

A Ph.D. Dissertation on

TRAJECTORY-BASED PROACTIVE SAFETY ASSESSMENT OF ROAD TRAFFIC

Submitted by

Suvin P. Venthuruthiyil

(Roll No. 166104028)

Under the guidance of

Mallikarjuna Chunchu



CIVIL ENGINEERING DEPARTMENT
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI, ASSAM, INDIA- 781039
October, 2021



Copyright© 2021 Suvin P. Venthuruthiyil
All Rights Reserved.

Certificate

This is to certify that the work presented in the thesis entitled “**Trajectory-Based Proactive Safety Assessment of Road Traffic**” is carried out by **Mr. Suvin P. Venthuruthiyil** for the award of Doctor of Philosophy in the Department of Civil Engineering from Indian Institute of Technology Guwahati. The thesis embodies result of original work and studies carried out by student himself under my supervision and has not been submitted elsewhere for the degree.

Mallikarjuna C.

Professor

Department of Civil Engineering

Indian Institute of Technology Guwahati

Date: Saturday 23rd October, 2021

Acknowledgements

Finally, the time has come to write the first page of this dissertation. The most read page by family, colleagues, and friends, looking for their own name. Well, I hope I will not forget anybody to thank and that you find your name when you expect it to be here.

This dissertation would not have been possible without the support of a number of people who have contributed in some way to the success of this research project. It's been a long and exhausting journey, but one that has also been extremely enlightening. I would like to express my profound gratitude to my supervisor Prof. Mallikarjuna Chunchu, for his inspiration, invaluable guidance, and continuous encouragement during this work. His suggestions, comments, and most importantly, his patience with a beginner like me improved my study experience and made it less difficult. Thank you for your genuine care about my well-being and challenging me to take a broader perspective on my research. I consider myself blessed to have him as a mentor.

I am sincerely grateful to the chairman of my Doctoral Committee, Prof. Sreedeeep S., for his constructive suggestions and timely help during the entire period of this research work. I would also like to express my gratitude to the Doctoral Committee members, Dr. Anjan Kumar S., and Dr. Rajesh A., for their fruitful discussions and invaluable suggestions during the research process.

My stay at IIT Guwahati was more comfortable since you were around me. Thank you, Dr. Sreeja P., Dr. Suresh Kartha, Dr. Ravi K., Dr. Archana M, Prof. Vibin Ramakrishnan, Prof. A. B. Kunnumakkara, Dr. Resmi Suresh, Dr. John Jose, Dr. Benny George K., Dr. Tony Jacob, Dr. John Thomas, Prof. P. S. Robi, Prof. R. Ganesh Narayanan, Prof. Poullose Poullose, Dr. Ganesh Natarajan, Prof. Anamika Barua, Prof. Sajith, G, Mrs. Anusha. I wouldn't be able to feel IIT Guwahati as my second home without you all around.

Civil Engineering Department offered me several warm moments to cherish. I am very much thankful to the faculty members and staff for the positive environment. I sincerely acknowledge the Indian Institute of Technology Guwahati for providing me the opportunity to study at IITG.

I thank my research group members Pranab, Bharat, Regulus, Dr. Syed Omar, Dr. Dibyendu Pal, with whom I have shared my moments of anxiety and excitement. I wish to thank all my friends at IIT Guwahati, Aparimita, Chandan, Anjaneyulu, Ibaiahun Nongbet Sohlang, Sanhita, Arunabha, Suresh Nama, Shantanu, Nishant, Bhaskar, Saswathi. I would also like to thank all my colleagues in the field of transportation research, Vinayaraj, Avani, Richa, Rushikesh, Bhupali, Irene, Agnivesh, for all the fruitful conversations we had during our meetings.

Thank you, Vishnu Venugopal, Reshma, Dileep P., Jomi Dileep, Jiya, Mohamed Sajeer, Ayesha Wafa, Vivek Lukose, Anoop Balachandran, Jiss Joseph, Arun Mathew, Anu, Vasudevan, Aswathi, Hrishi, Aswin, Jith, Gadha, Christy, Riya, Rahul Nair, Vijith, Mathew Francis, Tinu, Sandeep, Sonu, Haris, Hijas, Tom, Sudhi, Vivek, Ansel, George, Moses, Akhil, Anjani, Aswanth, Vimal, Geetanjali, Nithin, Robin, Harris, Kashinath, Devasree, Divakar. I enjoyed and celebrated my stay at IITG with you the most.

At this point, I remember and thank all my teachers and well-wishers for showing me the path to my dreams.

I should pay my tribute to my beloved parents and in-laws for giving me the whole strength to continue these studies while staying thousands of miles away. I thank them for their persistent faith in my decisions. Thank you, my ever-supportive siblings, Suveesh and Suvina. You gave me strength with your emotional support during all tough times in my life. Shravan Viswa and Parthiv Krish, my baby nephews, your innocent and cute smile were enough to melt away all my final-year stress. I thank my brother-in-laws Prabeesh, Priyesh, Vinay, and Shinoj and sister-in-laws Swapna and Sumitha for your support. Aditi, your weekend calls made me feel loved.

Finally, thanks to my wife, Amrutha Suresh, I have reached this point in my academic career because of you. She was a constant source of inspