM chip. External devices can be connected through one Ethernet port, 4 USB 2.0 ports and a variety of General-purpose input/output (GPIO) pins. The GPIO pins offer the possibility to connect input and output devices and include an inter-integrated circuit (I2C) and a serial peripheral interface (SPI) bus. The main benefit of the RPi is the comparably low cost and high computational power which has resulted in usage in previous research projects [4, 48]. To be able to use the actual time for the time stamps in the data, the Adafruit DS1307 real time clock (RTC) was interfaced to the RPi via I2C.

#### Sensor overview

The sensors connected to the RPi are summarised in Table 3.3. For range and bearing measurements, a 2D LiDAR (light detection and ranging) with a minimum and maximum angle of  $\pm 95$  deg is used which has a maximum detection range of 120 m and a guaranteed detection range of 30 m. The PhidgetSpatial IMU includes a gyroscope, accelerometer and magnetometer. The LiDAR is powered separately with a 12 V Yuasa NPW45-12 battery while

the RPi is powered with a standard 5 V power bank from Deltaco, using the 2 A USB output port.

Sensor	Product name	Connection type	Sample rate
LiDAR	Hokuyo UXM-30LAH-EWA	Ethernet	20 Hz
IMU	PhidgetSpatial 1044_0	USB	250 Hz
GPS	Globalsat BU-353S4	USB	1 Hz
Camera	Creative Live! Cam Sync HD	USB	15 fps
Flag button	-	GPIO	10 Hz

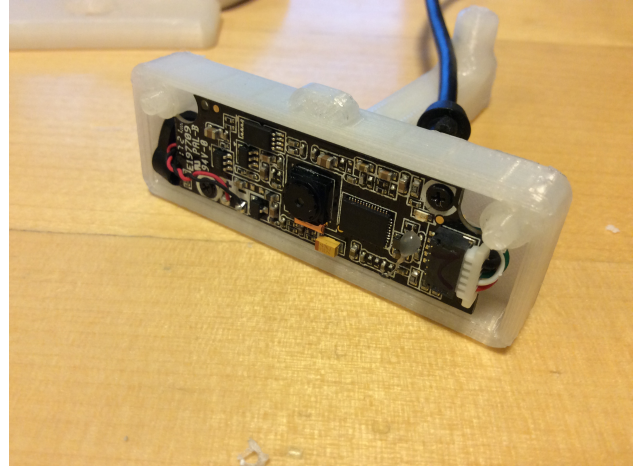
Table 3.3: Overview of sensors connected to main computer

## Physical arrangement

The LiDAR and the IMU are statically connected with a custom-made adaptor, shown in Figure 3.15a. The adaptor was designed in Autodesk Inventor 2018 and printed with polylactic acid (PLA) filament on a Dagoma DiscoEasy200 3D printer. The arrangement (Figure 3.15a) was chosen such that the IMU gyroscope axis is roughly aligned with the LiDAR axis. For adjusting the roll angle, a construction was added in form of an adaptor mounted to the backside of the LiDAR. This adaptor is fixed on one side at the person's belt and on the other side screwed to the back of the LiDAR. A rotation between the two sides is enabled by a joint, such that the angle can be adjusted with sub-degree precision. Other 3D printed components of the PDL include a case for the flag button (Figure 3.15c) which is to be held in the hand by the person wearing the PDL as well as cases for the RPi (Figure 3.15d), the camera (Figure 3.15b) and the battery.



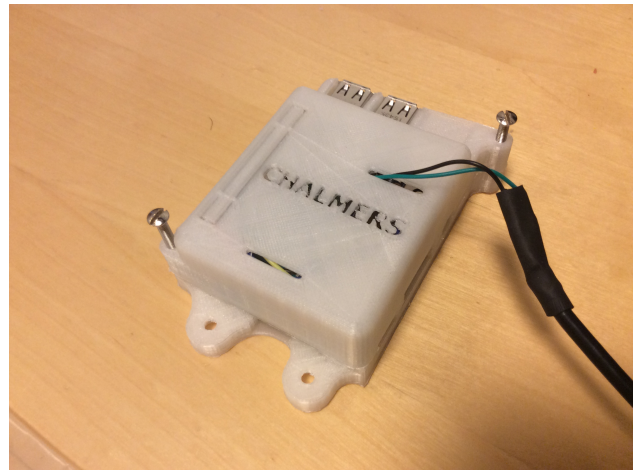
(a) LiDAR and adaptor



(b) Camera case



(c) Flag button case



(d) RPi case

Figure 3.15: Photos of the 3D printed physical components of the pedestrian data logger

### 3.3.2 Logger software

The PDL software is developed to allow logging data from a set of the listed sensors. The software is realised using robot operating system (ROS) and used via a server client connection with a web app on a smart phone.

#### ROS package

ROS is a flexible framework for all kinds of applications involving reading values from sensors or actuation. There exists a variety of open-source ROS packages online ranging from drivers over filters to complex robot control applications. The PDL software is developed as a ROS package called *div\_datalogger* which includes a web socket from the *rosbridge* package and a user controller to record data to a .bag file which is a data storage type in ROS. The main node *div\_datalogger* subscribes to the topics published by the sensors. It further advertises a service called *user\_cmd* which can be called to change the state in *div\_datalogger* node and receive a response code, possibly a message and the current recording time. The state machine is depicted in Figure 3.16. The logger is initialised in the *idle* state. From *idle*, the user can

request a state change to either start recording, state logging, or shut down the program, state shutdown. Once in logging state, the user can request to stop the logging and return to idle state.

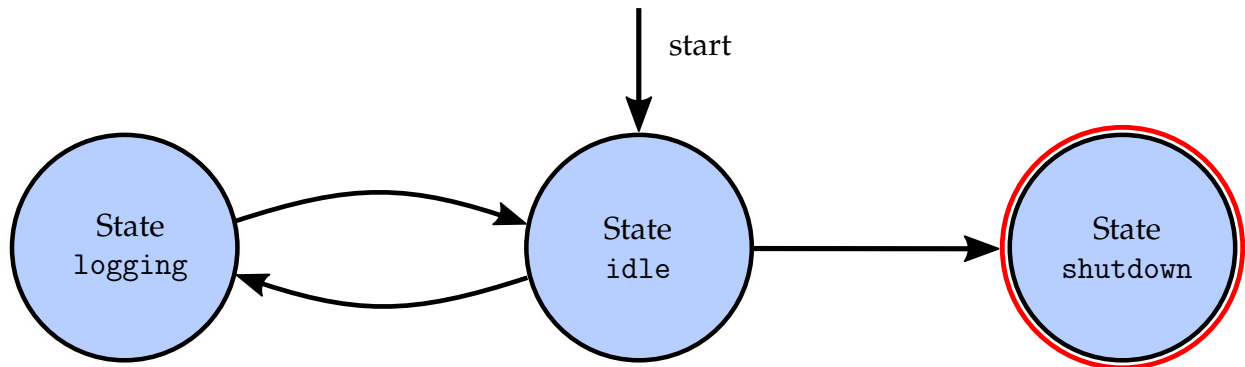


Figure 3.16: *Pedestrian data logger state machine*

The recorded topics are summarized in Table 3.4. The user control is enabled by running a ROS web socket (package *rosbridge*) in parallel.

Topic	Message type	Sample rate
/scan	Range measurements	20 Hz
/imu/data_raw	Accelerometer, gyroscope, magnetometer data	250 Hz
/time_reference	GPS time stamp	1 Hz
/vel	GPS speed (not derived from position)	1 Hz
/fix	GPS position (longitude, latitude)	1 Hz
/flagbutton_pressed	1 for pressed, 0 for released	10 Hz
/cv_camera/image_raw/compressed	Compressed image stream from camera	15 fps
/rosout	ROS master log output	-

Table 3.4: Overview of the recorded topics by the PDL

The *div\_datalogger* package is made publicly available <sup>1</sup>.

### Frame definitions

ROS packages typically deal with spatial data from several physical components which are placed at different positions in space with changing orientations. One usually defines a frame for each component of the body which is defined by the corresponding position and

<sup>1</sup>[https://github.com/ruvigroup/div\\_datalogger](https://github.com/ruvigroup/div_datalogger)

rotation. By defining the transformations between these frames, it is possible to construct a “frame tree” from which coordinates in any frame can be expressed in any other frame. In Figure 3.17, the different frames from *base\_footprint* onwards are visualised in the *rviz* tool in ROS, together with the camera output. The original frame tree is shown in Figure A.1 in Appendix A.

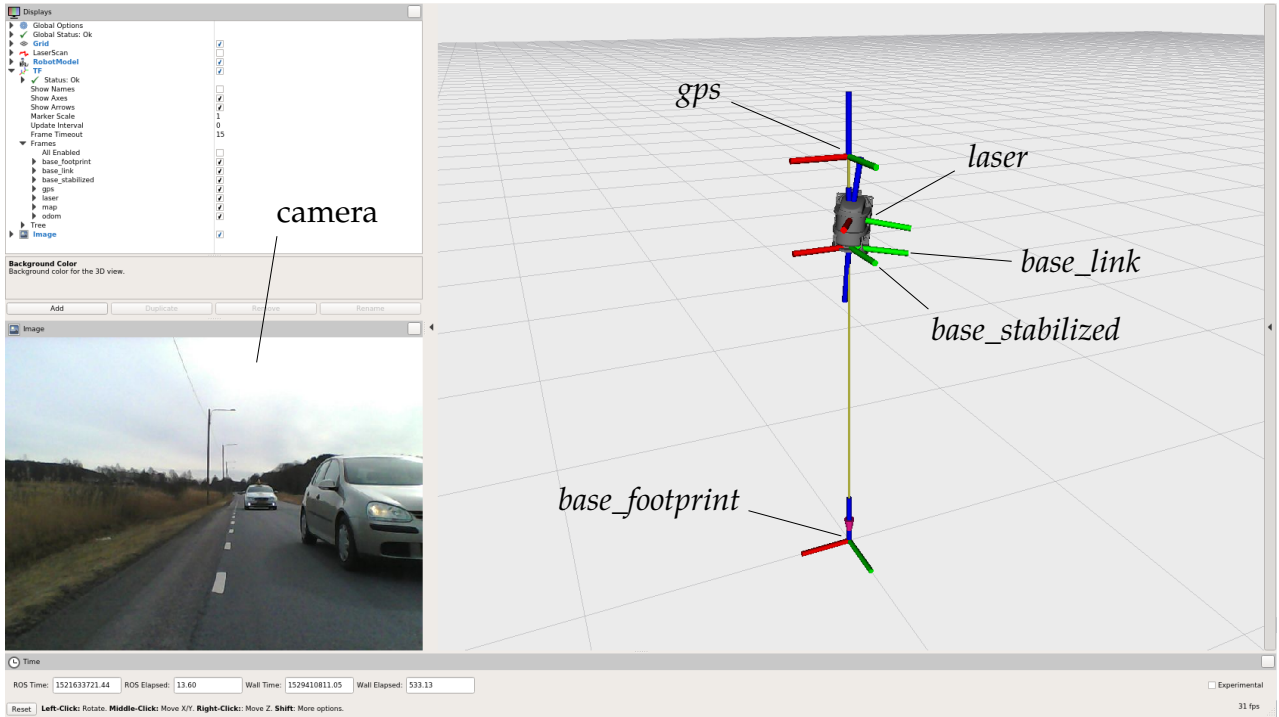


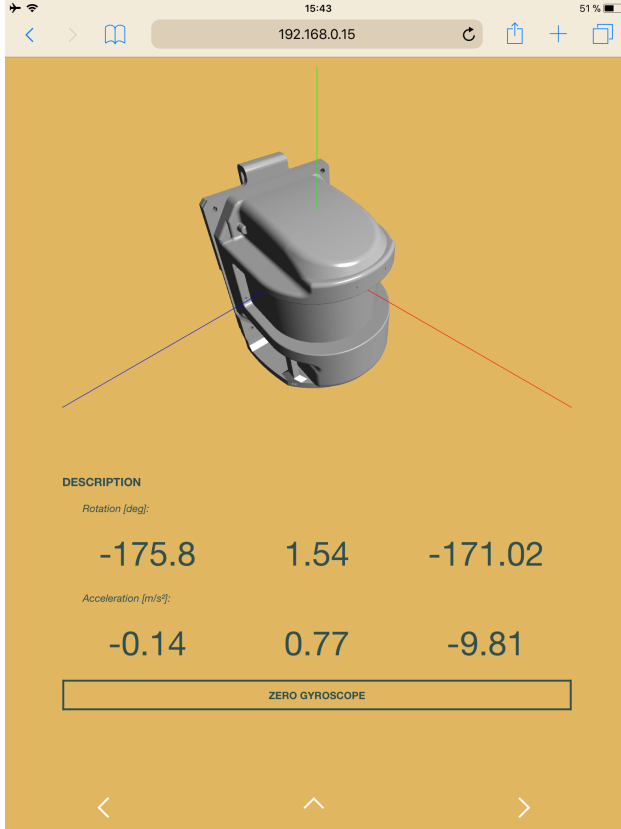
Figure 3.17: Screen shot from *rviz* visualisation tool in ROS, showing the camera image (lower left corner) and the frame tree (right side)

The original frame tree includes the *map* and *odom* frames which are not used in this thesis but could serve as reference frames for global positioning and odometry of the data logger, for example using GPS or IMU data. The branch of the frame tree used in this thesis and displayed in Figure 3.17 starts at the frame *base\_footprint* which is located on the road below the PDL. 1.2 m up in z-direction from *base\_footprint*, the *base\_stabilized* frame is located, which marks the stabilised position of the IMU. The actual IMU orientation is represented by the *base\_link* frame which is oriented according to the IMU orientation with respect to the *base\_stabilized* frame. The *laser* frame is located inside the LiDAR of the PDL by a static translation from the *base\_link* frame. The *gps* frame is translated 1.5 m in z-direction up from *base\_footprint* frame and represents the location of the GPS sensor attached to the pedestrian.

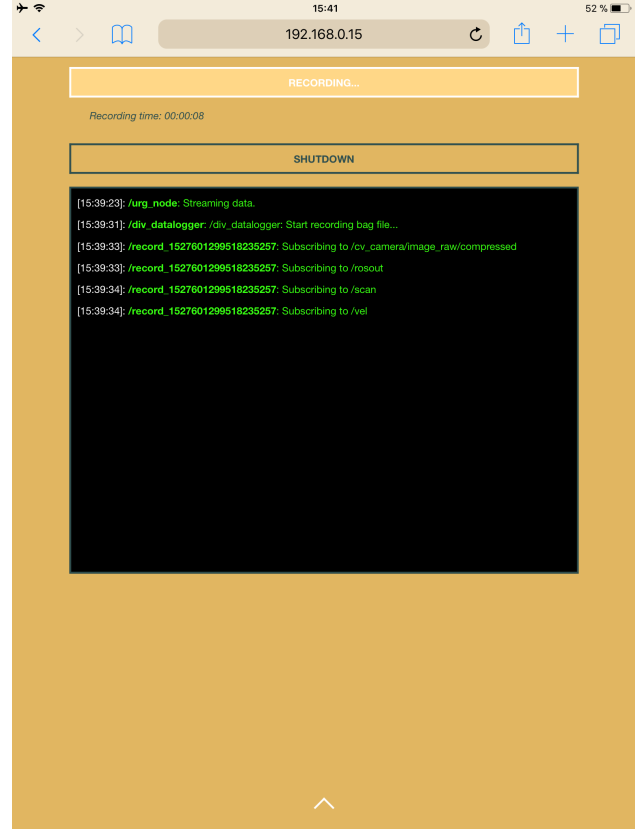
### Usage via web app

To facilitate the usage of the PDL, a web application is used which connects to the web server via WiFi. A WiFi hot spot is started by the RPi at system boot up and the user can connect to it for example with a smart phone. The user can then open the website which serves as an interface to the *rosbridge* web socket. The web app interface contains the options to start and stop recording data, to shutdown the logger and to visualise the readings of the LiDAR, the IMU orientation as well as the camera. Figure 3.18 shows two screen shots of the web app,

the visualisation mode for the IMU orientation (3.18a) and the recording screen (3.18b). Due to the physical characteristics of the gyroscope sensor, the readings can drift over time, an effect called bias [49]. In the IMU visualisation it is possible to reset the gyroscope readings via button click to zero (“zero gyroscope”).



(a) IMU visualisation



(b) Recording screen

Figure 3.18: Screen shots of the PDL control web app

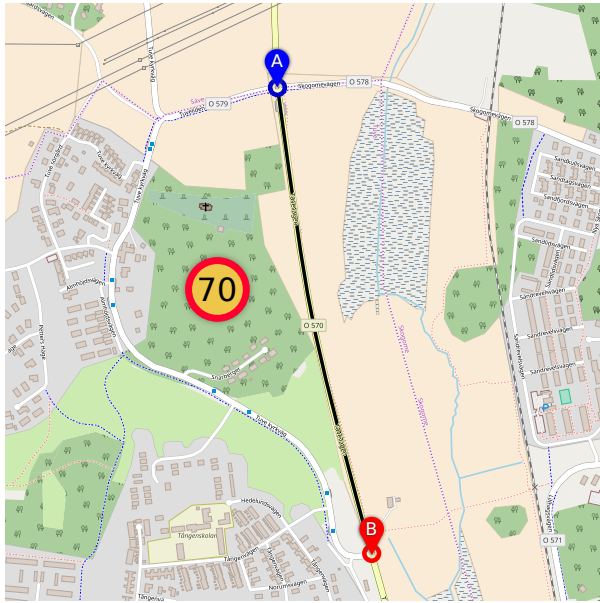
The web app code is made publicly available <sup>2</sup>.

### 3.3.3 Data collection

The data collection consists of two days of measurements on the road “Tuvevägen” in Tuve, north of Gothenburg, Sweden. The road is about 1 km long, straight and has a speed limit of 70 km/h. Figure 3.19 shows the road with the two corner points A and B between which the collection was performed.

<sup>2</sup>[https://github.com/ruvigroup/div\\_data\\_logger\\_webapp](https://github.com/ruvigroup/div_data_logger_webapp)





(a) OpenStreetMaps overview



(b) Google Street View

Figure 3.19: Data collection site at Tuvevägen, GPS coordinates point A (57.7682, 11.9357) and point B (57.7594, 11.9390)

A pedestrian was equipped with the PDL system and instructed to walk four scenarios. The first scenario includes walking on the lane marking line, case “line”, in the same direction as the traffic. In the case of line, the pedestrian was walking on the line with the foot which was closer to the traffic such that the LiDAR was almost positioned inside the lane. The second scenario includes walking on the line while facing the traffic. The third and fourth scenario was performed just next to the road edge, case “curb” while walking in the same and the opposite direction as the traffic, respectively. Figure 3.20 shows the two lateral positions of the pedestrian, line and curb.

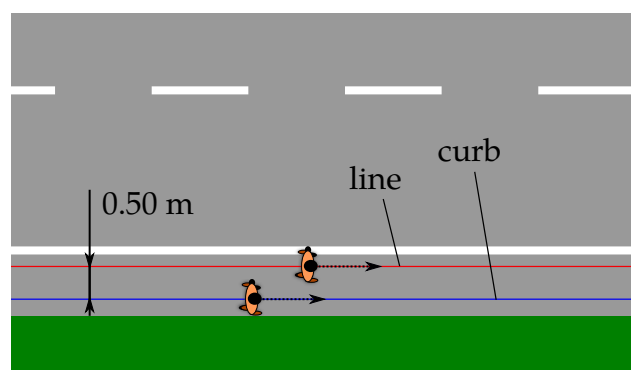


Figure 3.20: Two lateral pedestrian positions tested, in case line the pedestrian is walking and curb

In all scenarios the LiDAR which was mounted at the hip of the pedestrian (see Figure 3.21a) faced into the road, hence, when switching pedestrian direction, the LiDAR needed to be mounted at the opposite hip. During the data collection, a measurement team was close by and at all times connected to the pedestrian via cell phone. About 630 overtaking and passing manoeuvres were recorded in total. Figure 3.21b shows the cockpit of the

measurement team car, provided by Autoliv AB, during an overtaking manoeuvre. Figure 3.21c shows the pedestrian walking on the line in opposite direction to the traffic.



(a) Pedestrian with the PDL



(b) Cockpit view from measurement crew car



(c) Pedestrian in line scenario, walking direction opposite to traffic

Figure 3.21: Photos from data collection on 13th March, 2018

### 3.3.4 Data preparation and analysis

The aim of processing the recorded LiDAR and IMU data was to be able to track vehicles over time, identify the metrics speed and MC during the overtaking manoeuvre, and to identify the factors oncoming traffic and piggy backing. The LiDAR and IMU data was used in a C++ program together with point cloud library (PCL). PCL contains useful implemented functions for filtering, segmenting and clustering point clouds. Figure 3.22 shows the flow chart of the analysis process.



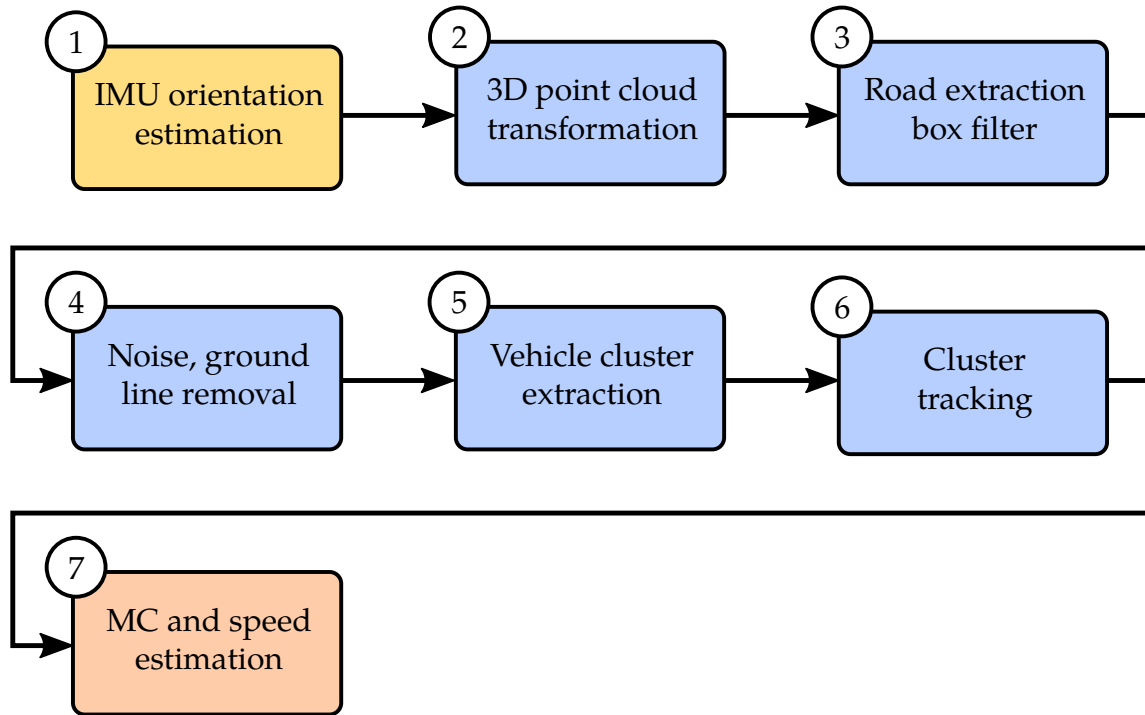


Figure 3.22: Flow chart of the data analysis process, yellow background indicates use of IMU data, blue background use of LiDAR data, red background symbolises the final output of the analysis

### Orientation estimation

The transformation from the *base\_link* to the *base\_stabilised* frame was estimated from the IMU data. The Madgwick attitude and heading reference system (AHRS) filter was used for the estimation. The filter takes accelerometer, gyroscope and magnetometer readings as input and calculates the orientation in quaternion representation using an optimised gradient descent algorithm. The advantage in usage of the Madgwick filter compared to other AHRS filters like the Kalman filter is the low computational cost in terms of mathematical operations with a simultaneously high accuracy. Furthermore, the Madgwick filter is effective at low sampling rates down to the order of 10 Hz and has a built in magnetic distortion compensation and gyroscope bias correction. [50]

Figure 3.23 shows the Euler angle distributions during a 17 min record of walking. The wobbling motion inherited from the walking style can be seen from the bi-modal distributions in roll and pitch angle (Figure 3.23a and 3.23b). The yaw angle distribution (Figure 3.23c) indicates the two walking directions, from point A to B and from point B to A.

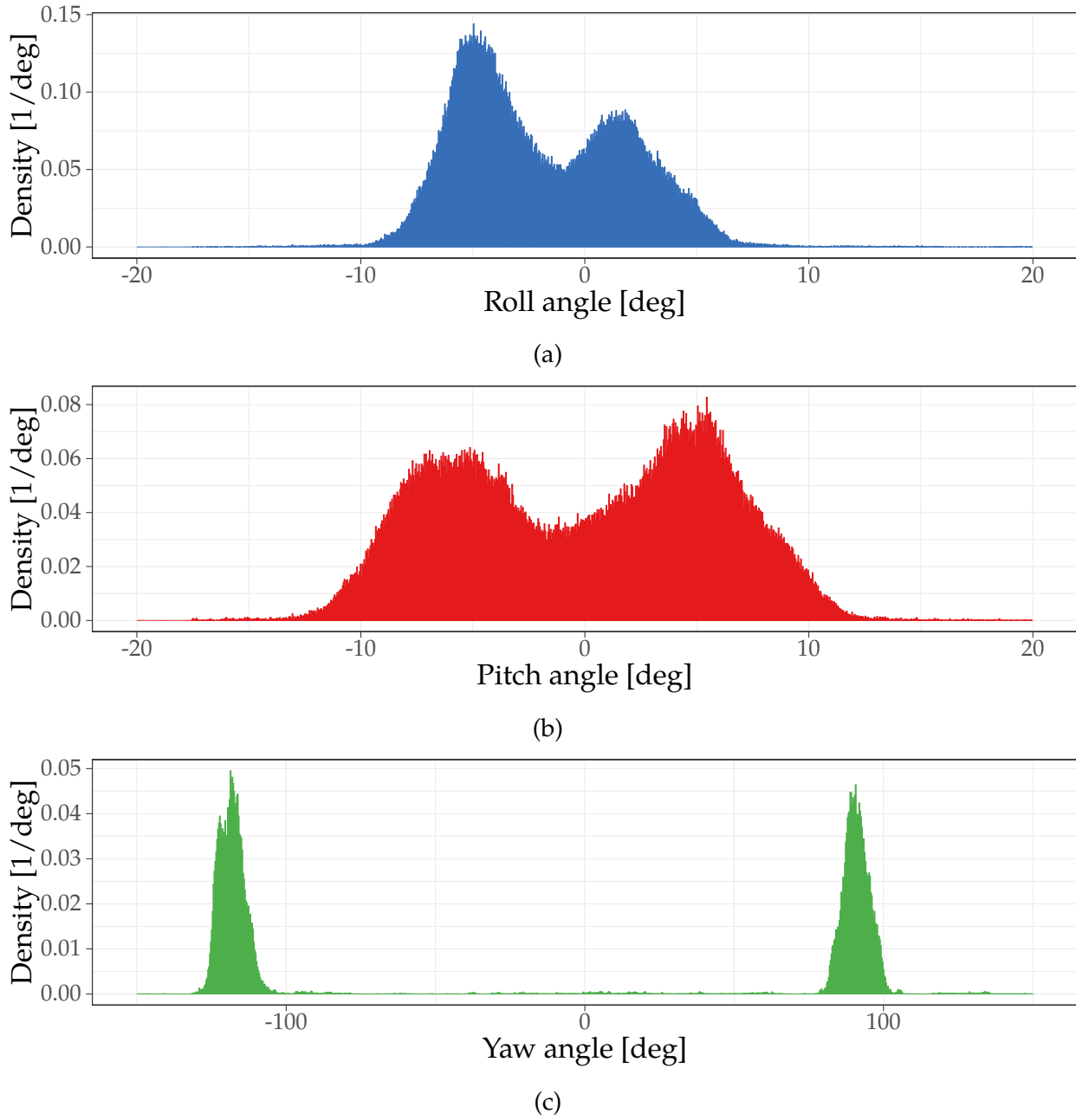


Figure 3.23: Normalised histograms for IMU orientation in Euler angles, retrieved from Madgwick filter

### 3D point cloud transformation

The LiDAR data was recorded as a set of 1521 range bearing measurements per sample,  $(\mathbf{r}, \boldsymbol{\theta}) = (r_i, \theta_i)$  with  $i \in [1, 2, \dots, 1521]$  in the laser frame. In the *laser* frame, the 3D point cloud  $\mathcal{P}_{\text{laser}}$  can be formulated as

$$\mathcal{P}_{\text{laser}} = \{\mathbf{p}_i = [x_i, y_i, z_i] \mid x_i = r_i \cos \theta_i, y_i = r_i \sin \theta_i, z_i = 0, i \in [1, 2, \dots, 1521]\}. \quad (3.5)$$

In order to transform the point cloud into the *base\_stabilised* frame, the orientation quaternion from the Madgwick filter  $q_{\text{ori}}$  was applied, together with the static transformation from

*base\_link* frame to the *laser* frame,  $q_{\text{link} \rightarrow \text{laser}}$ . Due to the yaw orientation while walking, stemming from the given road angle in the *map* frame, the yaw angle was corrected with an angle offset. The offset was determined manually to align the curb of the road, which could be observed in the LiDAR data, with the x-axis, by applying a constant rotation quaternion for the rectification of the point cloud,  $q_{\text{rect}}$ . The total transformation,  $q_{\text{trans}}$ , can be expressed as a product of quaternions,

$$q_{\text{trans}} = q_{\text{rect}} \cdot q_{\text{ori}} \cdot q_{\text{link} \rightarrow \text{laser}}. \quad (3.6)$$

### Road extraction, noise and ground line removal

The resulting point cloud contained a substantial amount of noise which is due to detections of the environment including mainly the vegetation next to the road curb and the road itself, the ground. The vegetation was first removed by limiting the rectified and transformed point cloud in y direction to keep only points which lie within 0 and 8 m. A larger distance than the actual road width was chosen to make sure to include all points which were detected on the road, also at larger distances in x direction. The point cloud was limited in z direction to include all points which were located above the ground by filtering out all points which lied below -0.4 m from *base\_link* origin. Equation 3.7 defines the resulting point cloud  $\mathcal{P}_{\text{road}}$ .

$$\mathcal{P}_{\text{road}} = \{\mathbf{p}_i \in \mathcal{P}_{\text{base}} \mid y_i \in [0, 8], z_i \geq -0.4, i \in [1, 2, \dots, 1501]\} \quad (3.7)$$

RANSAC was used on top of limiting the point cloud to detect lines in  $\mathcal{P}_{\text{road}}$ . Lines which had a slope corresponding to an absolute angle within the range [15, 75] degrees were marked as noise as they corresponded most likely to ground detection or the road curb or vegetation. Hence, all inliers of those lines were removed from  $\mathcal{P}_{\text{road}}$ .

### Vehicle cluster extraction

Point cloud clustering techniques are used to divide a point cloud which is unorganised into smaller parts, clusters. The aim is to either save processing time in comparison to using the whole point cloud or to keep track of clusters over time, like in this thesis. Euclidean clustering is an algorithm which extracts sub sets of points from a point cloud, called clusters. The clusters can be described as regions of nearest neighbours, hence sets of points which have a small euclidean distance between each other. [51]

In its standard implementation in PCL, the euclidean cluster extraction is based on a k-d tree representation of the input point cloud. K-d trees are binary search trees to associate points to clusters of nearest neighbours. The nodes of the search tree represent k-dimensional points which split the space along the k-th dimension into two half-spaces which are represented by two child nodes. The child nodes each split their corresponding spaces into half-spaces along the (k-1)-th dimension etc. until the leaves for the last dimension (k = 1) are reached. The splitting hyper plane is the plane which includes the node point and is perpendicular to the dimension represented by the node. The split points can be chosen for example as the medians of the corresponding dimension. Using the k-d search tree, new points can be sorted comparably fast, however the algorithm is approximate and does not guarantee that points in each neighbourhood are not closer to points in other neighbourhoods. [52]

In this case, a 2-d tree was used to estimate the nearest neighbourhoods. The tree was used as input to a euclidean cluster algorithm with a cluster tolerance of 1.3. The maximum

cluster size was set to 1521, i.e. the number of LiDAR points per sample, and the minimum cluster size to 30, in order to prevent that agglomerated noise clusters were classified as vehicle clusters.

## Vehicle tracking

Figure 3.24 shows the cluster shapes which typically occur when a vehicle overtakes the pedestrian. The phases (1)-(5) correspond to the defined overtaking phases, i.e. (1) approaching, (2) steering away, (3) passing and (4) returning. In (5), the overtaking manoeuvre is finished. When the car changes heading angle, i.e. in (2) and (4), the detected cluster is shaped like an “L”, containing detections from both front and side. In (1) and (5), the vehicle detections are shaped roughly like a vertical straight line corresponding to front and rear, in (3) like a straight line corresponding to the side.

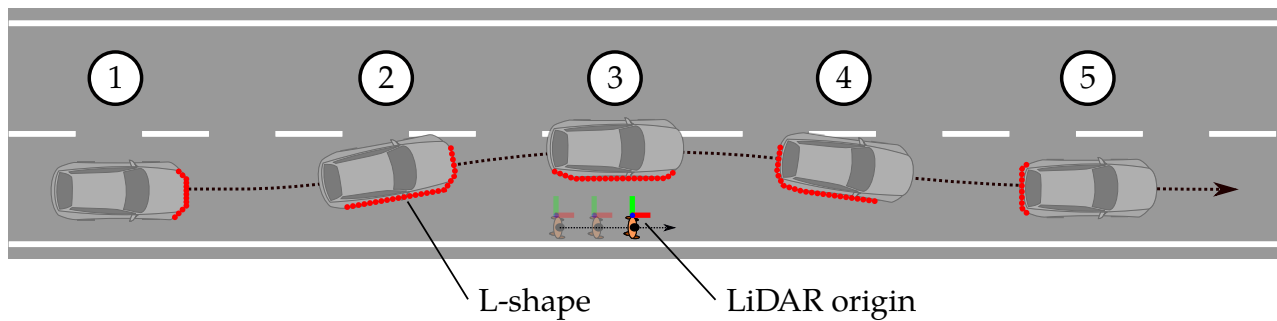


Figure 3.24: LiDAR point clusters corresponding to detections of an overtaking car at different phases of the overtaking event, recorded by the LiDAR mounted to the hip of the pedestrian. Approaching phase (1) produces a rather vertical line, while the steer away phase (2) causes an L-shape cluster. When passing (3), the cluster becomes a rather horizontal line, when steering back (4) an L-shape appears. At the end of the overtaking manoeuvre (5), the cluster appears as a vertical line

Figure 3.25 shows the L-shape cluster as an example of which extreme points are estimated from each cluster. These points consist of minimum and maximum x-coordinate,  $p_{\min-x}$  and  $p_{\max-x}$ , as well as minimum and maximum y-coordinate,  $p_{\min-y}$  and  $p_{\max-y}$ . The point  $p_{\text{inter}}$  marks the intersection between the line fits of front and side which is calculated from the RANSAC optimised line parameters.

To track vehicles over time for MC and speed estimation, the vehicles were assigned cluster IDs. Clusters which moved at a reasonable speed, estimated from the change in x-coordinate  $x_{\max-x}$  over a certain amount of time, were assigned the same ID. Oncoming traffic was identified from comparing the y-coordinate  $y_{\max-x}$  with a threshold of 3.5 m to decide whether the cluster was in the adjacent lane.

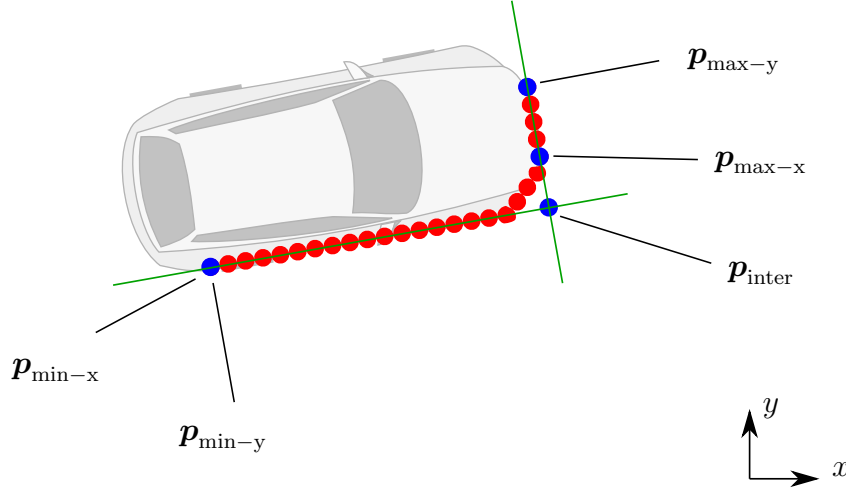


Figure 3.25: Detected points in cluster when L-shape is present (red), blue points mark the extreme points of the cluster used for size and velocity estimation. The extremity points are the points with minimum/maximum  $x$ - and  $y$ -coordinate as well as the intersection point of the perpendicular line fits for front and side.

### MC and speed estimation

MC was calculated from the minimum range measured among all occurrences of the corresponding cluster. The minimum range is thus a directly measured metric. It is equal to the MC under the assumption that the pitch angle does not have a significant influence on the actual clearance which was measured in the *laser* frame.

To estimate a more precise speed of the overtaking vehicle  $v_{\text{veh}}$ , the  $x$ -coordinate of the points with the minimum and maximum  $x$ -coordinate,  $p_{\min-x}$  and  $p_{\max-x}$ , corresponding to front and rear, respectively, were used. Using RANSAC, a linear line fit was applied to each of the two sets of points to estimate two constant relative speeds between the front or rear of the overtaking vehicle and the pedestrian, according to

$$(v_{\text{rel}}^{\text{F}}, N_{\text{in}}^{\text{F}}) \leftarrow \text{RANSAC}(x_{\min-x}(t), N_{\text{it}} = 1000, k = 2, \tau = 0.01, p_{\text{out}} = 0.2), \quad (3.8)$$

$$(v_{\text{rel}}^{\text{R}}, N_{\text{in}}^{\text{R}}) \leftarrow \text{RANSAC}(x_{\max-x}(t), N_{\text{it}} = 1000, k = 2, \tau = 0.01, p_{\text{out}} = 0.2), \quad (3.9)$$

where  $v_{\text{rel}}^{\text{F}}$  and  $v_{\text{rel}}^{\text{R}}$  are the relative speeds of the front and rear, respectively, equal to the slopes of the RANSAC estimated lines.  $N_{\text{in}}^{\text{F}}$  and  $N_{\text{in}}^{\text{R}}$  are the corresponding numbers of inliers. Using the ratio between the number of inliers of the front data fit and the rear data fit, a weighted average was then performed to extract the combined relative speed. The absolute speed of the vehicle  $v_{\text{veh}}$  was estimated by adding the mean GPS speed of the pedestrian  $\hat{v}_{\text{ped}}^{\text{GPS}}$  over one record to the combined relative speed.

$$v_{\text{veh}} \approx \frac{N_{\text{in}}^{\text{F}}}{N_{\text{in}}^{\text{F}} + N_{\text{in}}^{\text{R}}} v_{\text{rel}}^{\text{F}} + \frac{N_{\text{in}}^{\text{R}}}{N_{\text{in}}^{\text{F}} + N_{\text{in}}^{\text{R}}} v_{\text{rel}}^{\text{R}} + \hat{v}_{\text{ped}}^{\text{GPS}} \quad (3.10)$$

Figure 3.26 shows examples of the line fit for an overtaking vehicle (3.26a) and an oncoming vehicle (3.26b).

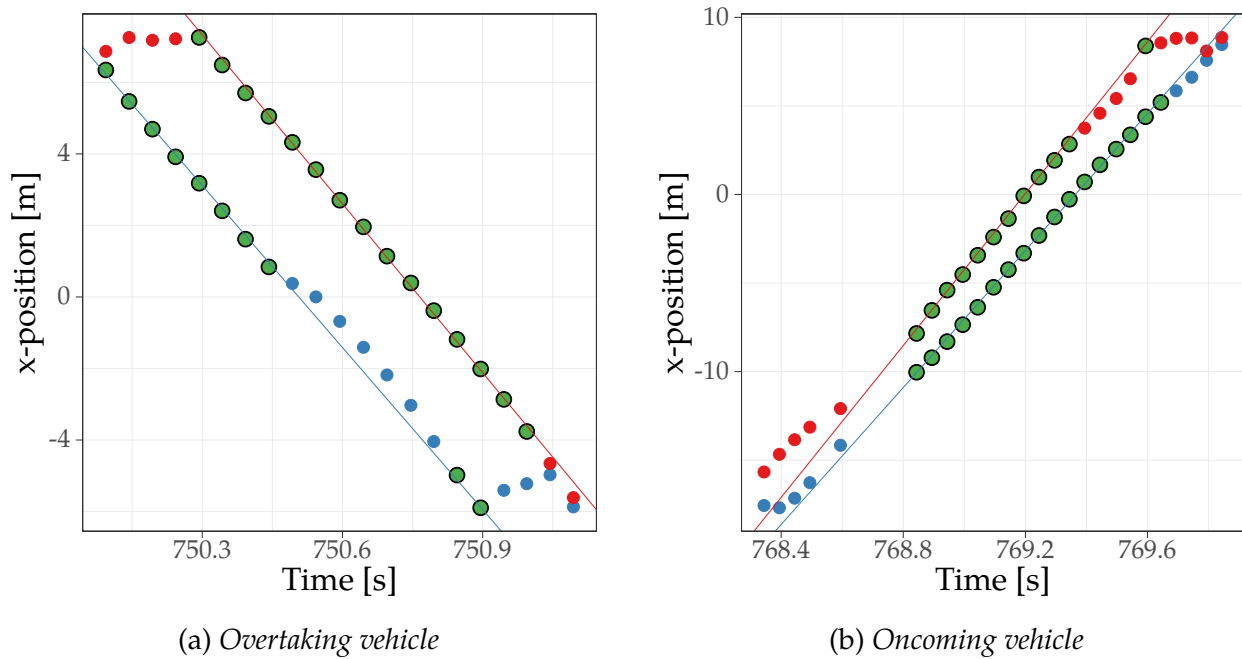


Figure 3.26: RANSAC line estimation for front (blue) and rear (red) detections for two different vehicles, inliers are marked in green

## 3.4 Bayesian modelling of overtaking metrics

There exist two main concepts in statistics of defining what probability is, the frequentist and the Bayesian approach. In frequentist statistics, unknown quantities are deterministic and probability is defined as a proportion of outcomes in a repeated test. In contrast, in Bayesian statistics, unknown quantities are modelled random and probability quantifies the likelihood of a certain outcome, i.e. a probability density function of the outcome conditioned on prior knowledge from e.g. data. In this thesis, a model is developed based on a Bayesian regression model in order to identify changes in overtaking metrics MC and TTC under certain conditions.

### 3.4.1 Bayesian statistics

The fundamental principle in Bayesian statistics is based on Bayes' rule to estimate the posterior density  $p(\theta|D)$  of parameters  $\theta$ , conditioned on the data  $D$ . In equation (3.11), Bayes' rule is shown. [53, 54]

$$p(\theta|D) = \frac{p(\theta) p(D|\theta)}{p(D)} \quad (3.11)$$

In (3.11),  $p(\theta)$  is the prior density of the parameters and  $p(D|\theta)$  is called the likelihood of the data.  $p(D)$  is the prior density of the data which is defined as

$$p(D) = \int_{\theta} p(D|\theta) p(\theta) d\theta, \quad (3.12)$$

and assumed to be a constant normalisation factor. Hence, the posterior density can be expressed as proportional to the prior of the parameters  $p(\theta)$  and the likelihood of the data  $p(D|\theta)$ .

$$p(\theta|D) \propto p(\theta) p(D|\theta) \quad (3.13)$$

### 3.4.2 Bayesian regression model

Bayesian regression models (BRMs) aim at approximating the posterior density of a model with certain parameters conditioned on data. BRMs can be designed with linear and non-linear distributions. As general statistical models, BRMs can be of multilevel character, meaning that there is a nested structure or hierarchy of populations in the available data set. An example is a model which involves students who are each within different classes, within different schools within different cities, i.e. different levels. Such a model is referred to as multilevel model (MLM). [54, 55]

Within the Bayesian framework, a linear regression model is expressed in a probabilistic way which means that model parameters are sampled from probability densities. The  $i$ -th of  $N$  responses in the response vector  $\mathbf{y}$  of a linear BRM,  $y_i$ , is assumed to be distributed according to

$$y_i \sim \mathcal{D}(f(\eta_i), \theta), \quad i = 1, \dots, N \quad (3.14)$$

where  $D$  is an arbitrary statistical probability function with a set of parameters  $\theta$  and the  $i$ -th element of the linear predictor  $\eta_i$ . The linear predictor is for each response  $i$  composed of  $p$  fixed effects, contained in  $X_i$ , and  $q$  random effects, contained in  $Z_i$ , weighted by the model parameters  $\beta$  and  $u$ , respectively.  $X$  and  $Z$  are also referred to as population level and group level effects, respectively. [55]

$$\eta_i = X_{i,1}\beta_1 + \dots + X_{i,p}\beta_p + Z_{i,1}u_1 + \dots + Z_{i,q}u_q \quad (3.15)$$

The full response description can be expressed with fixed effect matrix  $\mathbf{X}$  and random effect matrix  $\mathbf{Z}$ ,

$$\eta = \mathbf{X}\beta + \mathbf{Z}u, \quad \mathbf{X} \in \mathbb{R}^{n \times p}, \mathbf{Z} \in \mathbb{R}^{n \times q} \quad (3.16)$$

In a BRM, the model parameters are, in contrast to other regression models, assumed to be sampled from distributions which are characterised by a mean and a lower and upper credible interval (CI). Credible intervals can be roughly expressed as the pendant to confidence intervals in frequentist statistics but in contrast tell between which bounds a random variable is sampled with a certain probability. Having parameters of the model expressed as densities represents an advantage in contrast to frequentist methods like linear mixed effect models because one can quantify the uncertainty in parameters with various distribution models, whereas in linear mixed effect models only a single point estimate is determined. [54]

To estimate BRM parameter distributions as well as the distribution of the response variable, modern solvers make use of so called Markov chain Monte Carlo (MCMC) algorithms. MCMC algorithms aim at approximating the posterior density using [54, 55]

### 3.4.3 Implementation in R

The BRM was fitted in the statistical programming language R using the package *brms*. The package *brms* is developed in the statistical modelling language Stan which allows performing Bayesian inference and includes implementations of different MCMC sampling algorithms [56]. *brms* contains tools for fitting MLM with a variety of distribution families. The package contains methods for fitting, prediction and comparison of MLM and allows users to set custom prior densities in case of available knowledge about parameters [55].

The factors oncoming traffic, pedestrian direction and pedestrian lateral position, as well as the speed are used for the models since they were found to be the main influencing factors from a first qualitative analysis in frequentist perspective, described in the first of the results chapter.

#### Model of minimum clearance

Since the distribution of MC was assumed to have a clear minimum at zero, meaning that drivers are assumed to leave an  $MC > 0$  to the pedestrian, an asymmetric distribution family was chosen, in this case a log-normal distribution with link function identity. The  $i$ -th response for one of the data sets  $y_{MC,j,i}$  is modelled as

$$y_{MC,j,i} \sim \log \mathcal{N}(\eta_{j,i}, \sigma_{MC,j}^2), \quad \eta_{j,i} = \mathbf{X}_{MC,j,i} \beta_{MC} + \mathbf{Z}_{MC,j,i} u_{MC}, \quad (3.17)$$

where  $j \in \{\text{PDL}, \text{NDS}\}$  is the index of the data set.

The fixed parameters  $\beta$  and effects  $\mathbf{X}$  for the MC model are for both data sets defined as

$$\begin{aligned} \mathbf{X}_{MC,j,i} &= \begin{bmatrix} 1 & X_{\text{onc},j,i} & X_{\text{dir},j,i} & X_{\text{lat},j,i} & X_{\text{onc} * \text{dir},j,i} & X_{v,j,i} \end{bmatrix} \\ \beta_{MC}^T &= \begin{bmatrix} \beta_0 & \beta_{\text{onc}} & \beta_{\text{dir}} & \beta_{\text{lat}} & \beta_{\text{onc} * \text{dir}} & \beta_v \end{bmatrix}, \end{aligned} \quad (3.18)$$

where  $\beta_0$  is the parameter of the intercept, i.e. effect  $X_{0,i} = 1$ . The factor oncoming traffic is represented by  $\beta_{\text{onc},i}$  with effect  $X_{\text{onc},i}$ , the factor pedestrian direction represented by the parameter  $\beta_{\text{dir},i}$  with effect  $X_{\text{dir},i}$  and the pedestrian lateral position with  $\beta_{\text{lat},i}$  and  $X_{\text{lat},i}$ .  $\beta_{\text{onc} * \text{dir},i}$  and  $X_{\text{onc} * \text{dir},i}$  are parameter and effect, respectively, of the interaction between oncoming traffic and pedestrian direction.  $\beta_v$  is the parameter of the vehicle speed effect  $X_{v,i}$ .

The linear BRM used for modelling the response variable MC from the field test data is designed as described in (3.19).

$$y_{MC,\text{PDL},i} \sim \log \mathcal{N}(\mathbf{X}_{MC,\text{PDL},i} \beta_{MC}, \sigma_{MC,\text{PDL}}^2), \quad (3.19)$$

where  $\sigma_{MC}^2$  is the variance of the log-normal distribution.

The distribution for the UDRIVE data is shown in equation 3.20. Since the data set contains multiple events from same drivers, the driver ID was set as a random effect.

$$y_{MC,\text{NDS},i} \sim \log \mathcal{N}(\mathbf{X}_{MC,\text{NDS},i} \beta_{MC} + \mathbf{Z}_{MC,\text{NDS},i} u_{MC}, \sigma_{MC,\text{NDS}}^2), \quad (3.20)$$

with the random effect and parameter, respectively, being defined as

$$\begin{aligned} \mathbf{Z}_{MC,\text{NDS},i} &= \mathbf{Z}_{\text{ID},i} \\ u_{MC} &= u_{\text{ID}}, \end{aligned} \quad (3.21)$$



where  $u_{ID} \sim \mathcal{N}(0, \sigma_{ID}^2)$ .

Both models are fit in *brms* with 4 MCMC chains, each with a maximum number of 2000 iterations out of which the first 1000 are used for sampler warm-up. Non-informative default prior distributions are set due to a lack of prior knowledge about the parameters, those are student's t-distributions with a larger variance. For the model from UDRIVE data, an adaptive step size parameter of 0.95 is chosen for finer sampling in the MCMC algorithm in order to avoid diverging transitions. Figure 3.27 illustrates a Kruschke diagram of the linear BRM used for the MC model, including the linear predictor, priors and posterior of the response variable.

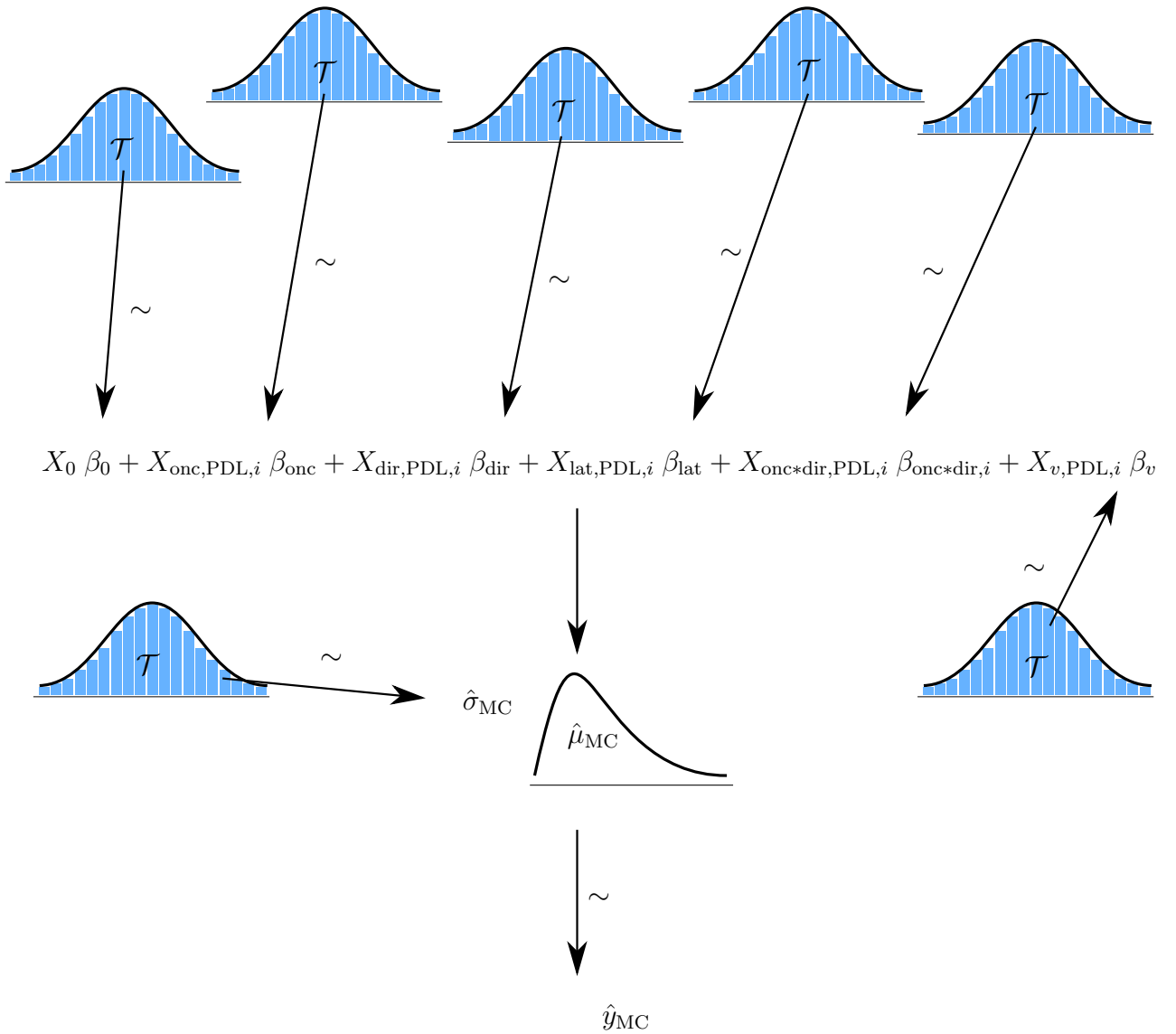


Figure 3.27: Kruschke diagram, adopted from [54], of used BRM for MC

### Model of time to collision at steer away

TTC at steer away was modelled from the UDRIVE data set with a log-normal distribution, assuming as for MC that drivers generally stay above  $TTC = 0$ . The parameters were taken the same as for the MC model. The  $i$ -th response is

$$y_{TTC,i} \sim \log \mathcal{N} \left( \mathbf{X}_{TTC,i} \beta_{MC} + Z_{ID,i} u_{ID}, \sigma_{TTC}^2 \right), \quad (3.22)$$

with  $\mathbf{X}_{TTC,i}$  being the TTC fixed effect matrix and  $u_{ID} \sim \mathcal{N}(0, \sigma_{ID}^2)$ .

### Model of speed next to pedestrian

The speed of the vehicle next to the pedestrian  $v_{veh}$ , i.e. the approximated passing phase speed, is modelled from field test data with a student's t-distribution  $\mathcal{T}$ , assuming that the speed distribution among drivers is approximately normally distributed with outliers far above and far below the speed limit. For the field test data, the model is assumed as

$$y_{v,i} \sim \mathcal{T} \left( X_{v,i} \beta_v, \nu_v, \sigma_v^2 \right), \quad (3.23)$$

where  $\nu_v$  is the number of freedoms.

The parameters and effects were set similarly to the ones used for the MC model, however excluding the vehicle speed,

$$\begin{aligned} \mathbf{X}_{v,i} &= \begin{bmatrix} 1 & X_{onc,i} & X_{dir,i} & X_{lat,i} & X_{onc * dir,i} \end{bmatrix}, \\ \beta_v^T &= \begin{bmatrix} \beta_0 & \beta_{onc} & \beta_{dir} & \beta_{lat} & \beta_{onc * dir} \end{bmatrix}. \end{aligned} \quad (3.24)$$

# 4

## Results

The following sections present the results from the analysis of the two data sets. Trends and uncertainties for different parameter constellations are expressed from a Bayesian regression model.

### 4.1 Results from UDRIVE data

In this section, the CZB results obtained from the extracted UDRIVE events are presented. Firstly, general observations from looking at the available signals as annotator are mentioned. Secondly, the quantitative trends are presented which are product of the explained methodology. The data used to generate the models excludes cases in which an overtaking car was piggy backing.

#### 4.1.1 Qualitative observations

Throughout the data extraction and annotation period, several observations were made after viewing the UDRIVE data. Generally, it was observed that drivers tend to pass pedestrians rather than overtake, meaning that drivers do not apply much steering input especially during flying overtaking manoeuvres. If the visibility of the road permits and no oncoming traffic is present, drivers slightly steer away long time ahead of reaching the pedestrian. Furthermore, drivers seemed to generally apply the steering wheel input in a continuous motion throughout the whole manoeuvre. This means that as soon as drivers evaded from the collision path, they steered back which gave rise to the assumption that there is no passing phase as being defined in previous studies [2, 6]. The time of steering back seemed to happen already before reaching the pedestrian. After each completed overtaking manoeuvre, drivers looked at the pedestrian through the central rear view mirror or a side view mirror. Drivers who were piggy backing were observed to imitate the lead vehicle's trajectory, often even though there was oncoming traffic present during the event.

#### 4.1.2 Temporal results

##### Overtaking phases duration

The overtaking phases duration is shown in form of box plot diagrams in Figure 4.3. Figure 4.1a shows the duration of the four phases for accelerative and Figure 4.1b the phases for flying overtaking strategy. It can be seen that with the accelerating strategy, drivers remain for a longer time in the approaching phase (1). For flying manoeuvres, the spread of the returning phase (4) is larger than for accelerative manoeuvres. In both manoeuvres, the passing phase (3) is significantly shorter than all other phases with the median being zero. The returning phase is significantly longer than all other phases.

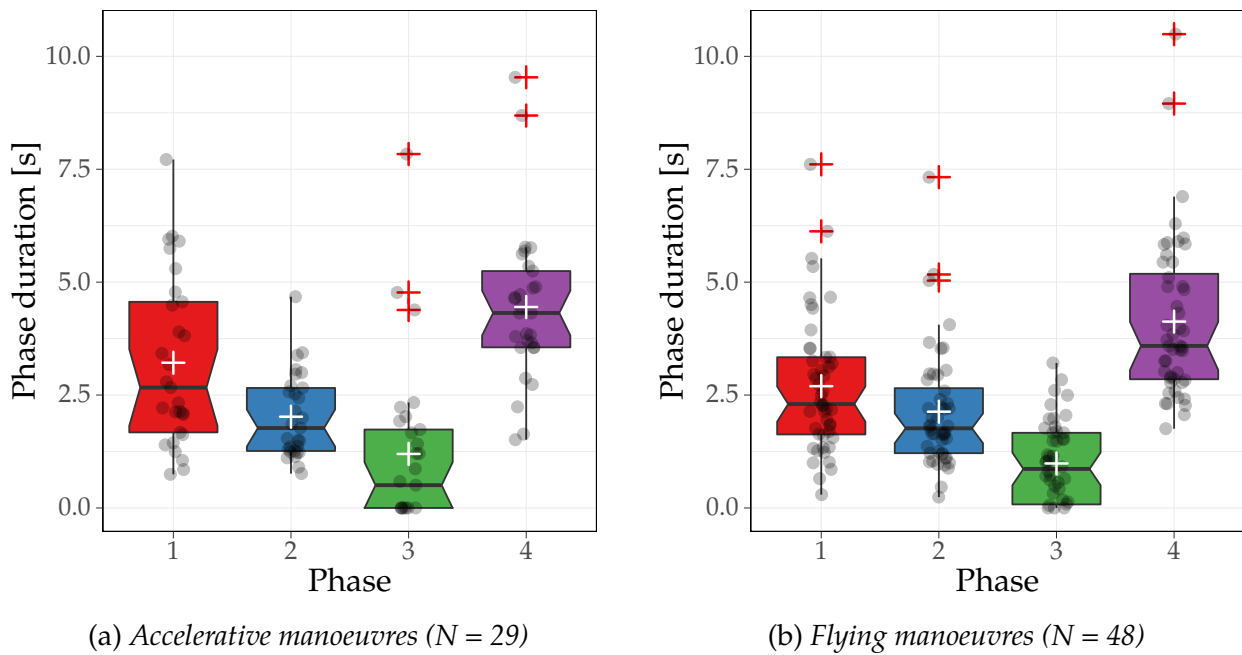


Figure 4.1: Box plot diagrams showing the phases duration, outliers marked with red plus symbols, mean marked with white plus

The quantitative results for the overtaking phases duration are summarised in Table 4.1.

		Phase 1 [s]	Phase 2 [s]	Phase 3 [s]	Phase 4 [s]	Total
Accelerative	mean	3.22	2.02	1.20	4.45	10.88
	std	1.88	0.94	1.81	1.73	2.95
Flying	mean	2.70	2.13	0.99	4.13	9.94
	std	1.51	1.32	0.89	1.77	2.96

Table 4.1: Overtaking phases duration for accelerating and flying manoeuvres, mean and standard deviation (std)

### Time to collision when steering away

In Figure 4.2, the empirical cumulative distribution function (ECDF) of the TTC at start of steering away is presented for the two factors oncoming traffic (Figure 4.2a) and pedestrian direction (Figure 4.2b). Only flying overtaking manoeuvres are included in the analysis of TTC to prevent braking effects from influencing the TTC at steer away. In the set of flying overtaking events, one driver had a TTC below the Euro NCAP threshold for FCW triggering of 1.7 s, corresponding to a share of 4%. In case of oncoming traffic present as well as pedestrian direction opposite, a tendency towards greater values of TTC is visible.

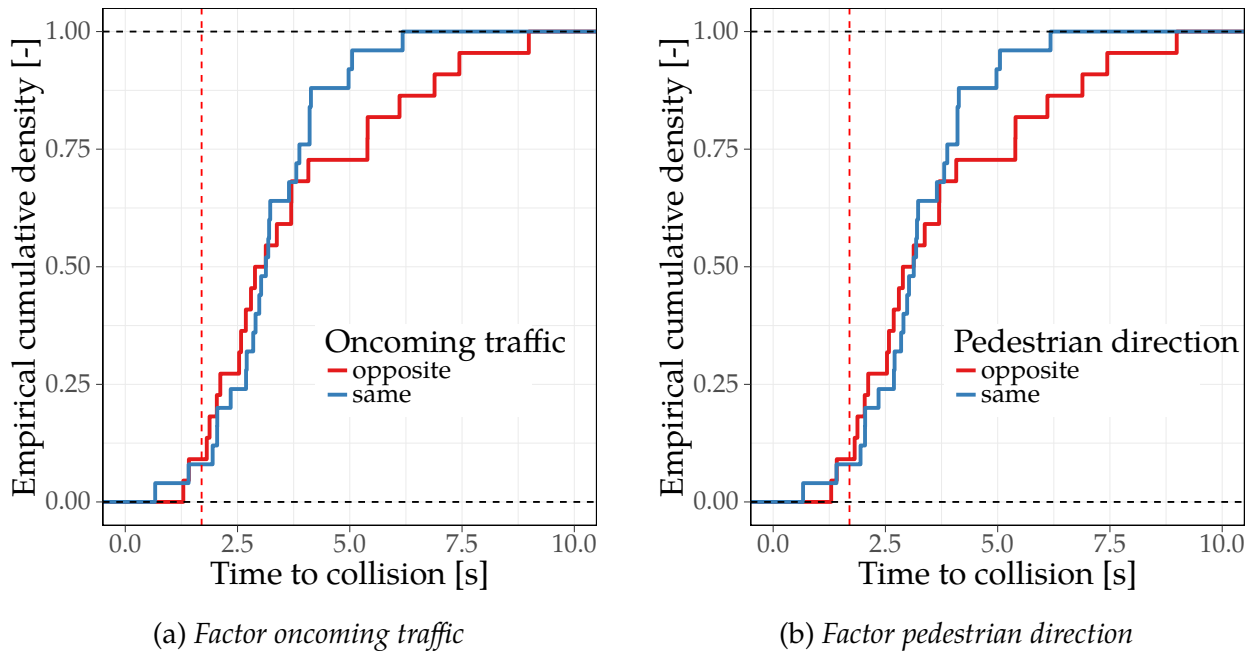


Figure 4.2: Empirical cumulative density function of time to collision for factor oncoming traffic and pedestrian direction

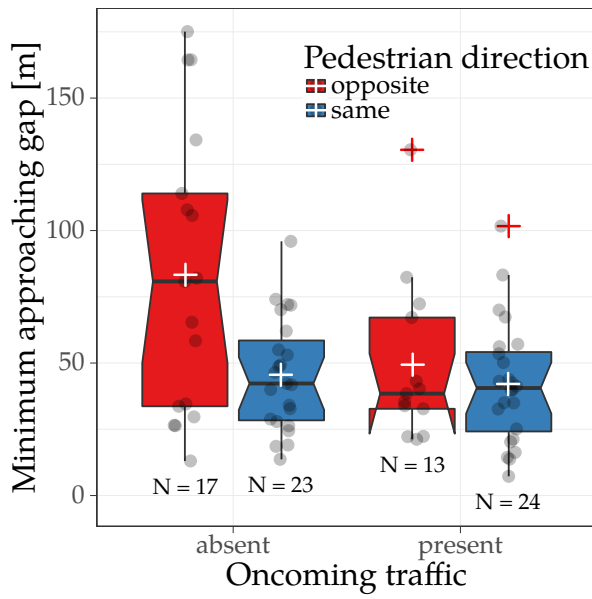
In Table 4.2, the values for TTC are shown under different factor variations for both overtaking strategies. The overall mean value for TTC is 3.61 s (1.74 s standard deviation).

		Oncoming traffic		Pedestrian direction		Total
		absent	present	opposite	same	
Flying	N [-]	12	11	11	14	48
	mean [s]	4.46	3.32	3.89	3.13	3.68
	std [s]	2.47	0.99	3.55	1.38	2.26

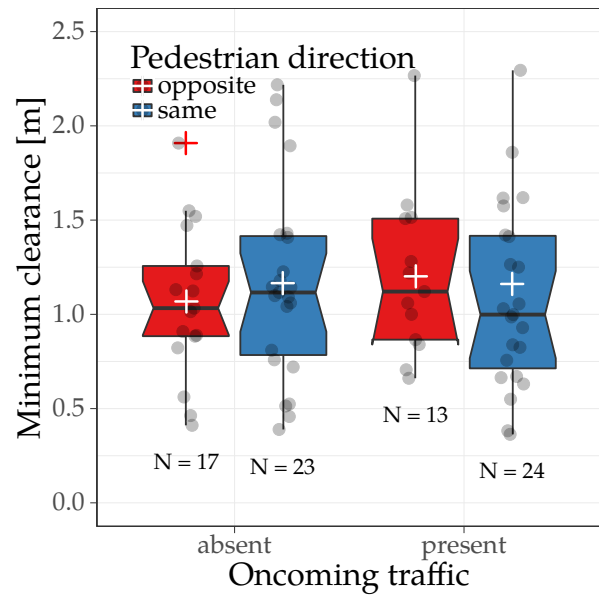
Table 4.2: Time to collision at steer away moment for accelerating and flying manoeuvres, number of observations (N), mean and standard deviation (std)

### 4.1.3 Spatial results

The spatial CZBs estimated for the UDRIVE data consists of MAG and MC. These metrics are evaluated for the factors oncoming traffic and pedestrian direction, as well as for both overtaking strategies accelerative and flying. Figure 4.3a shows the box plot diagrams of the distributions of MAG dependent on the factors oncoming traffic and pedestrian direction. From the box plots it is not possible to state any significance. However, it appears that the MAG is larger if oncoming traffic is absent and the pedestrian is walking in opposite direction to the traffic. For MC (Figure 4.3a), there is neither a trend nor any significant change observable.



(a) MAG results



(b) MC results

Figure 4.3: Box plot diagrams showing the MAG and MC distributions depending on the factors oncoming traffic and pedestrian direction, outliers marked with red plus symbols, mean marked with white plus

The ECDF of MC is shown in Figure 4.4 for factor oncoming traffic (Figure 4.4a) and factor pedestrian direction (Figure 4.4b). Both distributions for each factor are similarly shaped. It can be stated that about 75% of the drivers stayed below 1.5 m which is a minimum distance set by law in several European countries [2].

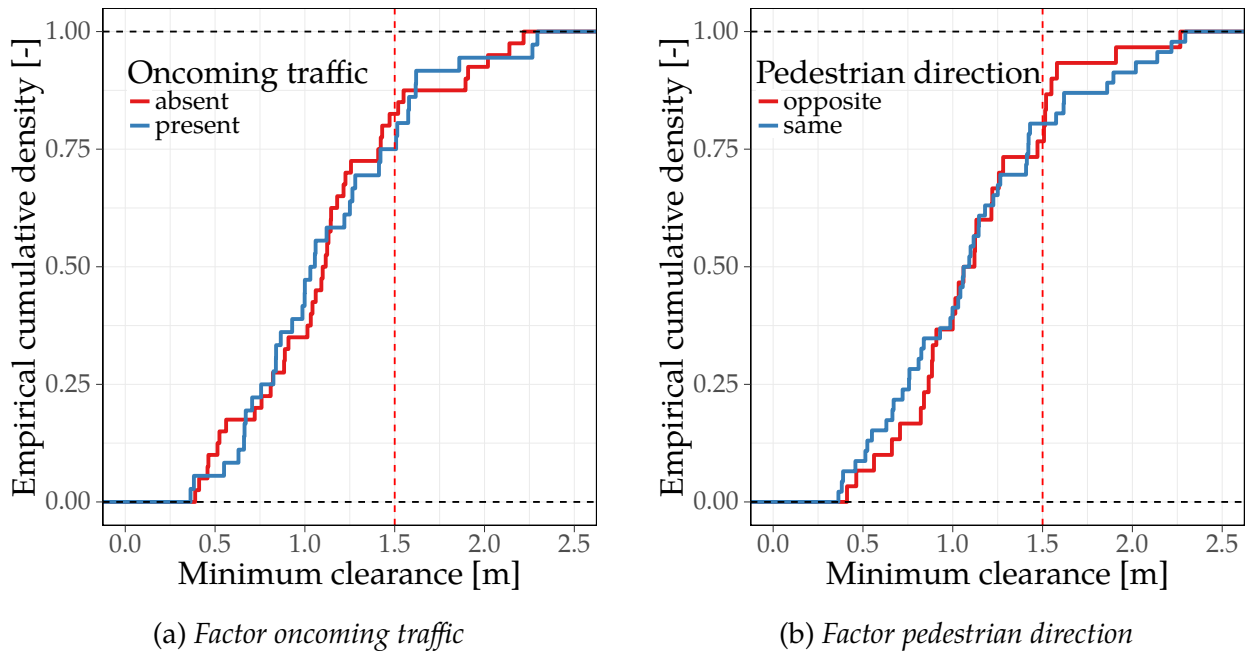


Figure 4.4: Empirical cumulative density function of minimum clearance for factor oncoming traffic and pedestrian direction, the red dashed line marks 1.5 m distance which is included in most of the European countries' policies as a minimum distance [2]

In Table 4.3, the mean and standard deviation values of MC are summarised. The overall mean value is 0.97 m (standard deviation 0.41 m).

		Oncoming traffic		Pedestrian direction		
		absent	present	opposite	same	Total
Accelerative	N [-]	5	12	2	10	29
	mean [m]	0.87	1.33	0.98	0.99	1.11
	std [m]	0.47	0.53	0.20	0.36	0.47
Flying	N [-]	12	11	11	14	48
	mean [m]	1.15	0.99	1.24	1.28	1.17
	std [m]	0.36	0.47	0.47	0.72	0.53

Table 4.3: Minimum clearance from UDRIVE events for accelerating and flying manoeuvres, number of observations (N), mean and standard deviation (std)

#### 4.1.4 Spatio-temporal results

In Figure 4.5, vehicle speed (4.5a) and vehicle longitudinal acceleration (4.5b) when the driver is passing the pedestrian is shown in form of box plots for the interaction of oncoming traffic and pedestrian direction. There is a trend observable that with oncoming traffic being present, the speed is reduced when the pedestrian is walking opposite to the traffic. However, it must be taken into account that the speed limits on the corresponding roads are not necessarily

the same which might have had an influence on the passing speed. Acceleration is more negative and has a larger variance when the pedestrian is walking in the same direction as traffic.

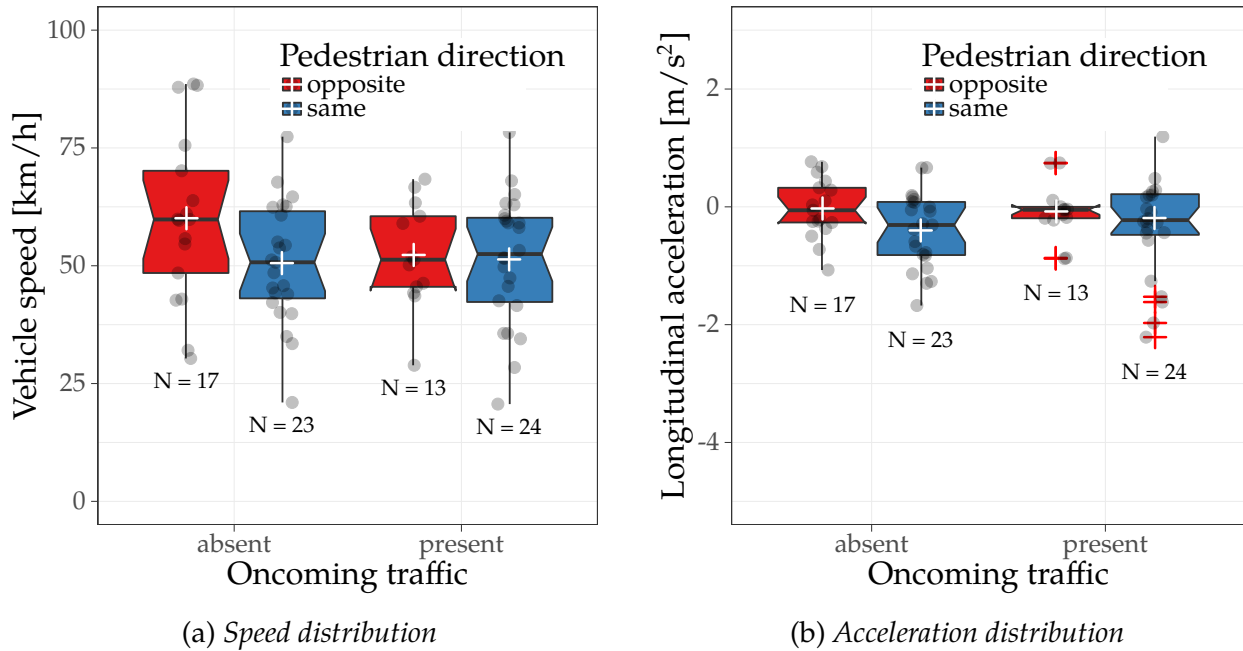


Figure 4.5: Box plot diagrams showing vehicle speed and longitudinal acceleration when being next to the pedestrian, including interaction of factors oncoming traffic and pedestrian direction

In Figure 4.6, the speed change during the four phases of the overtaking manoeuvre is shown for accelerative (Figure 4.6a) and flying (Figure 4.6b) manoeuvres. For accelerative manoeuvres, the speed change between in the approaching phase is -10.73 km/h (14.27 km/h standard deviation) and 9.36 km/h (6.81 km/h standard deviation) in the return phase. For flying manoeuvres, the speed change mean and median is located around zero with the largest deviations in approaching and returning phase.



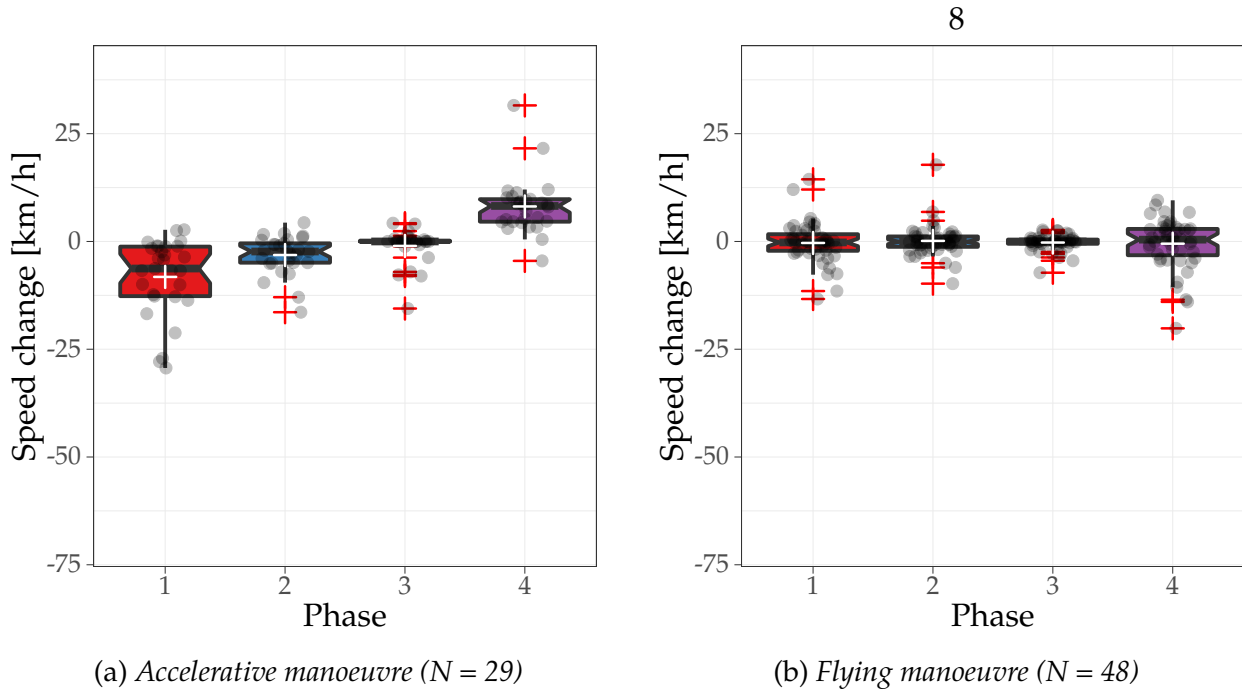


Figure 4.6: Box plot diagrams showing the speed change during the four different overtaking phases, for accelerative and flying manoeuvres

Figure 4.7 shows scatter plots for TTC (Figure 4.7a) and MC (Figure 4.7b) over speed for accelerative and flying overtaking manoeuvres. The lines show the trends approximated by a linear model as well as the 95% confidence interval bands. For flying manoeuvre strategy, the TTC increases with speed. For flying manoeuvres, also MC has a trend to increase with speed. However, for accelerative manoeuvres, there is no clear trend recognizable.

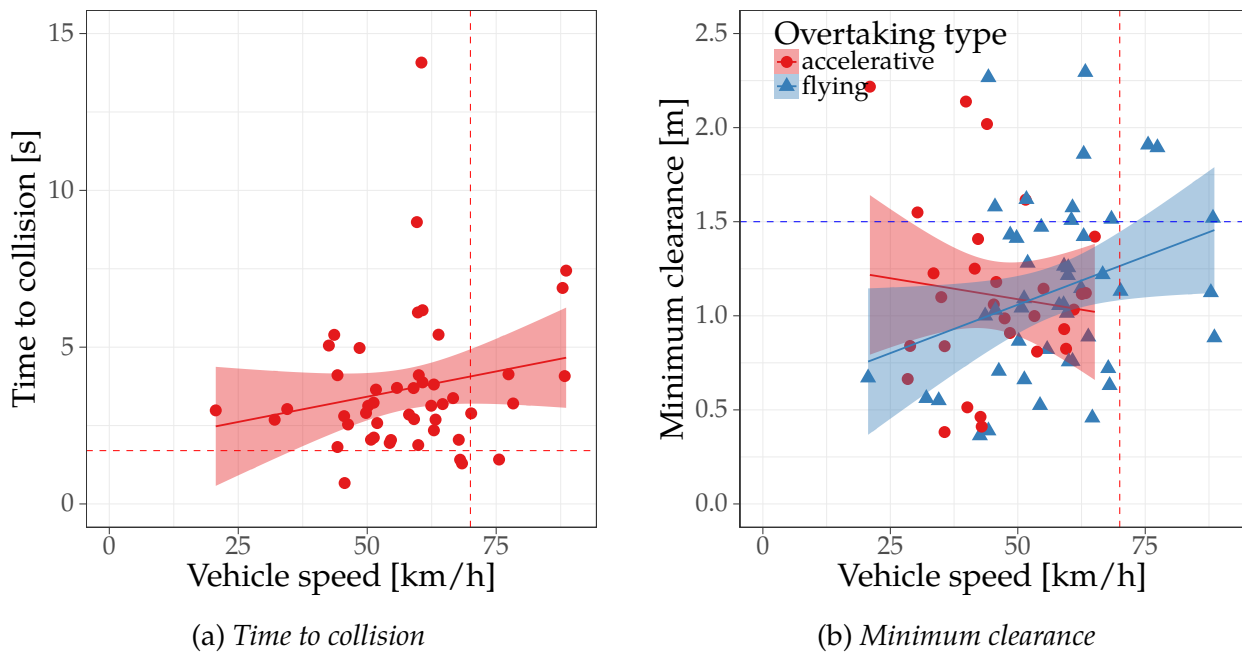


Figure 4.7: Scatter plots of time to collision (only flying overtaking manoeuvres) and minimum clearance over speed for accelerative and flying overtaking strategy

## 4.2 Results from field test data

The following section includes the results from the field data recorded with the PDL after tracking the vehicles and extracting the metrics MC and speed as well as the factors overtaking traffic and overtaking type. The data used for the results excludes vehicles which are classified as piggy backers as well as vehicles with a detected length greater than 5 m in order to include only passenger cars.

### 4.2.1 Spatial results

#### Minimum clearance

Figure 4.8 shows the MC distribution for the factors oncoming traffic and pedestrian direction in both scenarios, line (Figure 4.8a) and curb (Figure 4.8b). It can be seen that in case oncoming traffic is present during the overtaking, MC decreases in terms of mean and median. The notches around the medians of both box plot diagrams only overlap each other when the pedestrian is walking in the same direction, hence it can be assumed that there is a significant difference in median values with a confidence of 95%. If a pedestrian is facing the traffic while walking (factor value opposite), MC generally decreases in comparison to walking in the same direction as the traffic, independently of the oncoming traffic. This trend is more evident for the line than the curb scenario.

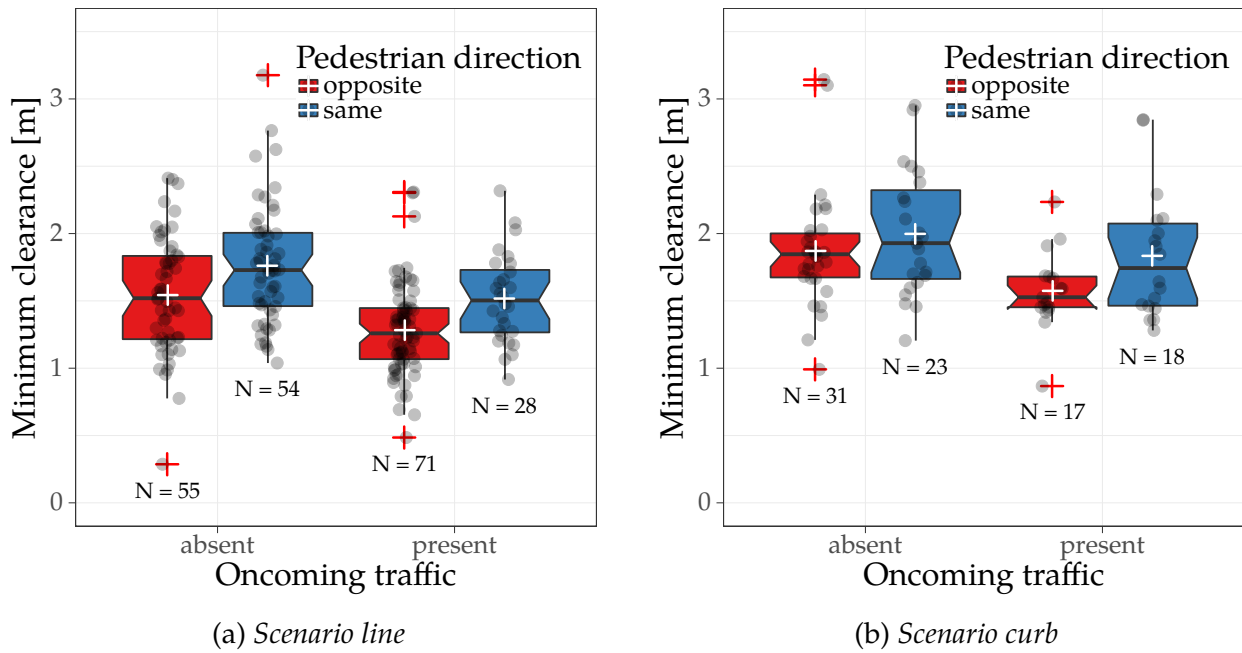


Figure 4.8: Box plot diagrams for minimum clearance distribution in field test data, line and curb scenario in dependence of factors oncoming traffic and pedestrian direction. Means are marked with white plus symbols, outliers marked with red plus symbols.

In Figure 4.9, the ECDF of the MC is shown for factors oncoming traffic and pedestrian direction. From the plot it can be concluded that when the pedestrian is walking opposite to the traffic and oncoming traffic is present, the MC distribution is shifted towards lower MC values. Almost 75% of the drivers are willing to pass the pedestrian under 1.5 m in that condition. Cases pedestrian direction opposite and oncoming traffic absent as well as pedestrian direction same and oncoming traffic present yield similar MC distributions. In the condition with the pedestrian walking in the same direction and oncoming traffic is absent, drivers leave the most distance to the pedestrian.

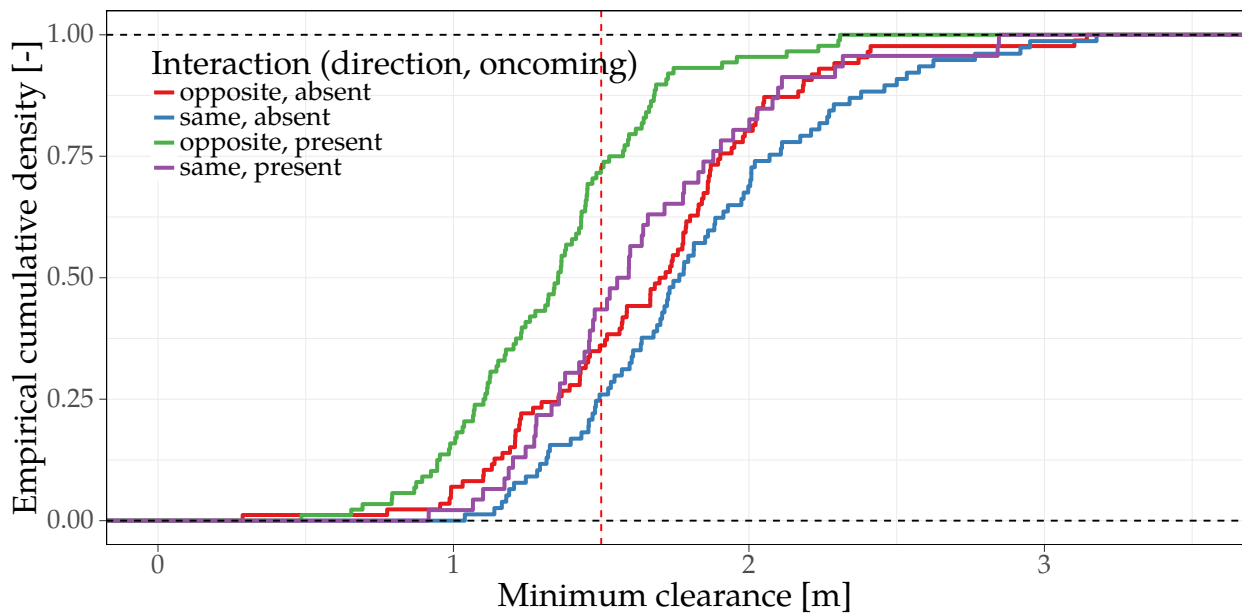


Figure 4.9: Empirical cumulative density function of minimum clearance under influence of the interaction between the factors oncoming traffic and pedestrian direction, red dashed line symbolises minimum distance proposed by other European countries

Table A.1 and A.2 in the appendix show the MC values for the line and curb scenario, respectively.

### Lateral position of overtaking cars

In Figure 4.10, the distribution of the lateral position of the centre of the vehicles is shown, for line (Figure 4.10a) and curb (Figure 4.10b) scenario, estimated from MC and the detected cluster width according to

$$y_{\text{veh}} = \text{MC} + \frac{W_{\text{veh}}}{2}. \quad (4.1)$$

In (4.1), which represents the line scenario,  $y_{\text{veh}}$  is the lateral position of the vehicle and  $W_{\text{veh}}$  the width of the vehicle. In case oncoming traffic is absent and the pedestrian is walking in traffic direction, drivers tend to drive further away from the pedestrian and partly evade into the adjacent lane. Since the notches do not overlap, it can be concluded with 95% confidence that drivers pass the pedestrian further away if the pedestrian is walking in the same direction than if the pedestrian is facing the traffic. If oncoming traffic is present, the MC distributions are less wide and tend to be centred more towards the middle of the lane. Again, the pedestrian walking direction influences the lateral position, as drivers evade further towards the adjacent lane when the pedestrian is walking in the same direction (95% confidence of difference in medians). It can be further stated that drivers tend to be less willing to evade to the adjacent lane in case of oncoming traffic.

The calculation of the lateral position in the curb scenario takes into account the  $W_{\text{off}} = 0.5$  m offset which the pedestrian maintained during the data collection from the lane marking

$$W_{\text{off}} = 0.5 \text{ m},$$

$$y_{\text{veh}} = \text{MC} + \frac{W_{\text{veh}}}{2} - W_{\text{off}}. \quad (4.2)$$

A similar trend of the lateral position as in the line scenario can be observed, i.e. when oncoming traffic is present during the overtaking manoeuvre, drivers tend to stay more in the centre of the lane and evade less from their path towards the adjacent lane. However, in contrast to the line scenario, none of the trends can be verified with statistical significance.

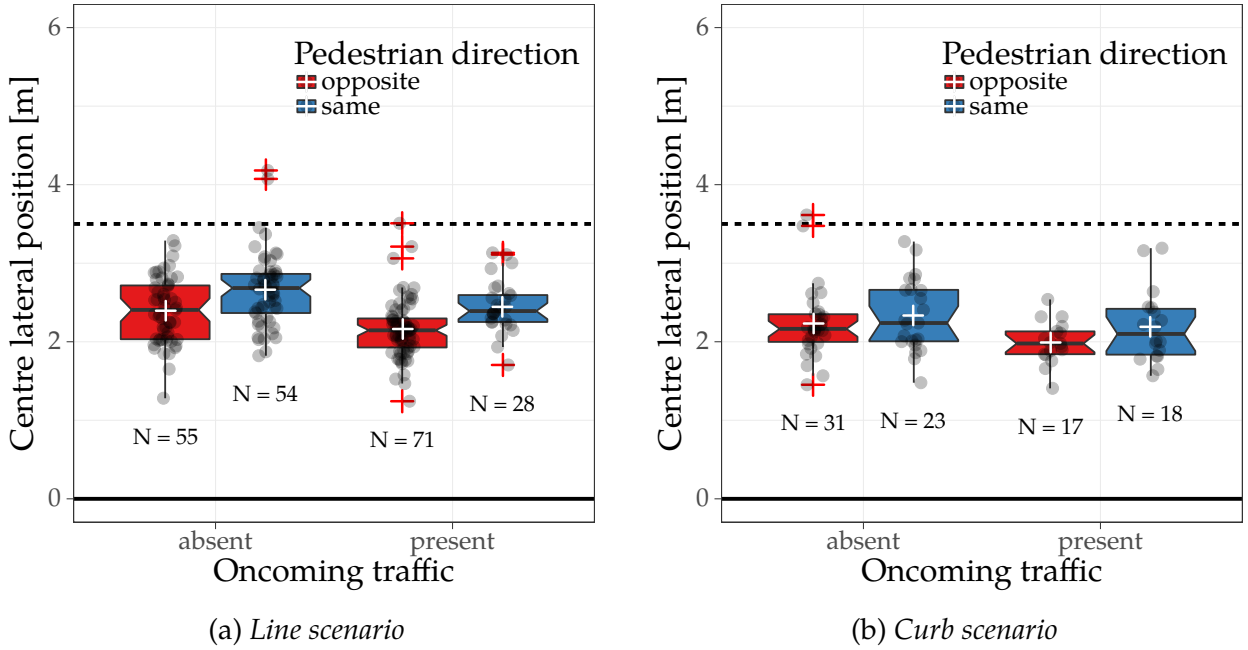


Figure 4.10: Box plot diagrams for the lateral position of the vehicle centre for walking on the line scenario, interaction between parameters oncoming traffic and pedestrian direction

Figure 4.11 shows results for the vehicle outer lateral position in the line (Figure 4.11a) and curb (Figure 4.11b) scenario. The corresponding equations for the vehicle outer lateral position, i.e. the left side, for line and curb scenario are for the line scenario

$$y_{\text{veh}} = \text{MC} + W_{\text{veh}}, \quad (4.3)$$

and for the curb scenario

$$y_{\text{veh}} = \text{MC} + W_{\text{veh}} - W_{\text{off}}, \quad (4.4)$$

respectively.

The distributions of the vehicle outer lateral position confirm findings for the centre lateral position. When the pedestrian was walking in the same direction as traffic, drivers evaded more into the adjacent lane than in case of opposite direction. In the curb scenario, no significant differences can be observed among factors oncoming traffic and pedestrian direction.

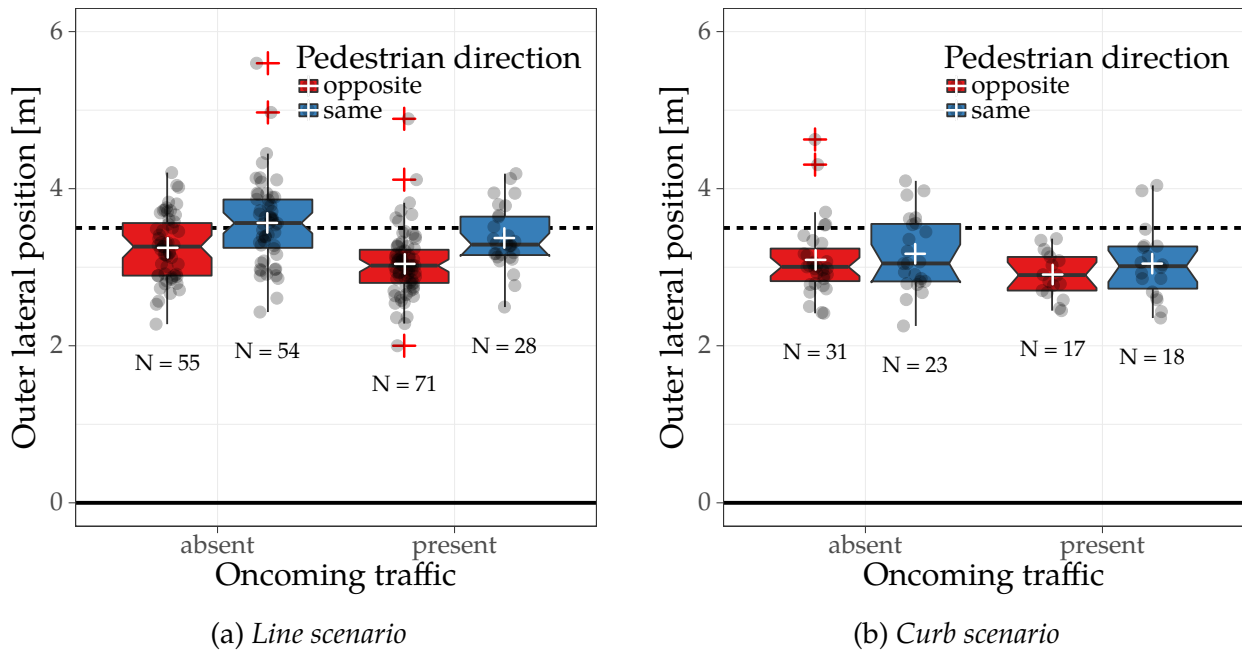


Figure 4.11: Box plot diagrams for the lateral position of the vehicle centre for walking on the line scenario, interaction between parameters oncoming traffic and pedestrian direction

## 4.2.2 Spatio-temporal results

Figure 4.12 shows the speed distributions for the factors oncoming traffic and pedestrian direction in the line scenario, for line (Figure 4.12a) and curb 8line (Figure 4.12b). There are neither trends nor significant differences in median speed between any of the four factor combinations.

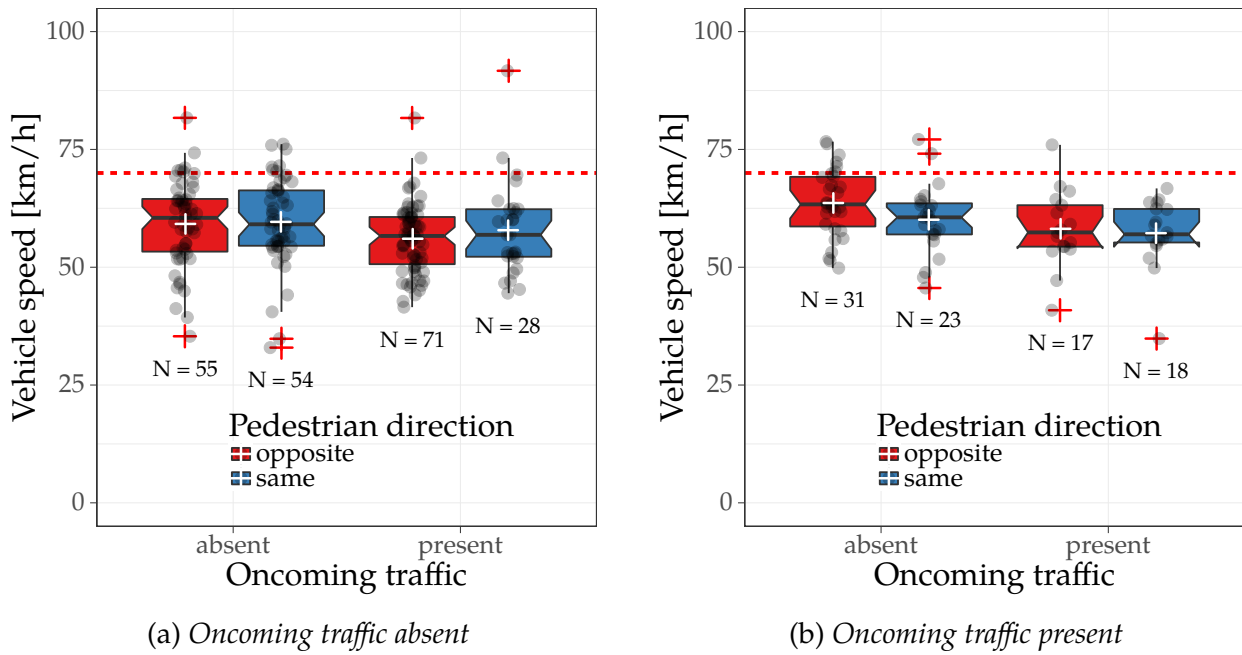


Figure 4.12: Speed of overtaking vehicles in walking on the line scenario with factors oncoming traffic and pedestrian direction, dashed line: speed limit

In Figure 4.13, the speed distributions for the curb scenario is shown. There is a visible trend that in case of oncoming traffic absent, drivers overtake at higher speeds if the pedestrian is walking opposite to the traffic direction. However, there is no statistically significant trend recognisable from the notches of the box plot diagrams in Figure 4.12.

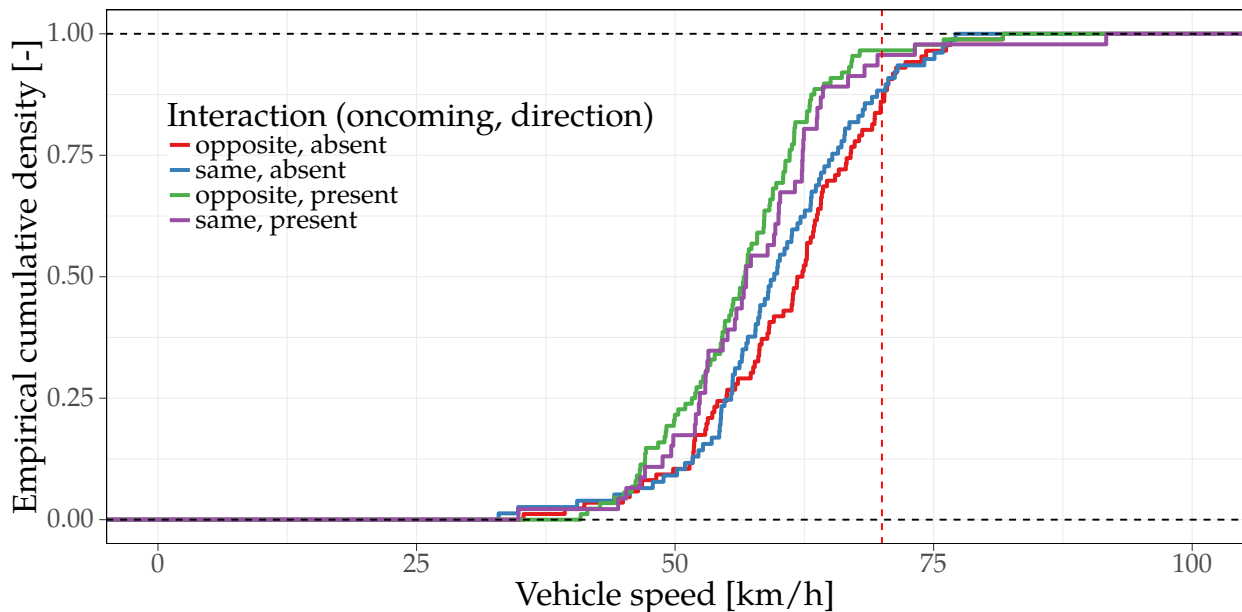


Figure 4.13: Empirical cumulative density of speed under influence of interaction between factors oncoming traffic and pedestrian direction

In Figure 4.14, MC is plotted against vehicle speed for the two different scenarios line

(4.14a) and curb (4.14b). The linear estimates are shown together with the corresponding 95% confidence bands. In both cases, MC tends to increase with vehicle speed and for oncoming traffic absent, the slope is slightly larger. In the curb scenario, the increase is more evident than in the line scenario.

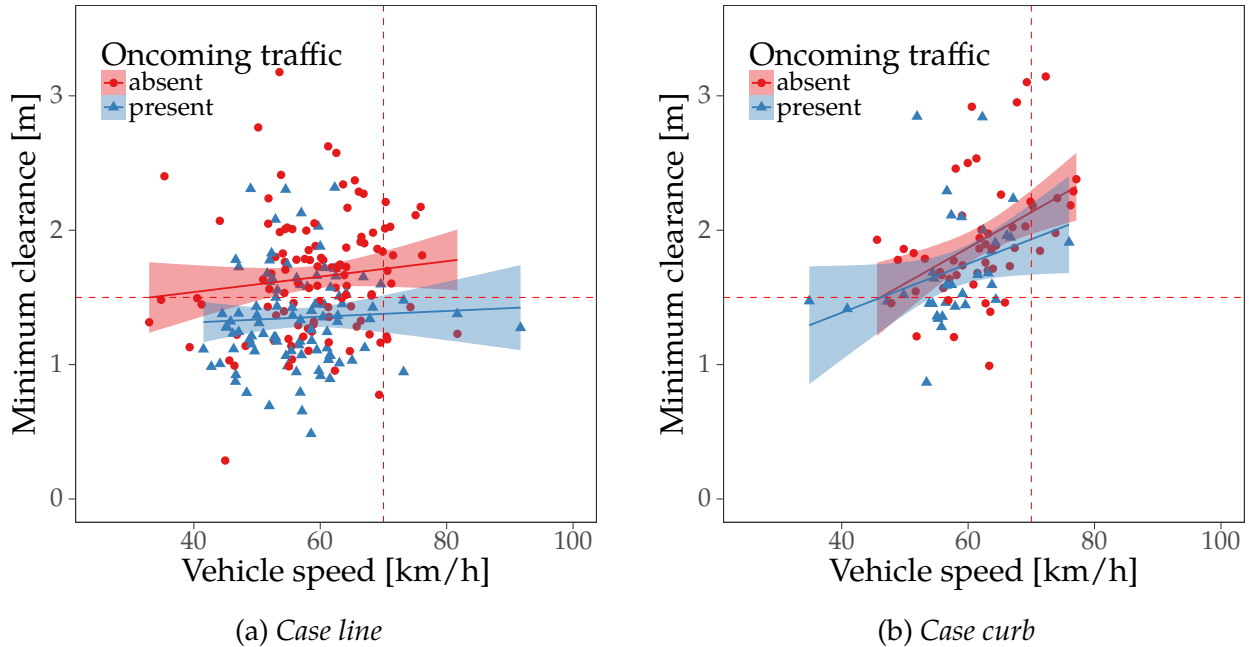


Figure 4.14: Minimum clearance over vehicle speed in walking on the line and curb scenario with factors oncoming traffic, vertical dashed line: speed limit, horizontal dashed line, 1.5 m limit set by other EU countries

### Results for different types of vehicles

In Figure 4.15, the ECDF of MC is shown for different vehicle types under the absence or presence of oncoming traffic. Vehicle type medium represents vehicles with a length of greater than 5 m and less or equal 10 m. Long vehicles are vehicles with a length more than 10 m. It can be seen that the longer vehicles leave less space to the pedestrian and almost three fourth of the long vehicles stay under the recommended distance of 1.5 m, line and curb scenario together. For all vehicle types, the trend can be observed that MC decreases under the presence of oncoming traffic. In Figure 4.16, the distribution of the lateral position of the centre of the vehicle is shown with respect to factors vehicle type and oncoming traffic. It can be noted that longer vehicles overtake the pedestrian using a similar lateral position in the lane as other vehicle types, which decreases towards the centre of the lane when oncoming traffic is present.



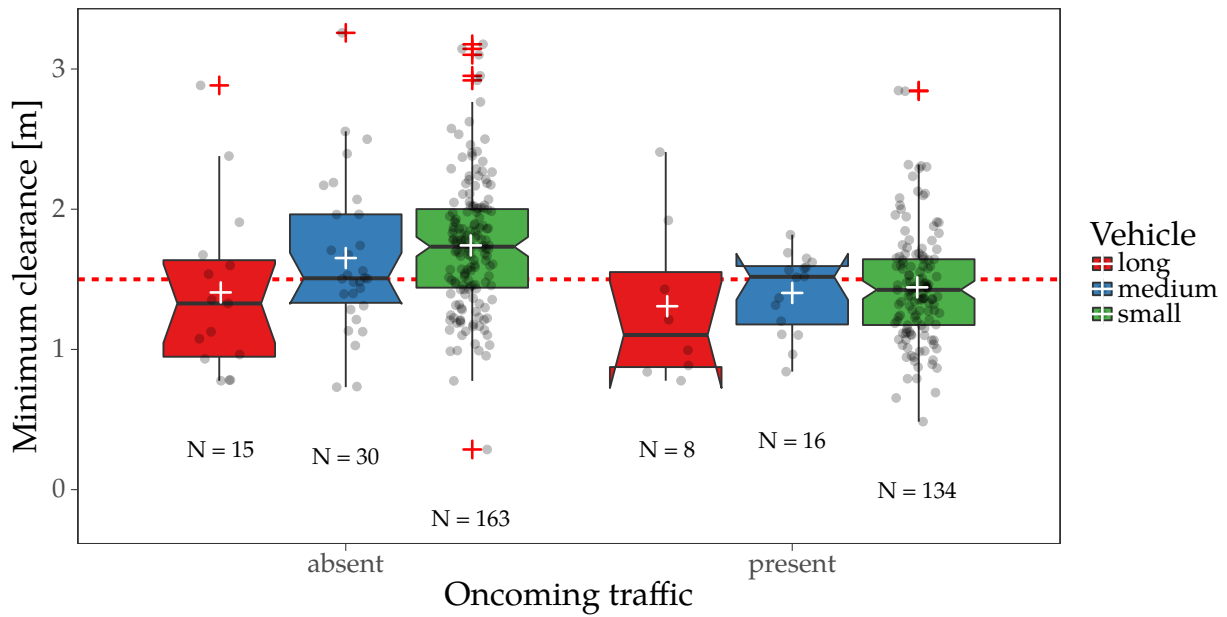


Figure 4.15: Empirical cumulative density of speed under influence of interaction between factors oncoming traffic and pedestrian direction

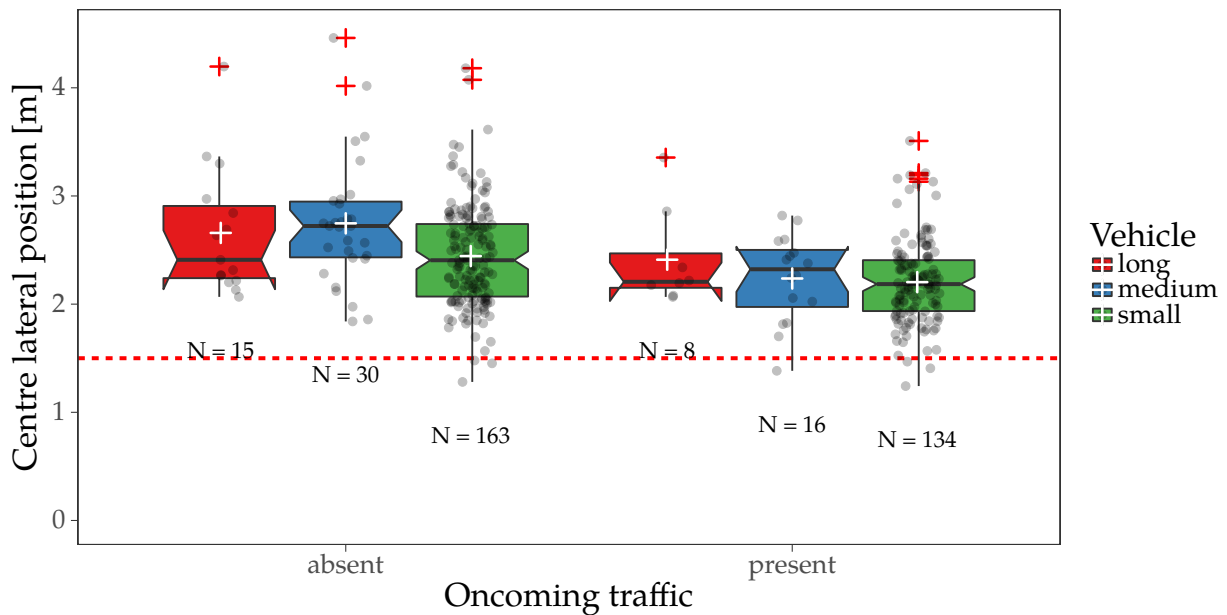


Figure 4.16: Empirical cumulative density of speed under influence of interaction between factors oncoming traffic and pedestrian direction

### 4.3 Driver modelling with Bayesian regression

The following section shows an attempt to model a pedestrian overtaking manoeuvre using linear Bayesian regression. The aim is to be able to predict posterior distributions of metrics like MC or TTC based on a linear formula with different factors. Using the Bayesian approach,

changes due to certain factor levels can be expressed as a density with an estimate as well as uncertainty to allow a realistic quantification of driver behaviour.

### 4.3.1 Minimum clearance model

#### Model from UDRIVE data

Figure 4.17 displays the density of the data and 100 posterior densities sampled, i.e. predicted, from the MC model fit in form of ECDFs (blue lines) for the UDRIVE data set, so called posterior predictive checks. The red line marks the real, measured data. As can be seen, the model is able to predict similar densities to the real MC ECDF, suggesting that the model fit is reasonable.

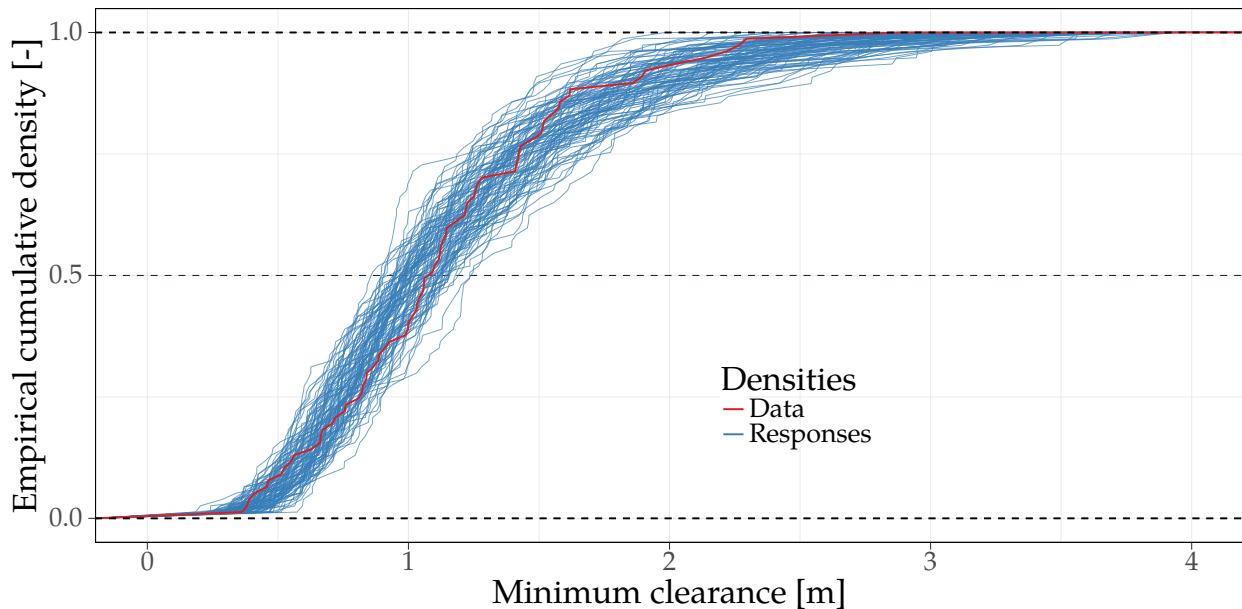


Figure 4.17: *Empirical cumulative distribution functions of posterior predictive checks from UDRIVE data using 100 samples from the fitted model (blue lines) for minimum clearance from UDRIVE data, red line marks the measured data*

The parameters oncoming traffic, pedestrian direction and lateral position have previously been identified as the main influencing factors on MC. Figure 4.20 shows the fitted and the predicted distributions from the MC model. The fitted distributions are obtained from applying the model to the original data while the predicted distributions are obtained from new data, hence contain generally a larger spread. For each interaction between the parameters, the fitted distribution is shown with the mean as well as the 1 and 2  $\sigma$  bounds, corresponding to 66 and 95 % of the distribution, respectively. As can be seen, the predicted posterior densities conditioned on the factor interactions, represent similar distributions as the fitted densities, supporting the model fit quality.

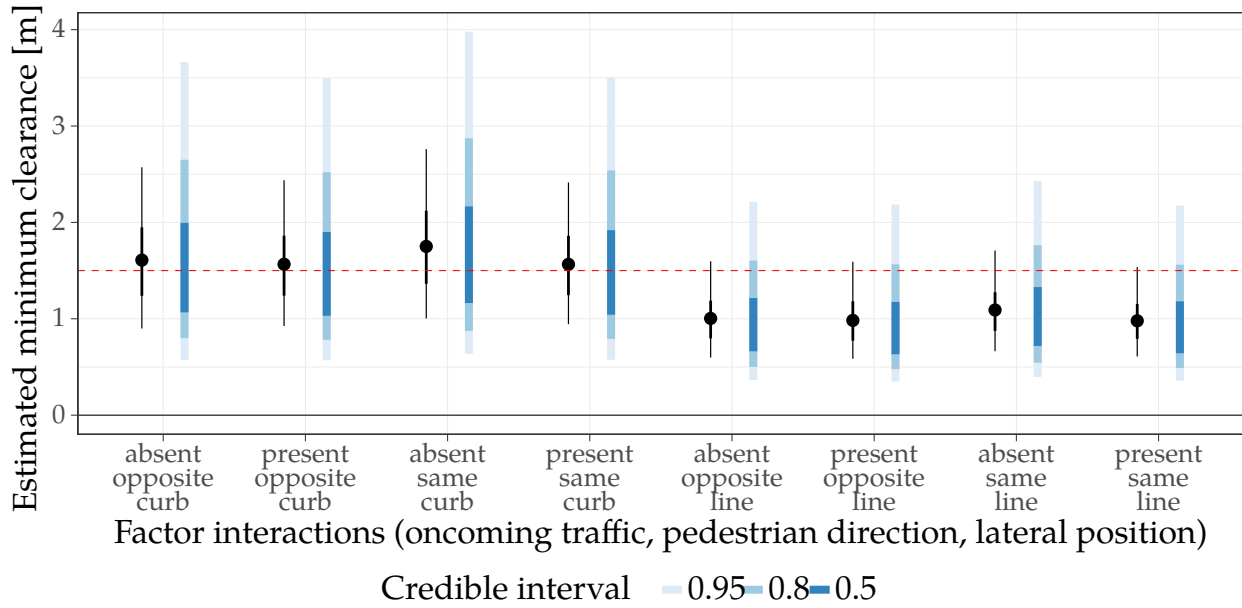


Figure 4.18: Model estimates of minimum clearance from UDRIVE data under different levels of the interaction between factors oncoming traffic, pedestrian direction and lateral position for field test data, fitted estimates are marked with black coloured bars. Fitted distributions are plotted with mean (circle), 66 % (thick line) and 95 % estimates (thin line). Predicted distributions are plotted with 50 %, 80 % and 95 % intervals. The red dashed line marks 1.5 m, the minimum distance set by law in other European countries.

### Model from field test data

In Figure 4.19, samples from the model of MC from the field test data are shown (blue lines) together with the real data (red line). The sampled densities represent a reasonable fit to the real data which leads to the conclusion that the model appropriately represents the behaviour of the drivers captured in the field tests.

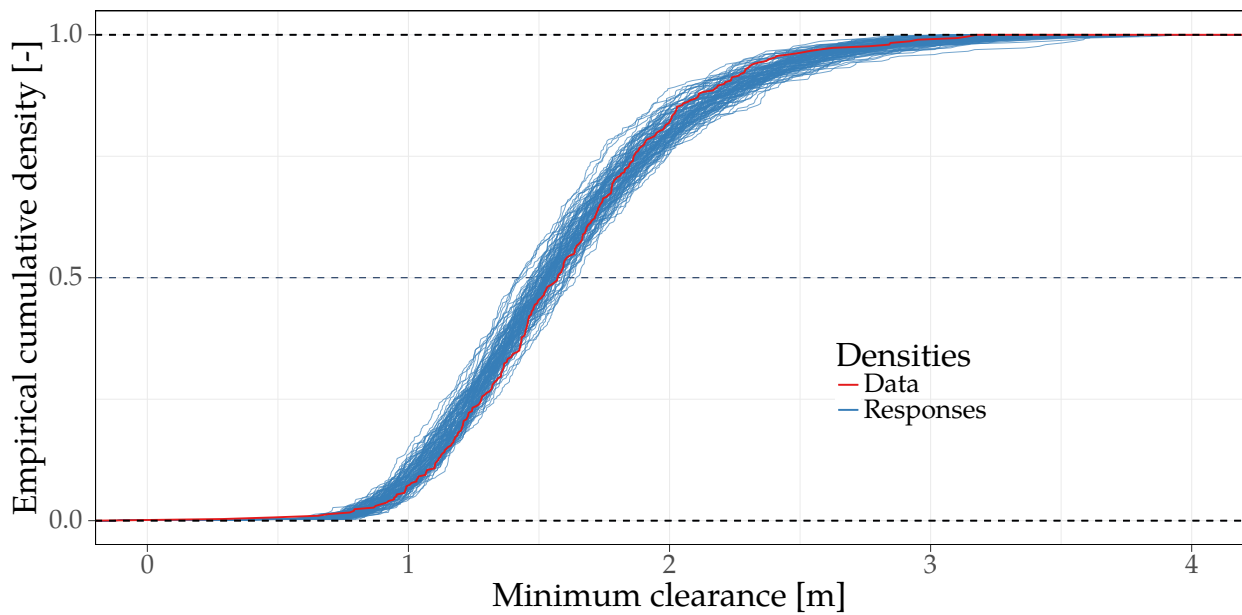


Figure 4.19: Empirical cumulative distribution functions of posterior predictive checks using 100 samples from the fitted model (blue lines) for minimum clearance from field test data, red line marks the measured data

Figure 4.20 shows the fitted and predicted distributions of the MC model from field test data under influence of the interaction between the three main influencing factors. In comparison to the UDRIVE model, the fitted and predicted densities have less spread. As for the UDRIVE model of MC, the predicted densities are similar to the fitted ones.

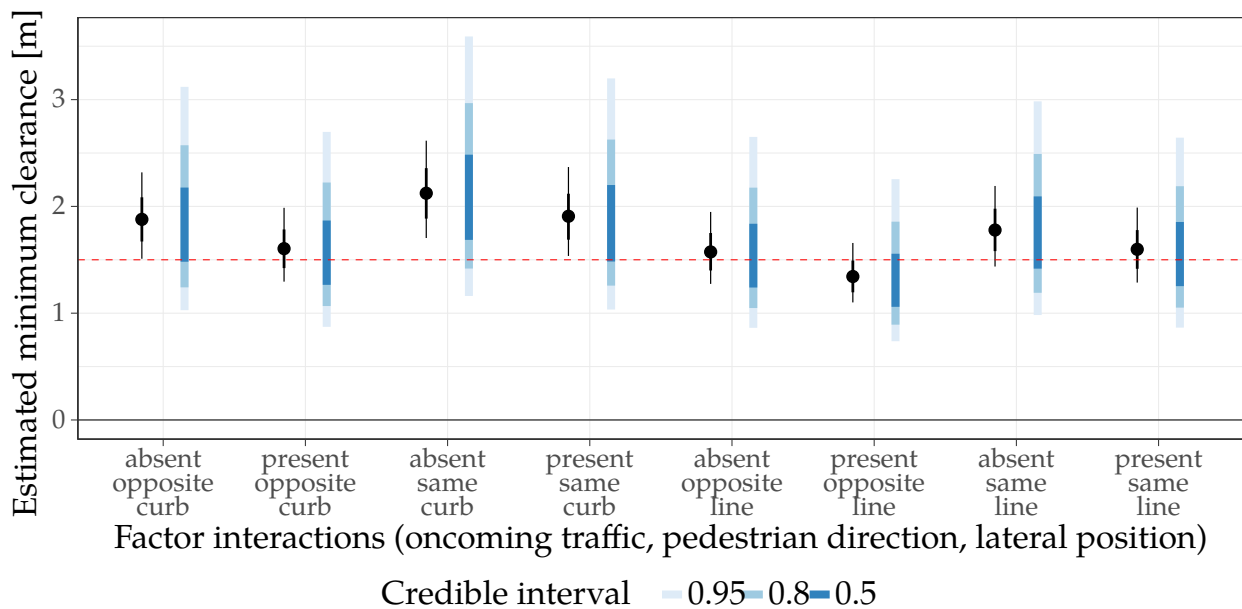


Figure 4.20: Model estimates of minimum clearance from field test data under different levels of the interaction between factors oncoming traffic, pedestrian direction and lateral position for field test data, fitted estimates are marked with black coloured bars, predicted with blue colour scale

### 4.3.2 Time to collision model

#### Model from UDRIVE data

In Figure 4.21, the data for TTC (red line) are shown together with 100 sampled posterior densities from the TTC model for UDRIVE data. The real data lies within the bounds of the estimated posterior densities, hence it can be assumed that the model is a fair representation of the TTC distribution.

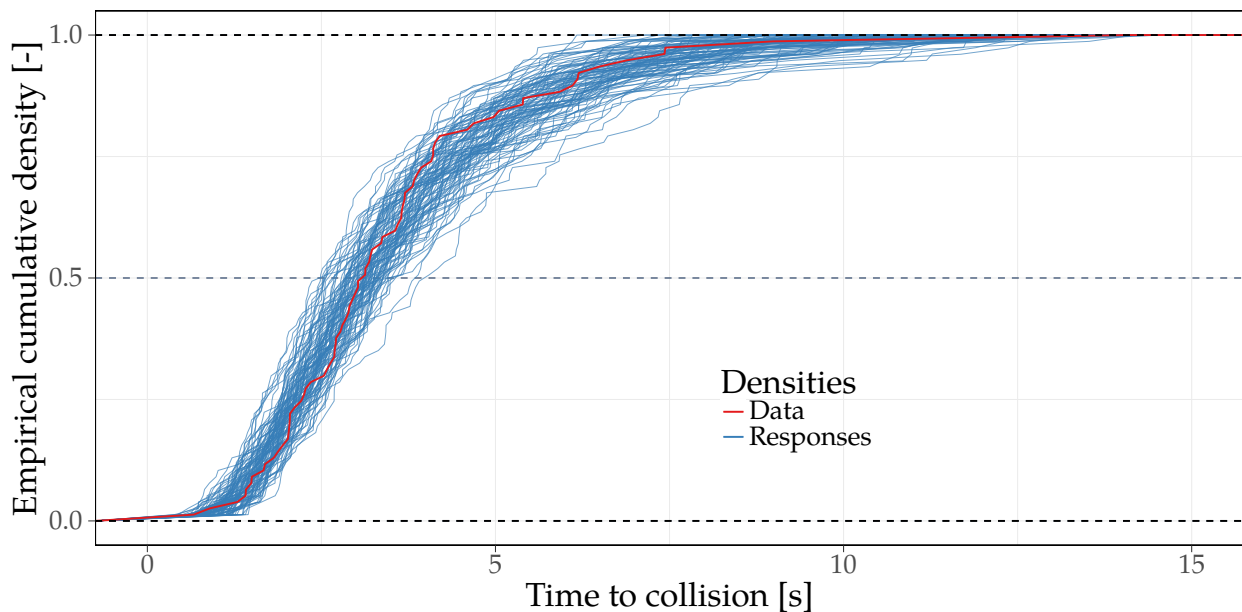


Figure 4.21: Empirical cumulative distribution functions of posterior predictive checks using 100 samples from the fitted model (blue lines) for minimum clearance from field test data, red line marks the measured data

Figure 4.22 shows the model estimates for TTC, fitted and predicted distributions. As for the previous models, the predicted densities are similarly shaped as the fitted ones with a larger spread. For some interactions, the 95% credible interval ranges about 8 s.

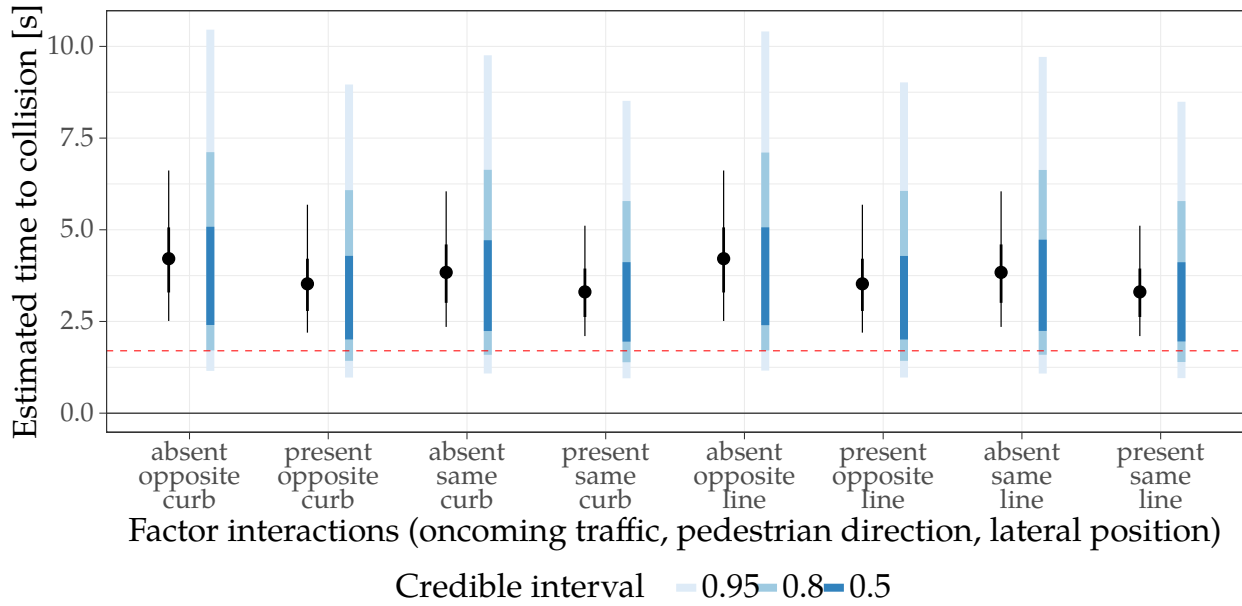


Figure 4.22: Model estimates of time to collision under different levels of the interaction between factors oncoming traffic, pedestrian direction and lateral position for UDRIVE data, fitted estimates are marked with black coloured bars, predicted with blue colour scale, red dashed line marks the Euro NCAP 1.7 s TTC minimum limit for FCW trigger

### 4.3.3 Speed model

The model for speed is tested for the field test data. Figure 4.23 shows the ECDF of the posterior predictive checks of the speed model together with the data. It can be observed that the model is able to predict the speed distribution in a reasonable manner.

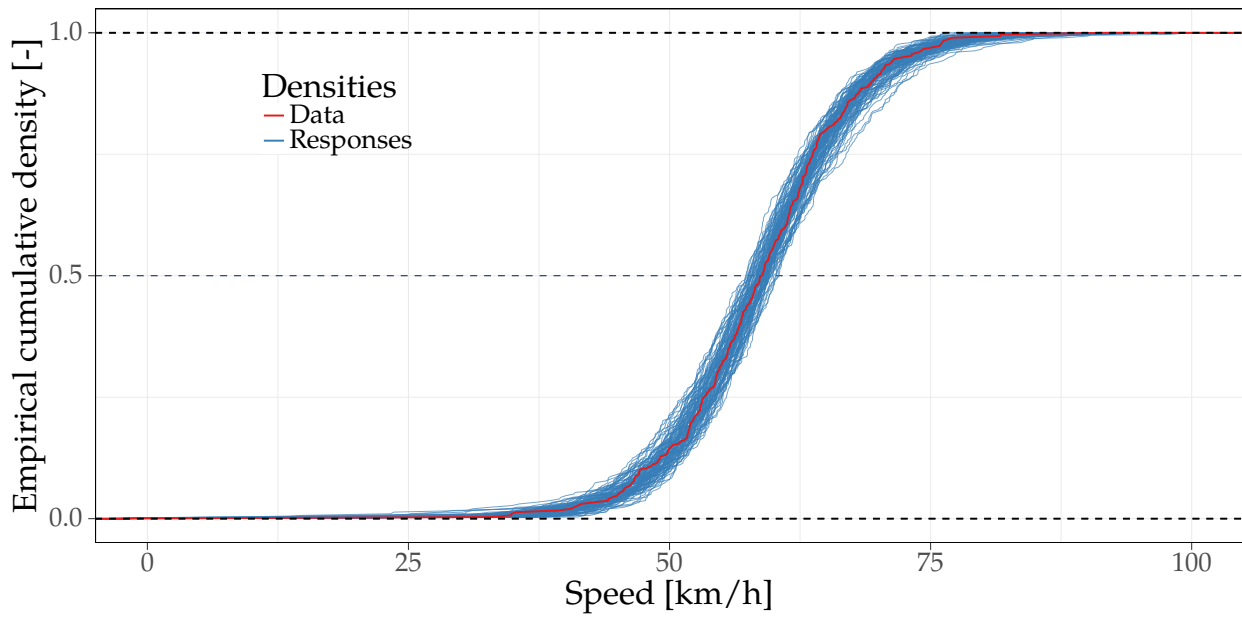


Figure 4.23: Empirical cumulative distribution functions of posterior predictive checks using 100 samples from the fitted model (blue lines) for vehicle speed from field test data, red line marks the measured data

Figure 4.24 shows the fitted and predicted densities of the model for the different factor interactions. The mean values of the predicted densities correspond roughly to the mean values of the fitted densities, supporting the fit of the model. The spread of values for the predicted densities is however clearly larger than for the fitted densities.

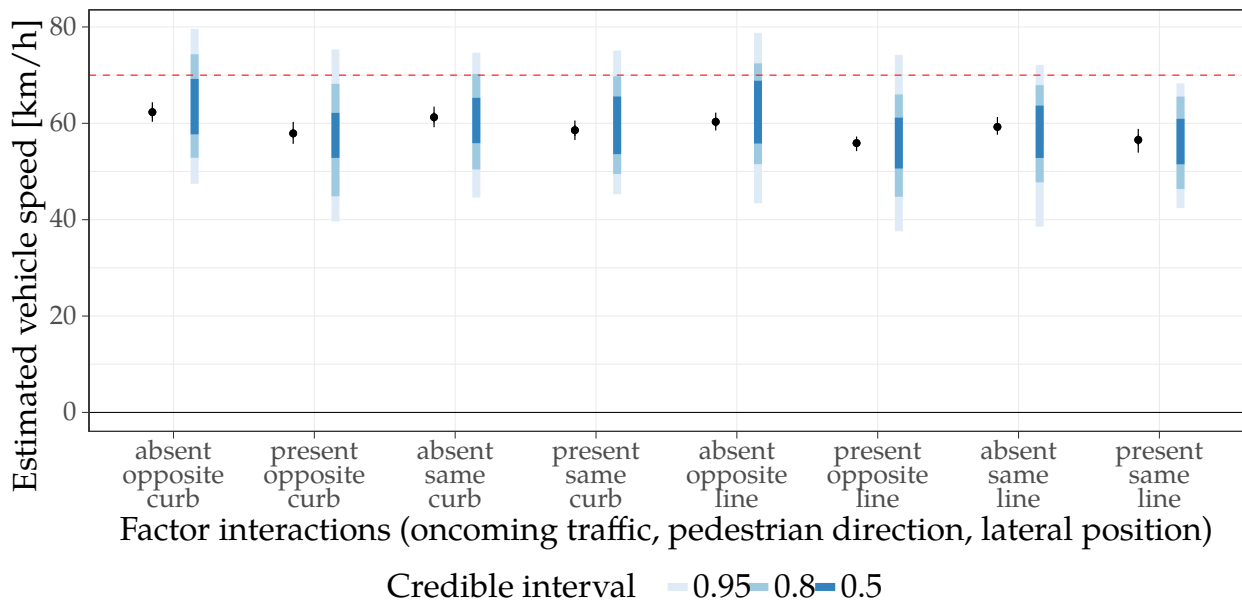


Figure 4.24: Model estimates of vehicle speed from field test data under different levels of the interaction between factors oncoming traffic, pedestrian direction and lateral position for field test data, fitted estimates are marked with black coloured bars, predicted with blue colour scale, red dashed line marks speed limit (70 km/h)





# 5

## Discussion

### 5.1 Hypotheses review

**Hypothesis 1.** *When the pedestrian walks closer to the middle of the lane, MC decreases*

The factor of the lateral position of the pedestrian could only be analysed for the field test data in which there is a sufficiently large amount of data for both line and curb conditions to indicate significant differences. The results from the field test data analysis show that MC decreases when walking on the line, in contrast to walking on the curb. However, analysis of the lateral position of the car centre from the lane marking reveals that drivers do not deviate much from the lane centre in the curb scenario. An interpretation could be that drivers see the lane marking between their lane and the pedestrian as a safety zone boundary and expect that the pedestrian will stay in his dedicated “lane”.

**Hypothesis 2.** *When driver and pedestrian face each other, MC decreases*

Results from the field tests show that when a pedestrian walks opposite to the traffic, i.e. facing the traffic, MC decreases which indicates a confirmation of the hypothesis. It seems that drivers regard a pedestrian differently based on its walking direction. One possible interpretation is that when drivers see a pedestrian walking towards them, they assume that the pedestrian sees them and is aware of the situation. Hence, there could be a mutual agreement between driver and pedestrian to not suddenly deviate from their path and accept a lower clearance. Pedestrians who are walking with their back facing the driver instead of their face, could be classified as more dangerous by the driver because the driver thinks that the pedestrian might not be aware of the situation because he cannot see the vehicle approaching from behind. The driver might fear sudden actions in terms of turning around or stepping into the lane in an attempt to cross the road. Hence, drivers give more space to the pedestrian. Driver uncertainty in the situation of not facing the pedestrian might also be inherited from the traffic law which prescribes pedestrians to walk opposite to the traffic in order to face the vehicles.

**Hypothesis 3.** *When the pedestrian walks slightly out of the lane, meaning next to the lane marking or on the curb, drivers do not steer away*

From the outer lateral position, obtained from the field test data, it is visible that drivers tend to evade less into the adjacent lane when the pedestrian is walking on the curb instead of the line, which indicates that the hypothesis is true. This relation is even clearer when the pedestrian is walking opposite to the traffic. Similarly, the lateral position of the vehicle centre decreases when changing from line to curb scenario. The UDRIVE data set could not be used to indicate significant results due to the low number of cases in which the curb scenario was present.

**Hypothesis 4.** *When drivers face oncoming traffic, MC decreases*

Results from the field tests show that MC decreases when drivers face oncoming traffic, hence the hypothesis can be verified. This result implies that drivers are willing to give less space to pedestrians who are walking in their lane in order to avoid a collision with the oncoming traffic. It can be interpreted that drivers see the oncoming traffic as a more serious threat than the pedestrian to be avoided. Furthermore, the large relative speed between vehicle and pedestrian makes it difficult for drivers to slow down sufficiently much without affecting the traffic coming from behind, to stay behind the pedestrian and wait for the oncoming traffic to pass. Hence, drivers are more willing to overtake the pedestrian even if oncoming traffic is present.

**Hypothesis 5.** *When drivers face oncoming traffic, speed is reduced*

Results from the field tests show that drivers reduce speed when facing oncoming traffic, which can be seen in both scenarios line and curb. It can therefore be assumed that the hypothesis is true. A significant speed reduction in presence of oncoming traffic is only observed when the pedestrian is walking opposite to the direction of the traffic. This fact underlines the importance of walking in opposite direction on rural roads as a pedestrian.

## 5.2 Comparison between UDRIVE and field test data results

In contrast to the field test data, the UDRIVE data is sparser in number of overtaking events. On the other hand one can argue that the events from UDRIVE are more realistic since they are recorded in everyday driving and not in an artificially created scenario like the field tests. On top of that, the pedestrian in the field tests wore a reflective vest for safety reasons which may have affected approaching drivers to keep a larger distance to the pedestrian. Another difference between UDRIVE and field test data is that the UDRIVE data contains repetitions of drivers, i.e. certain driver account for several overtaking events. While the field test data is assumed to contain only unique participants with potentially a few repetitions, the 77 events found in UDRIVE are distributed among 22 individual drivers. One participant was involved in 16 of the overtaking manoeuvres. Hence, the field test data is richer in variety of different drivers.

The results from UDRIVE reveal similar trends for the statistically significant findings from the field test data for the effects of factors oncoming traffic and pedestrian direction on MC and TTC. However, different mean values for MC are reported for the different data sets. Overtaking events from the field test data, line scenario and small vehicle type, were performed with a mean MC of 1.51 m and standard deviation 0.43 m, while in UDRIVE a mean MC of 1.15 m and a standard deviation of 0.50 m is reported. It can therefore be noted that the overtaking manoeuvres on the rural roads in France were generally performed under the recommended distance threshold of 1.5 m [2, 23]. A reason could be a higher exposure of those drivers to the scenario of having a pedestrian in the lane in the local region, meaning that those drivers are more used to overtake pedestrians. Lower deviations of the 1.5 m threshold also occurred during the field tests, with an alarming minimum MC of 0.29 m.

The UDRIVE results do not include any information about other vehicle types apart from passenger cars. From the field test data, it could be observed that with an increase in

vehicle length, MC decreases, especially when oncoming traffic is present. However, the lateral position of the vehicle centre was found to be similar among different vehicle lengths, giving rise to the conclusion that longer vehicles give less distance to pedestrians due to their geometric size. This fact emphasises the danger for pedestrians when walking on rural roads which are shared among different vehicle types.

## 5.3 Comparison to existing studies

### 5.3.1 Driver interaction with cyclists

The analysis of MC from UDRIVE and field test data reveals a decrease in MC when oncoming traffic is present which is in accordance with previous findings from studies investigating cyclist overtaking manoeuvres in [2, 6, 7, 11].

In [4], overtaking phases duration and CZBs were studied from field tests on the same road as in this thesis, including a bicycle which was equipped with the same LiDAR as in this thesis and overtaken by vehicles. Since similar definitions for the different phases were followed, a comparison to the UDRIVE findings from this study is possible, for which the duration of different phases is available. Both studies report a shorter approaching phase for flying overtaking manoeuvres in contrast to accelerative manoeuvres. Overall, shorter approaching phases are observed in this study, most probably due to the higher relative speed and a difference in definition of approaching phase start. While in [4], the start of the approaching phase was set to the time when a vehicle entered the field of view of the LiDAR, in this study the start was set to the time when a pedestrian became visible in the UDRIVE front camera video feed. Similar to this study, the steer away phase duration was roughly equal among both overtaking types. However, this study reports slightly larger values. The passing phase duration in [4] was found to be generally present and with lower variance than in this study, where the passing phase was often missing when drivers immediately returned after completing the steer away phase. The end of the return phase was reached faster in [4] than in this study, however, the contrast in definition has to be taken into account as for the approaching phase. Results in [4] further show that drivers keep about 2.03 m (0.28 m standard deviation) distance to cyclists in accelerative overtaking manoeuvres and 1.60 m (0.49 m standard deviation) in flying overtaking manoeuvres. Values obtained from the field tests in this thesis have a mean of 1.51 m and are therefore lower in average. However, most of the overtaking manoeuvres were executed in a flying manner, therefore the distances are similar to [4]. MC values from UDRIVE were found to be clearly lower, giving rise to the assumption that driver behaviour varies among different countries or cultures. In [4], CZBs were found to be not influenced by speed. In contrast, this study suggests that there is a slight proportionality between vehicle speed and MC, especially for the curb scenario.

In [2, 11], lateral clearance to cyclists, maintained by drivers, was obtained from UDRIVE data collected in France. Results showed that drivers on average left 1.29 m (0.5 m standard deviation) distance to cyclists which is under the recommended minimum distance of 1.5 m. In this study, similar but slightly lower values are reported from UDRIVE for MC with a mean of 1.15 m (0.5 m standard deviation). A trend towards lower MC values for larger vehicle types goes in accordance to [57], in which drivers of busses and heavy good vehicles drove significantly closer to the cyclists while overtaking.

The slight trend for the increase in MC with an increase in speed during the passing

phase, which is found in this work, is also reported for driver overtaking manoeuvres of cyclists in [23]. In the same study, it is reported that lateral clearance decreases when the distance between lane edge and cyclist increases which is a comparable result found in this thesis where drivers are found to give less space to a pedestrian who is walking closer to their lane instead of walking on the curb.

### 5.3.2 Driver interaction with pedestrians

In [28], interaction with pedestrians in road crossing situations was analysed under the influence of the presence and absence of eye contact. Eye contact was found to be a factor under which the TTC at brake onset of drivers increased, i.e. influenced the CA strategy in form of braking. TTC at brake onset was found to be decreased when the pedestrian looked at a smart phone while walking. In this study, similar results are reported for CA by steering manoeuvres in which drivers on average started to steer away about one second TTC earlier when facing the pedestrian. Similar to the field tests in this study, [28] reports an influence on driver behaviour when eye contact is present, which can be assumed as well when the pedestrian was walking opposite to the traffic.

## 5.4 Impact on ADAS design and evaluation

Models such as the BRMs of the overtaking manoeuvre metrics could be used in ADAS or autonomous driving to quantify the interaction between vehicle and pedestrian. The models derived in this thesis can be used as a prior distribution which is then trained to the individual driver to be able to draw samples from a distribution which resembles the driver behaviour as closely as possible. Predicting metrics close to the individual driver behaviour can enhance product acceptance and avoid that drivers switch off or ignore the system.

For example, by extending an FCW system in which the algorithm is TTC based with a driver model such as the one derived in this thesis, the warning could be triggered when TTC falls below the individual lower 95% credible interval. However, the AEB or a CA system for steering should still intervene before the DVE surpasses a general comfort zone boundary, which also takes into account the comfort of the pedestrian, and the safety zone boundary to avoid the collision with the pedestrian.

In [58], driver overtaking manoeuvres of other road users like cyclists were investigated in a simulator experiment with participants involved. Results showed that subjective ratings of automated driving features suggested longer lateral distances to the cyclist and earlier steer away times than in manual driving. Similar findings might need to be expected as well in pedestrian overtaking manoeuvres such that systems rather activate earlier when the vehicle is in automated mode.

# 6

## Conclusion

In this thesis, two data sets of overtaking manoeuvres involving were obtained and analysed. The first data set was extracted from the NDS UDRIVE using a similar methodology which has been used in a previous study on bicycle overtaking manoeuvres in UDRIVE. The extracted UDRIVE data set contains 77 pedestrian overtaking manoeuvres which happened on rural roads in France. Based on the available events, overtaking phases were annotated, based on a four phase manoeuvre approach. An analysis methodology was developed to reconstruct the vehicle and the pedestrian trajectory during the overtaking manoeuvre in order to retrieve the MC metric. Other metrics include speed, acceleration, TTC and MAG to the pedestrian.

The second data set was obtained from field tests with a data logger mounted on the hip of a pedestrian. The data logger was implemented using ROS running on an RPi to record data from sensors like LiDAR, IMU, GPS and camera. During two days of collection, 630 overtaking events were recorded on a rural road stretch in Sweden. Two different pedestrian lateral positions, walking on the lane marking line and walking 0.5 m away from it at the road curb, were tested. For both lateral positions, two possible walking directions were tested, walking in the same and opposite direction to the traffic. After transforming the 2D LiDAR data into a 3D point cloud with the estimated orientation from the IMU data, noise from vegetation and ground was removed and the point cloud was clustered to detect vehicles. By tracking the clusters over time, an average speed and an MC value could be estimated for an overtaking vehicle. For the remaining 481 vehicles, the factors overtaking strategy piggy backing present or not, and oncoming traffic were evaluated according to existing literature.

Results of the UDRIVE data set show that the passing phase of the analysed overtaking manoeuvres are significantly shorter than the other phases (mean 1.16 s for accelerative and 1.03 s for flying manoeuvres). For accelerative manoeuvres the median passing phase duration is 0 s and for flying overtaking manoeuvres 1.03 s, suggesting that in accelerative manoeuvres, drivers return to their original path as soon as they have steered away enough to pass the pedestrian. Qualitative observations from video and signal data of the steering wheel have confirmed this assumption. Flying manoeuvres usually occurred when the driver had a good overview of the scenery and evaded earlier into the adjacent lane, hence prolonging the passing phase. Drivers were often found to return even before having reached the pedestrian. The average time of returning was found to be 0.52 s (0.72 s standard deviation) before the front of the vehicle reached the pedestrian.

TTC at the moment of steering away was found to be increased in the presence of oncoming traffic as well as when the pedestrian was walking in opposite direction of traffic, giving rise to the assumption that eye contact or at least seeing a pedestrian from the front plays a key role in the interaction between driver and pedestrian. 96% of all flying overtaking manoeuvres in the UDRIVE data set were performed above the  $TTC = 1.7$  s limit set by Euro NCAP for the CPLA scenarios, i.e. the Euro NCAP scenario goes in accordance to the CZBs measured for the drivers in this thesis. Both MC and TTC were found to be slightly

increasing with vehicle speed.

The results from the field tests confirm an increase in MC to the pedestrian when oncoming traffic is present. When the pedestrian is walking outside of the lane, MC increases, hence drivers give more space to the pedestrian, but steer away less. Furthermore, a decrease of MC is observed when the pedestrian is walking in opposite direction to the traffic. This fact gives further rise to the assumption that eye contact between driver and pedestrian plays a crucial role in the interaction.

Linear Bayesian regression models for MC, TTC and passing speed were developed and estimated to reflect driver behaviour dependent on the most influencing factors oncoming traffic, pedestrian direction and vehicle speed. Model fits predict the measured values for the metrics well, meaning that the measured data lies within the credible intervals of the predicted densities. Estimated contrasts between different factor levels reveal distributions which coincide with the qualitative frequentist analysis.

In an ISS like AEB or FCW, Bayesian regression models could be used to take into account driver behaviour when approaching pedestrians to enhance acceptance of the systems and to prevent false positive interventions or alerts. The Bayesian regression model therefore needs to be computed for the individual driver, hence values presented in this thesis can only resemble defaults or used as a prior to estimate the individual driver's posterior densities for the different metrics.

# 7

## Limitations and further work

This work includes analysis of overtaking manoeuvres from two data sets of participants, one from drivers in France and the other one from a particular road stretch in Sweden. Further research should aim at investigating the generalizability of the results to include other regions or cultures of drivers. In addition, future tests should aim at quantifying the pedestrian's comfort zone to verify that it matches with the estimated comfort zone for the driver.

In the UDRIVE data analysis, the extraction of overtaking events and analysis depends on the amount of ME detections. Hence, there is an unknown set of pedestrian overtaking manoeuvres which was filtered out because the number of ME observations was too low. In a future work, the use of image analysis methods could be investigated to identify and localise pedestrians from the video cameras installed in the car, or to fuse that information with the ME position estimates. For UDRIVE, TTC is calculated using the minimum Euclidean distance between vehicle and pedestrian. This can only be a valid approximation if the road is straight, hence, further work could aim at estimating the actual path length to get a more accurate TTC estimate.

The extracted metrics from the PDL data contain neither spatial nor temporal information about the overtaking phases because it was not possible to extract those due to the fact that the LiDAR was wobbling too much during the field tests. In similar research in the future, it could be useful to either stabilise the LiDAR using a gimbal construction which is also used to stabilise cameras, or to use a 3D LiDAR. Further missing is the classification of overtaking manoeuvre strategy due to the lack of temporal data.

As an inheritance from the missing temporal data of the overtaking phases, it should be further investigated if the four phase approach is appropriate for a longitudinal manoeuvre including cars and pedestrians. From the UDRIVE results for the phases duration, there is a clear trend that drivers do not enter a passing phase but rather steer back immediately after gaining enough lateral clearance to the pedestrian. However, this hypothesis should be tested on a larger data set to permit a change in definition of phases or e.g. remove the passing phase from the definition.





# References

- [1] European Road Safety Observatory. *Traffic Safety Basic Facts 2016 - Pedestrians*. Tech. rep. 2016, pp. 1–9. DOI: 10.1136/bmj.330.7487.367.
- [2] J. Kovaceva et al. Drivers overtaking cyclists in the real-world: evidence from a naturalistic driving study. *Article under revision* (2017).
- [3] M. Brännström, E. Coelingh, and J. Sjöberg. Model-based threat assessment for avoiding arbitrary vehicle collisions. *IEEE Transactions on Intelligent Transportation Systems* **11.3** (2010), 658–669. DOI: 10.1109/TITS.2010.2048314.
- [4] M. Dozza, A. Rasch, and C. N. Boda. An Open-Source Data Logger for Field Cycling Collection: Design and Evaluation (2017). DOI: 10.6084/m9.figshare.5404918.v1.
- [5] R. Schindler and V. Bast. *Drivers' comfort boundaries when overtaking a cyclist: Set-up and verification of a methodology for field data collection and analysis*. 2015.
- [6] M. Dozza et al. How do drivers overtake cyclists? *Accident Analysis and Prevention* **88** (2016), 29–36. ISSN: 00014575. DOI: 10.1016/j.aap.2015.12.008.
- [7] C. Moretto. *Drivers' comfort zone boundaries during overtaking of bicycles in Japan*. 2017.
- [8] C. N. Boda et al. Modelling how drivers respond to a bicyclist crossing their path at an intersection: How do test track and driving simulator compare? *Accident Analysis and Prevention* **111**. July 2017 (2018), 238–250. ISSN: 00014575. DOI: 10.1016/j.aap.2017.11.032.
- [9] J. J. Gibson and L. E. Crooks. A Theoretical Field-Analysis of Automobile-Driving. *The American Journal of Psychology* **51.3** (1938), 453–471. ISSN: 00029556.
- [10] M. Ljung Aust. "Evaluation Process for Active Safety Functions: Addressing Key Challenges in Functional, Formative Evaluation of Advanced Driver Assistance Systems". PhD thesis. 2012.
- [11] G. Nero. *Quantifying Drivers' Behaviours when Overtaking Bicyclists on Rural Roads A Study Using Naturalistic Driving Data from a Vehicle's Perspective*. 2017.
- [12] J. Bärghman et al. *The UDrive dataset and key analysis results*. Tech. rep. 2017. DOI: 10.26323/UDRIVE.
- [13] World Health Organization. *Global Status Report On Road Safety*. Tech. rep. 2015.
- [14] M. Yanagisawa et al. *Estimation of Potential Safety Benefits for Pedestrian Crash Avoidance / Mitigation Systems*. Tech. rep. April. 2017.
- [15] National Highway Traffic Safety Administration. *2016 Fatal Motor Vehicle Crashes: Overview*. Tech. rep. October. 2017, pp. 1–9. DOI: DOTH812456.
- [16] C. V. Zegeer and M. Bushell. Pedestrian crash trends and potential countermeasures from around the world. *Accident Analysis & Prevention* **44.1** (Jan. 2012), 3–11. ISSN: 0001-4575. DOI: 10.1016/J.AAP.2010.12.007.
- [17] G. Markkula. "Driver behavior models for evaluating automotive active safety: From neural dynamics to vehicle dynamics". PhD thesis. 2015. ISBN: 9789175971537.
- [18] O. Benderius. "Modelling driver steering and neuromuscular behaviour". PhD thesis. 2014, p. 64. ISBN: 9789175970905. DOI: 10.1109/SMC.2015.274. arXiv: arXiv:1011.1669v3.
- [19] V. Papakostopoulos, N. Marmaras, and D. Nathanael. The "field of safe travel" revisited: interpreting driving behaviour performance through a holistic approach. *Transport Reviews* **37.6** (2017), 695–714. ISSN: 14645327. DOI: 10.1080/01441647.2017.1289992.

- [20] M. Ljung Aust and J. Engström. A conceptual framework for requirement specification and evaluation of active safety functions. *Theoretical Issues in Ergonomics Science* **12.1** (2011), 44–65. ISSN: 1463922X. DOI: 10.1080/14639220903470213.
- [21] T. Richter and S. Ruhl. *Untersuchung von Maßnahmen zur Prävention von Überholunfällen auf einbahnigen Landstraßen*. Tech. rep. 24. 2014.
- [22] Bundesministeriums der Justiz und für Verbraucherschutz. *Straßenverkehrs-Ordnung (StVO)*. 2017, pp. 1–72.
- [23] B. Jonas et al. The UDrive dataset and key analysis results (2017). DOI: 10.26323/UDRIVE.
- [24] Näringsdepartementet RS T. *Trafikförordning (1998:1276)*.
- [25] N. Lubbe and E. Rosén. Pedestrian crossing situations: Quantification of comfort boundaries to guide intervention timing. *Accident Analysis and Prevention* (2014). ISSN: 00014575. DOI: 10.1016/j.aap.2014.05.029.
- [26] N. Lubbe and J. Davidsson. Drivers' comfort boundaries in pedestrian crossings: A study in driver braking characteristics as a function of pedestrian walking speed. *Safety Science* (2015). ISSN: 18791042. DOI: 10.1016/j.ssci.2015.01.019.
- [27] S. F. Nejad. *Drivers' comfort zone boundaries during overtaking of bicycles in Japan without oncoming traffic*. 2017.
- [28] Z. Ren, X. Jiang, and W. Wang. Analysis of the Influence of Pedestrians' eye Contact on Drivers' Comfort Boundary during the Crossing Conflict. *Procedia Engineering* **137** (2016), 399–406. ISSN: 18777058. DOI: 10.1016/j.proeng.2016.01.274.
- [29] H. Jeppsson, M. Östling, and N. Lubbe. Real life safety benefits of increasing brake deceleration in car-to-pedestrian accidents: Simulation of Vacuum Emergency Braking. *Accident Analysis and Prevention* **111**. April 2017 (2018), 311–320. ISSN: 00014575. DOI: 10.1016/j.aap.2017.12.001.
- [30] M. Brännström, E. Coelingh, and J. Sjöberg. Decision-making on when to brake and when to steer to avoid a collision. *International Journal of Vehicle Safety* **7.1** (2014), 87. ISSN: 1479-3105. DOI: 10.1504/IJVS.2014.058243.
- [31] R. Chen, R. Sherony, and H. C. Gabler. Comparison of Time to Collision and Enhanced Time to Collision at Brake Application during Normal Driving. *SAE Technical Paper* 2016-01-1448 (2016). ISSN: 01487191. DOI: 10.4271/2016-01-1448.
- [32] M. Aron, M.-B. Biecheler, and J.-F. Peytavin. "Time headway on french rural roads". *IFAC Control in Transportation Systems*. 2003, pp. 221–226. DOI: 10.1016/S1474-6670(17)32423-0.
- [33] I. Batkovic et al. A Computationally Efficient Model for Pedestrian Motion Prediction. *CoRR* **abs/1803.04702** (2018).
- [34] C.-N. Boda. *Driver interaction with vulnerable road users: Understanding and modelling driver behaviour for the design and evaluation of intelligent safety systems*. Licentiate thesis. 2017.
- [35] R. Schram et al. Euro NCAP's first step to assess Autonomous Emergency Braking (AEB) for Vulnerable Road Users. *The 24th International Technical Conference on the Enhanced Safety of Vehicles (ESV)* (2015), 1–7. URL: [www.grsroad%20safety.org](http://www.grsroad%20safety.org).
- [36] M. Van Ratingen et al. The European New Car Assessment Programme: A historical review. *Chinese Journal of Traumatology - English Edition* **19.2** (2016), 63–69. ISSN: 10081275. DOI: 10.1016/j.cjtee.2015.11.016. URL: <http://dx.doi.org/10.1016/j.cjtee.2015.11.016>.

- [37] Euro NCAP. *European New Car Assessment Programme, Pedestrian Testing Protocol*. Tech. rep. November. 2017.
- [38] Euro NCAP. *European New Car Assessment Programme, Assessment Protocol – Pedestrian Protection*. Tech. rep. November. 2017.
- [39] R. Eenink et al. UDRIVE the European naturalistic driving study. *Transportation Research Arena* **32.0** (2014), 1–10.
- [40] L. Fridman et al. MIT Autonomous Vehicle Technology Study: Large-Scale Deep Learning Based Analysis of Driver Behavior and Interaction with Automation (2017), 1–17. arXiv: 1711.06976.
- [41] L. Fridman et al. Driver Gaze Region Estimation without Use of Eye Movement. *IEEE Intelligent Systems* **31.3** (2016), 49–56. ISSN: 15411672. DOI: 10.1109/MIS.2016.47. arXiv: 1507.04760.
- [42] RSS2017 - Road safety & simulation international conference. *Workshop: Naturalistic Driving*. <http://rss2017.org/workshop-naturalistic-driving/>. [Online; accessed 14-June-2018]. 2017.
- [43] c. b. S. UDRIVE SP and WP leaders. *UDRIVE Deliverable 63.3*. Tech. rep. 2017.
- [44] E. Union. *eUropean naturalistic Driving and Riding for Infrastructure & Vehicle safety and Environment*. 2018.
- [45] H. Tattgrain et al. eUropean naturalistic Driving and Riding for Infrastructure and Vehicle safety and Environment (2017).
- [46] M. a. Fischler and R. C. Bolles. *Random Sample Consensus: A Paradigm for Model Fitting with*. Tech. rep. 1981, pp. 381–395. DOI: 10.1145/358669.358692. arXiv: 3629719.
- [47] K. G. Derpanis. *Overview of the RANSAC Algorithm*. Tech. rep. 1. 2010, pp. 2–3. DOI: 10.1002/cne.901000107.
- [48] M. Ambrož. Raspberry Pi as a low-cost data acquisition system for human powered vehicles. *Measurement* **100** (2017), 7–18. ISSN: 0263-2241. DOI: <https://doi.org/10.1016/j.measurement.2016.12.037>.
- [49] Z. Diao et al. Analysis and compensation of MEMS gyroscope drift. *Proceedings of the International Conference on Sensing Technology, ICST 2* (2013), 592–596. ISSN: 21568065. DOI: 10.1109/ICSensT.2013.6727722.
- [50] S. Madgwick. *An efficient orientation filter for inertial and inertial/magnetic sensor arrays*. Tech. rep. 2010, p. 32. DOI: 10.1109/ICORR.2011.5975346.
- [51] R. B. Rusu. “Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments”. PhD thesis. Computer Science department, Technische Universität München, Germany, Oct. 2009.
- [52] J. L. Bentley. Multidimensional binary search trees used for associative searching. *Communications of the ACM* **18.9** (1975), 509–517. ISSN: 00010782. DOI: 10.1145/361002.361007. arXiv: arXiv:1011.1669v3.
- [53] P. Hoff et al. *A first course in Bayesian Statistics*. Vol. 102. 2006, p. 618. ISBN: 9780387781884. DOI: 10.1016/j.peva.2007.06.006. arXiv: arXiv:1011.1669v3.
- [54] J. K. Kruschke. *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan, second edition*. 2014, pp. 1–759. ISBN: 9780124058880. DOI: 10.1016/B978-0-12-405888-0.09999-2. arXiv: arXiv:1011.1669v3.
- [55] P.-C. Bürkner. brms: An R Package for Bayesian Multilevel Models using Stan. *Journal of Statistical Software* **80.1** (2017). ISSN: 1548-7660. DOI: 10.18637/jss.v080.i01.
- [56] S. D. Team. *Stan*. <http://mc-stan.org/>. [Online; accessed 14-June-2018]. 2018.

- [57] I. Walker. Drivers overtaking bicyclists: Objective data on the effects of riding position, helmet use, vehicle type and apparent gender. *Accident Analysis & Prevention* **39.2** (2007), 417–425. ISSN: 0001-4575. DOI: <https://doi.org/10.1016/j.aap.2006.08.010>.
- [58] G. Abe, K. Sato, and M. Itoh. Driver trust in automated driving systems: The case of overtaking and passing. *IEEE Transactions on Human-Machine Systems* **48.1** (2017), 85–94. ISSN: 21682291. DOI: 10.1109/THMS.2017.2781619.

# A

## Pedestrian data logger

### A.1 Software

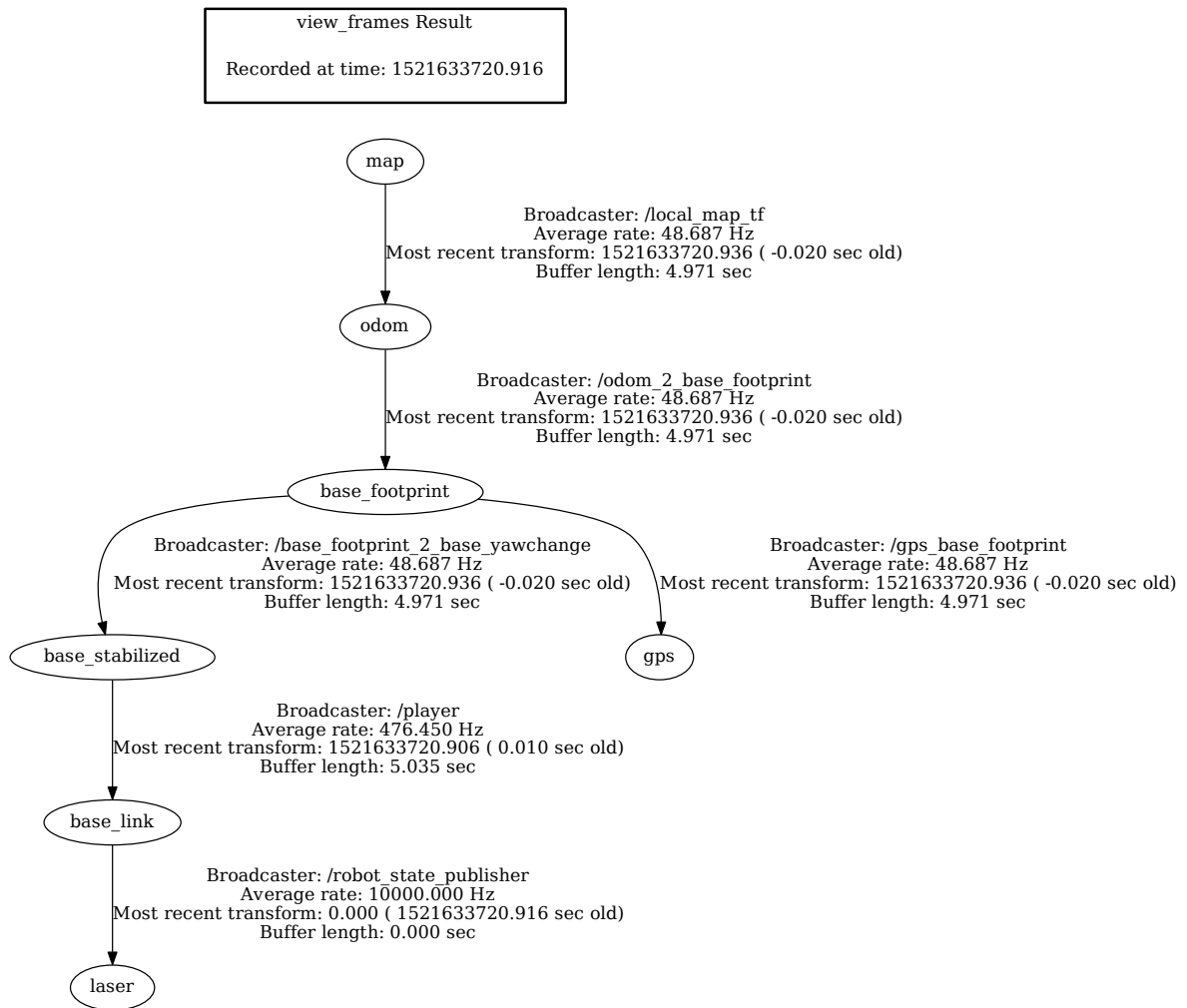


Figure A.1: Frame tree of pedestrian data logger, for each frame, the broadcasting node, the update rate, the time of the latest update and the buffer length is shown

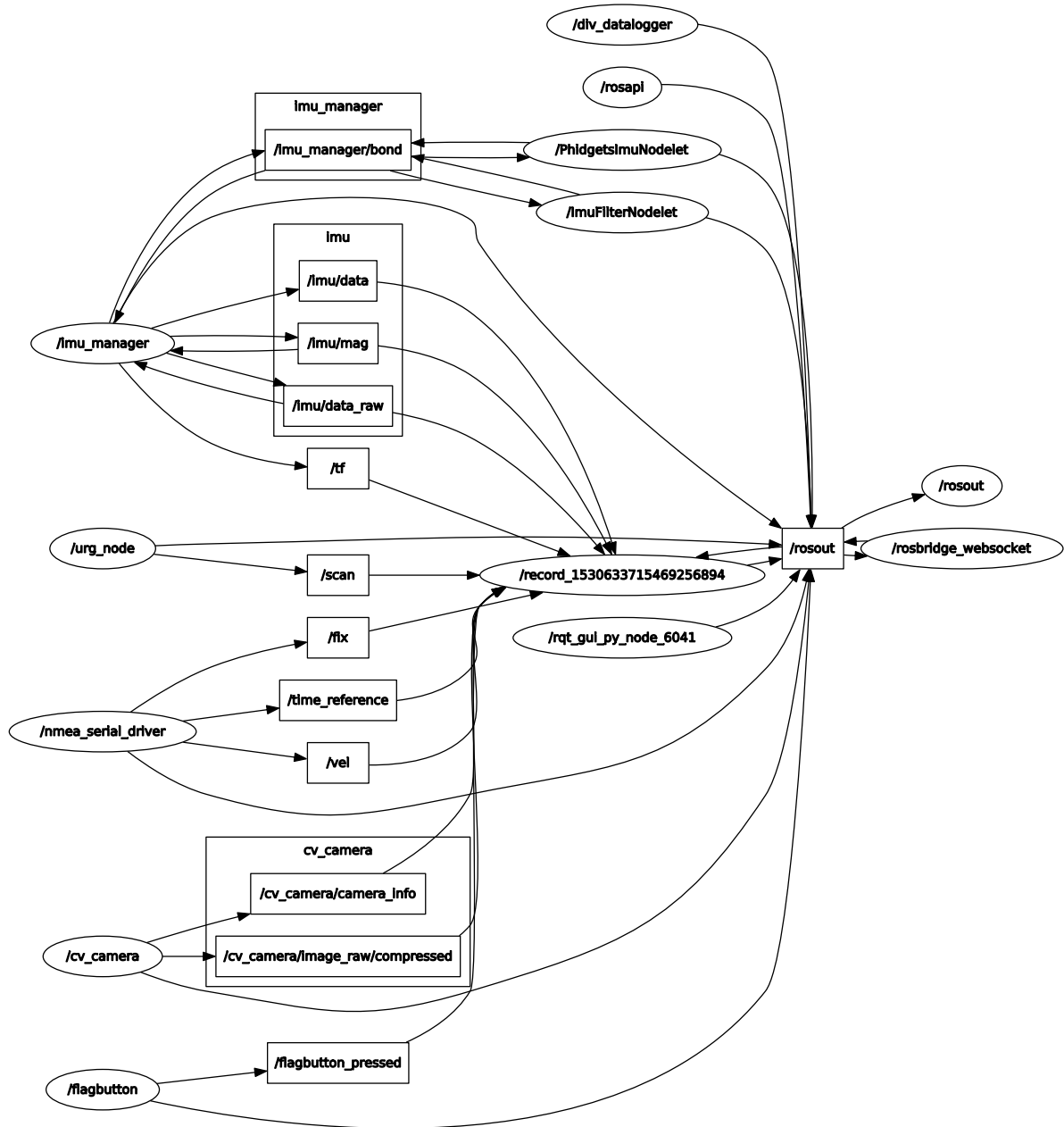


Figure A.2: ROS nodes and topics overview of `div_datalogger` package

## A.2 Field test data overview

		Oncoming traffic				
		absent		present		
		Pedestrian direction				
		opposite	same	opposite	same	
small	N [-]	55	54	71	28	208
	mean [m]	1.54	1.76	1.28	1.52	1.51
	std [m]	0.44	0.44	0.34	0.33	0.43
medium	N [-]	2	21	4	7	34
	mean [m]	1.50	1.63	1.41	1.22	1.51
	std [m]	0.09	0.60	0.22	0.28	0.52
long	N [-]	6	7	3	1	17
	mean [m]	0.97	1.62	0.91	1.78	1.22
	std [m]	0.32	0.63	0.08	0	0.55

Table A.1: Minimum clearance from field test overtaking manoeuvres for line scenario, including number of observations (N), mean and standard deviation (std)

		Oncoming traffic				
		absent		present		
		Pedestrian direction				
		opposite	same	opposite	same	
small	N [-]	31	23	17	18	89
	mean [m]	1.87	2.00	1.57	1.83	1.84
	std [m]	0.44	0.47	0.29	0.47	0.45
medium	N [-]	1	6	1	4	12
	mean [m]	1.13	1.88	1.69	1.64	1.72
	std [m]	0	0.47	0	0.13	0.39
long	N [-]	1	1	1	3	5
	mean [m]	2.38	1.54	1.43	1.85	1.81
	std [m]	0	0	0	0.60	0.50

Table A.2: Minimum clearance from field test overtaking manoeuvres for curb scenario, including number of observations (N), mean and standard deviation (std)





# B

## Bayesian regression models

### B.1 Minimum clearance model

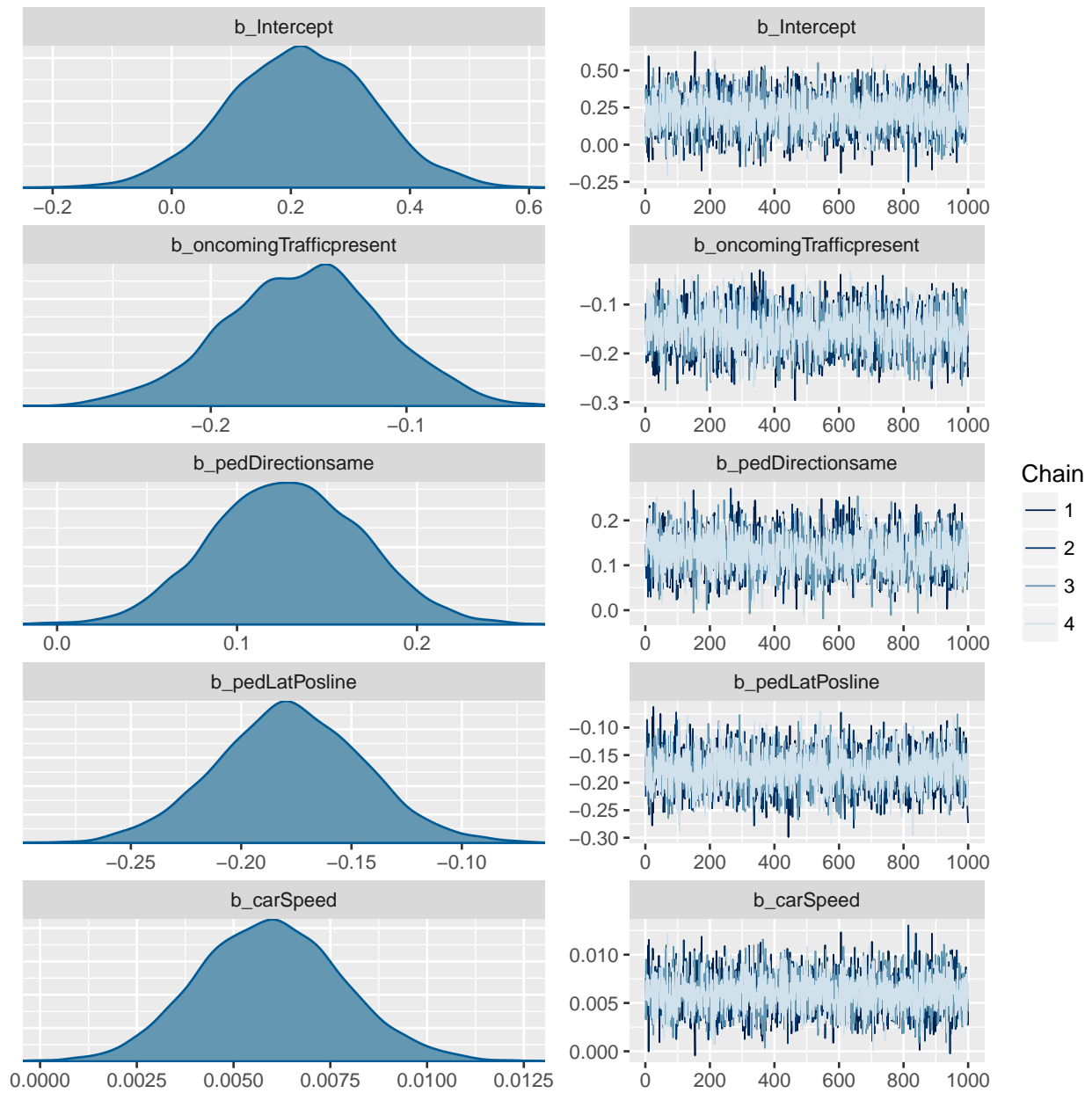


Figure B.1: MC model fit from field test data, including parameter posterior distributions (left) and MCMC chains (right), part 1

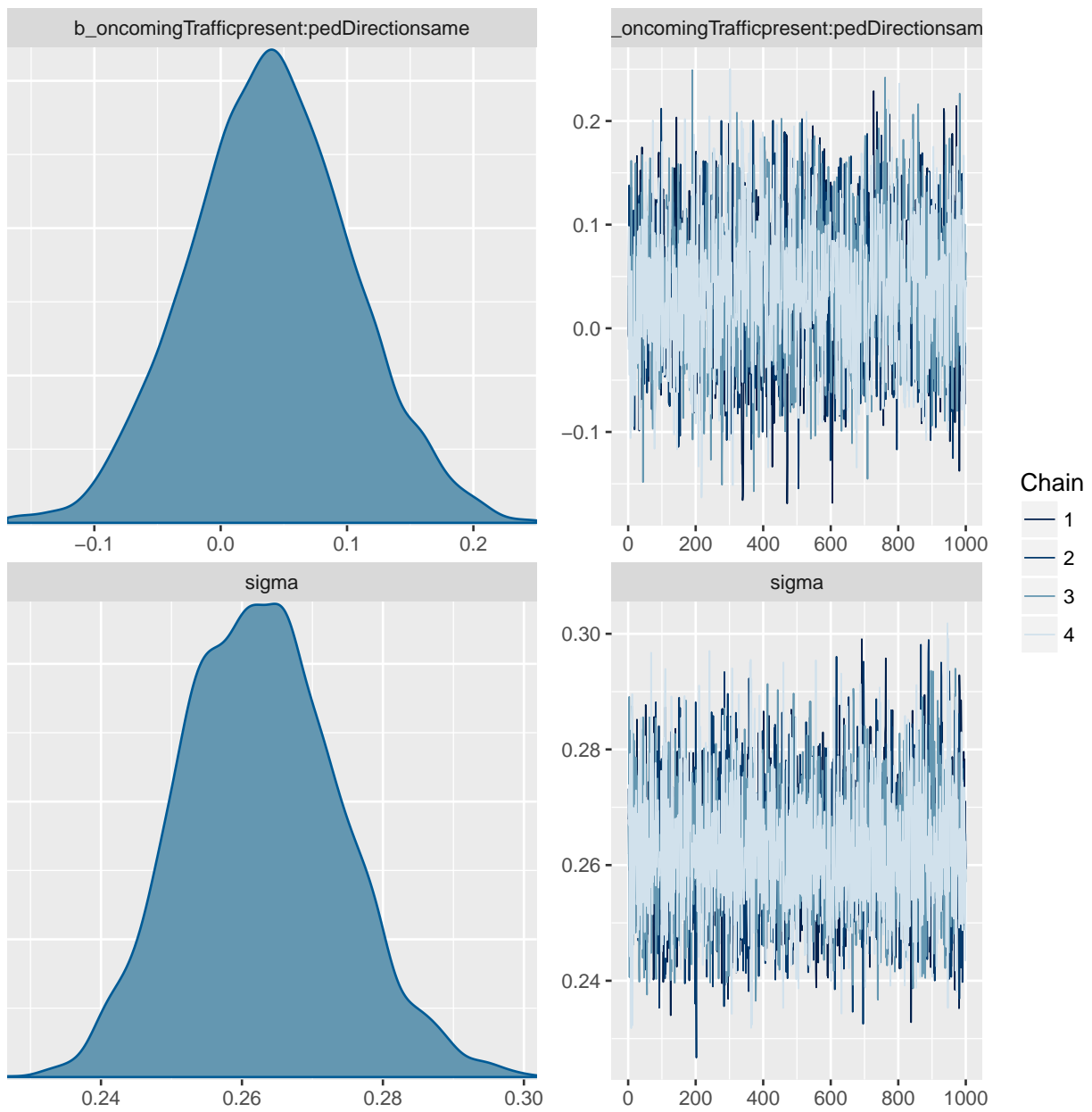


Figure B.2: MC model fit from field test data, part 2

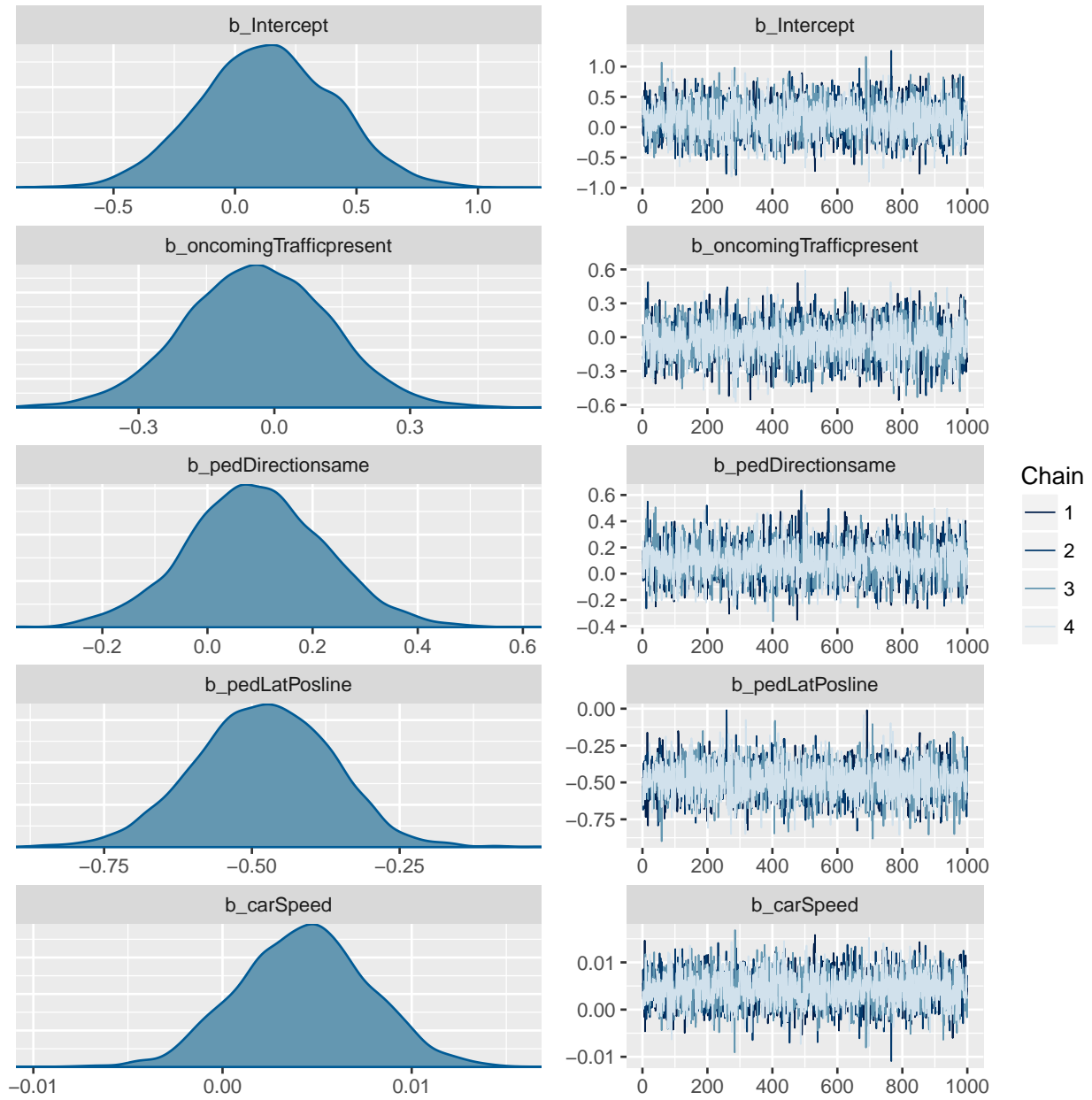


Figure B.3: MC model fit from UDRIVE data, part 1

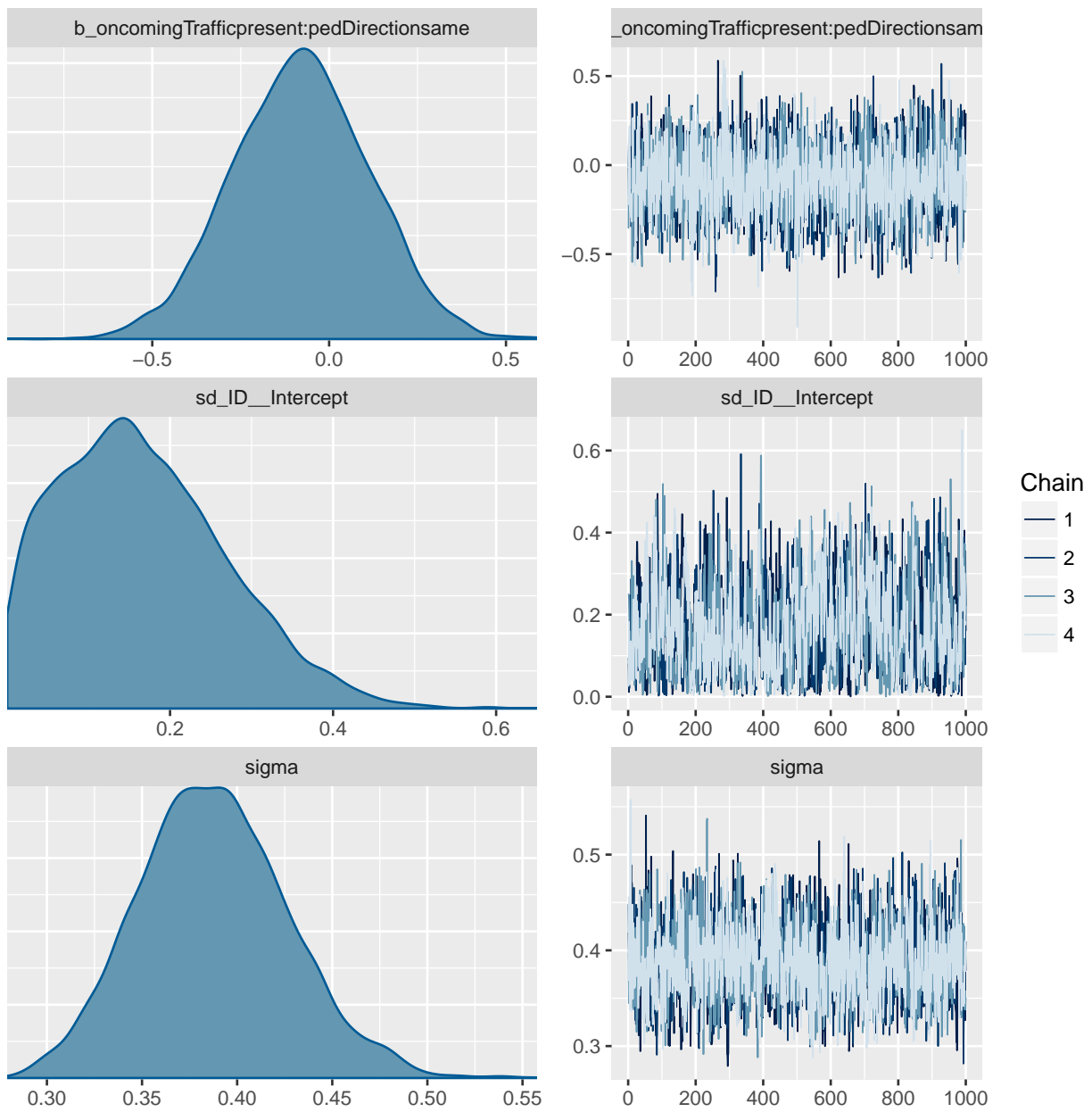


Figure B.4: MC model fit from UDRIVE data, part 2

## B.2 Time to collision model

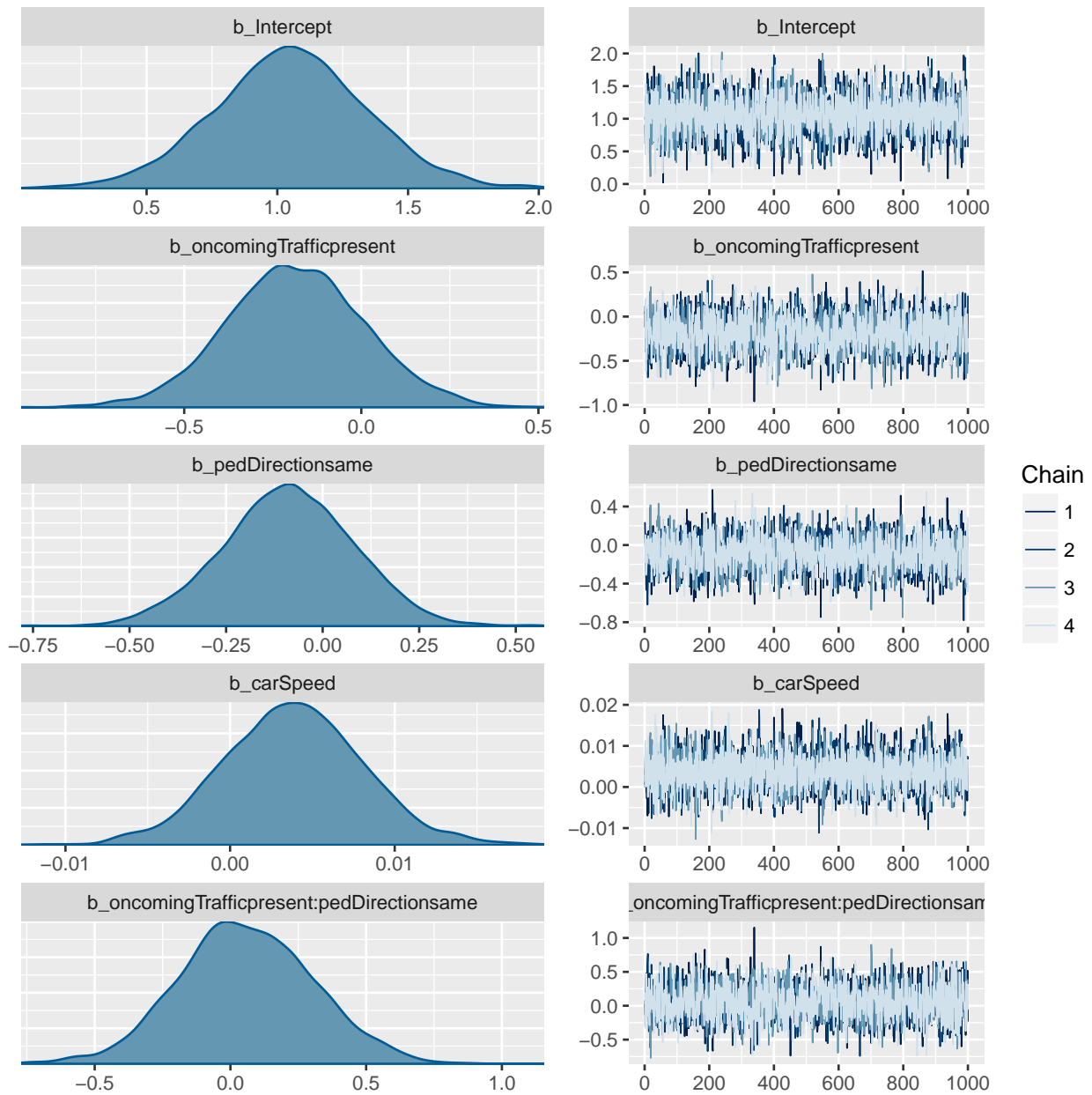


Figure B.5: TTC model fit from UDRIVE data, part 1

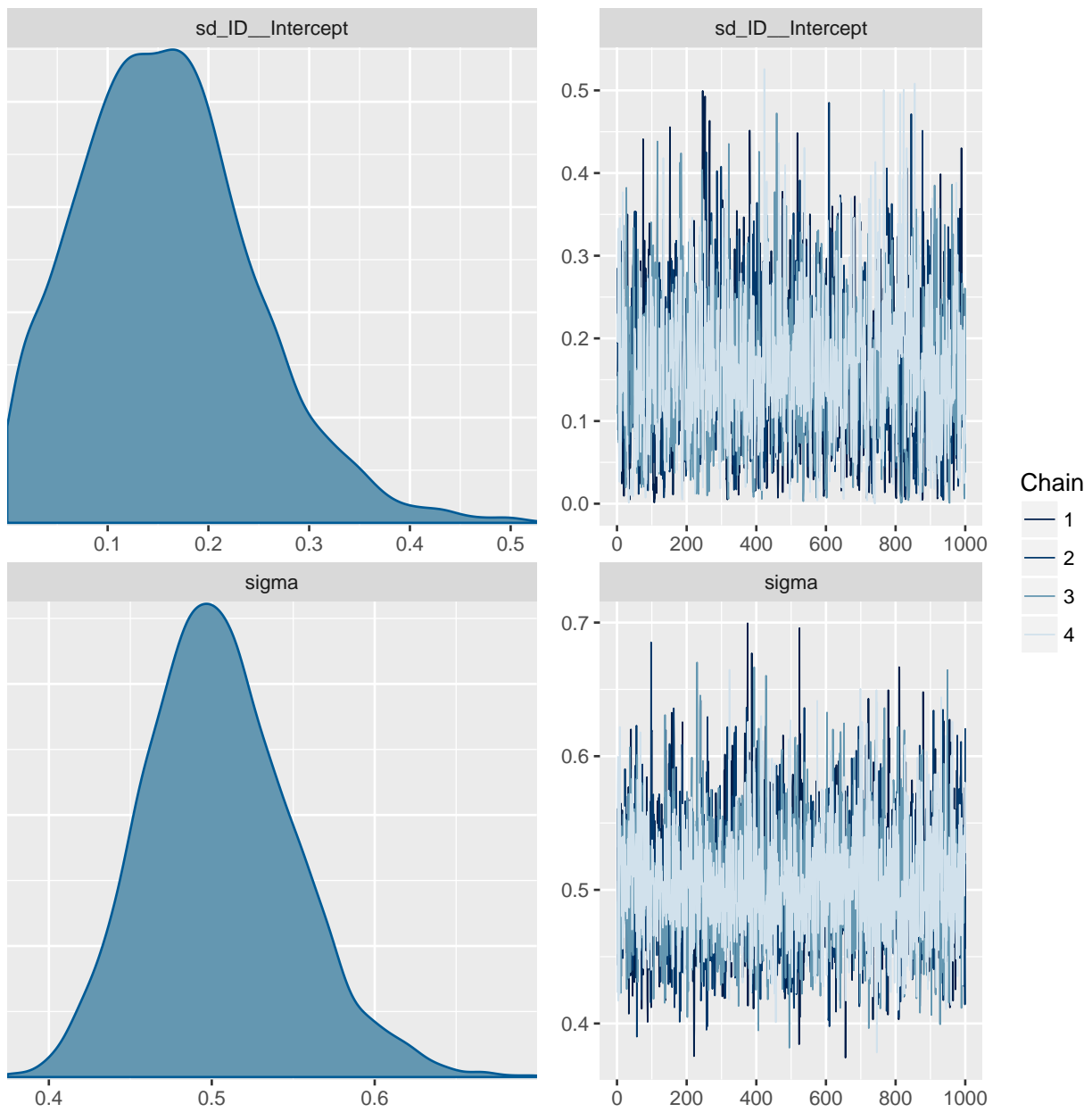


Figure B.6: *TTC model fit from UDRIVE data, part 2*

### B.3 Vehicle speed model

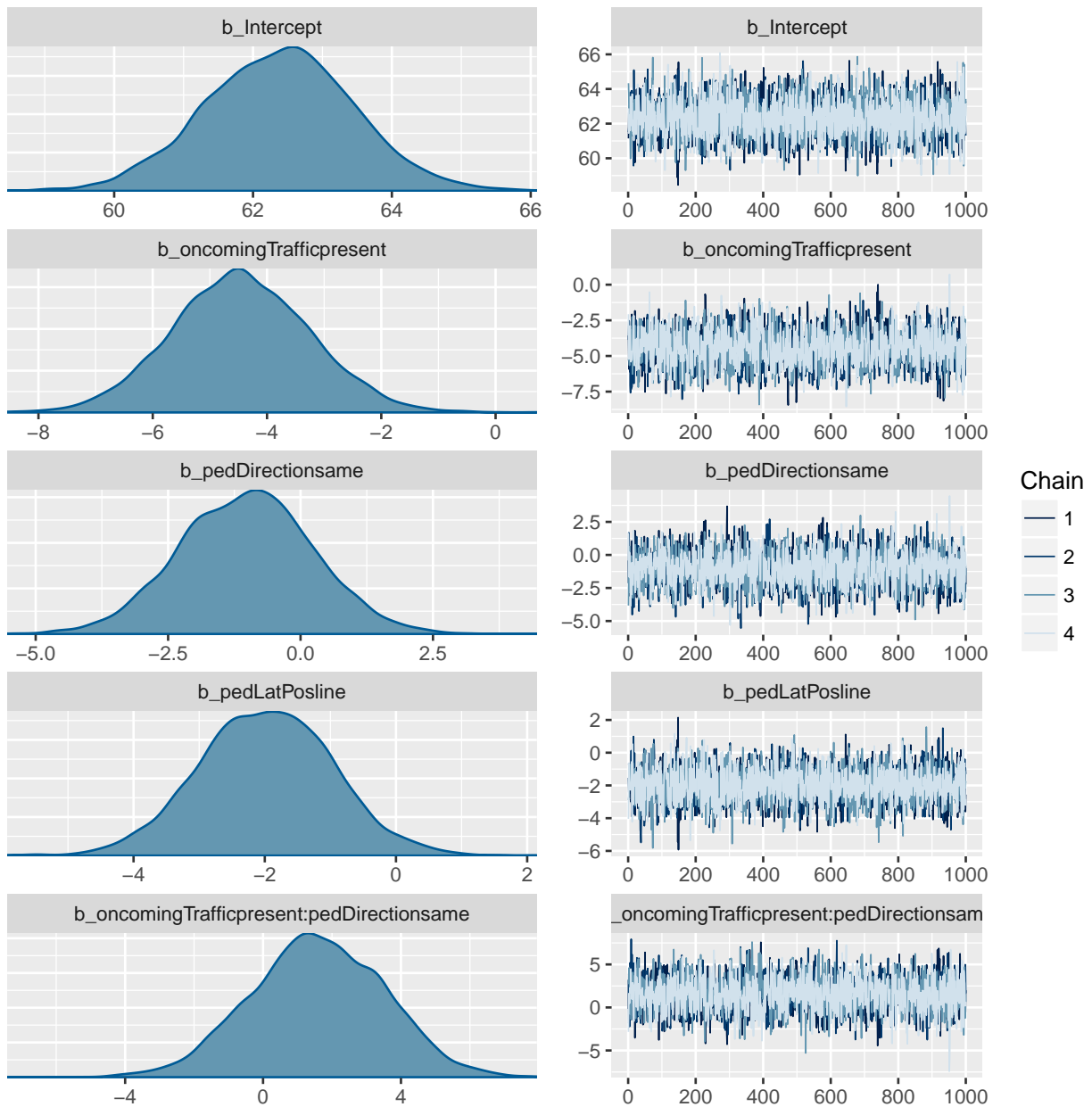


Figure B.7: Speed model fit from field test data, part 1

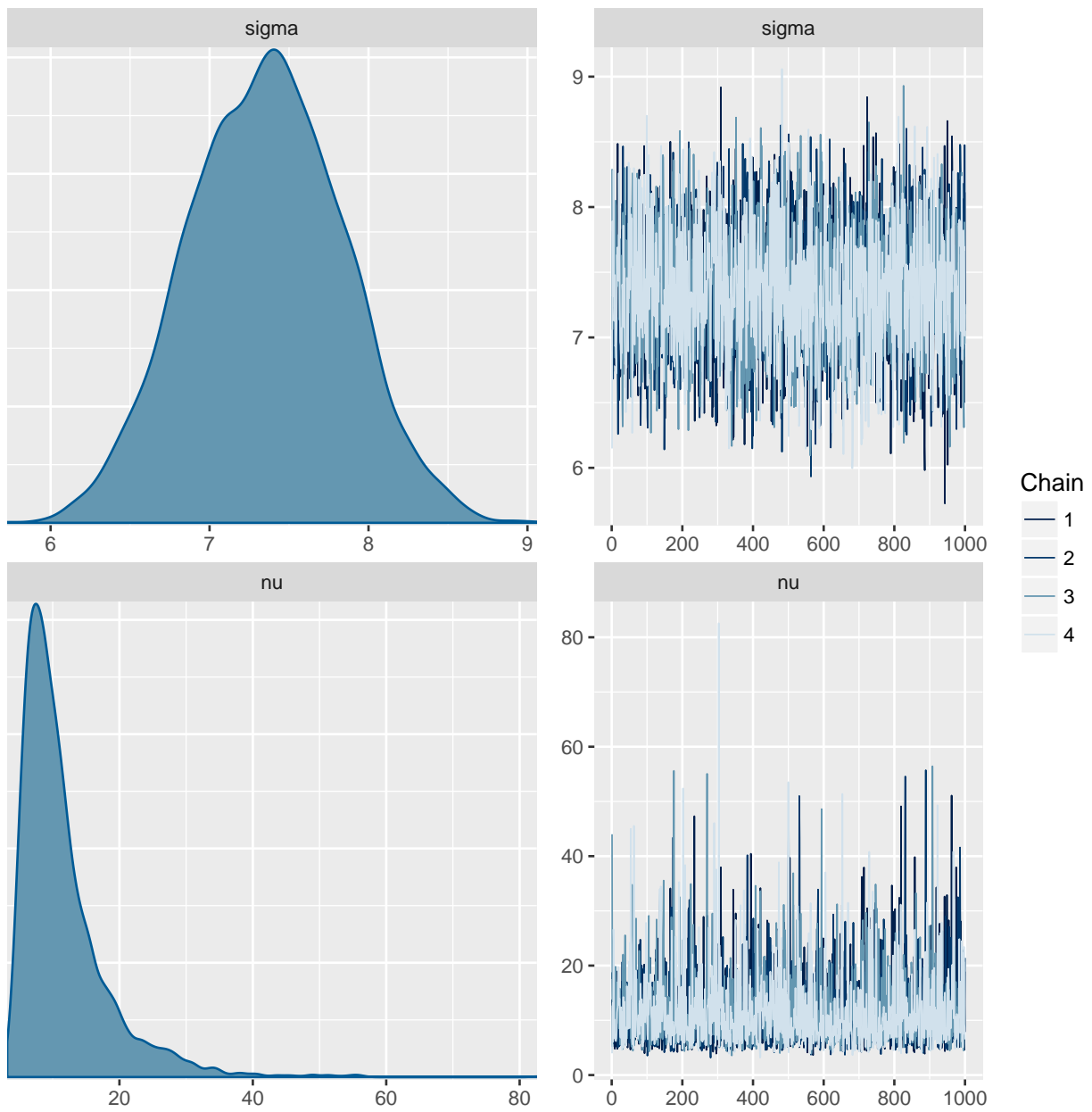


Figure B.8: *Speed model fit from field test data, part 2*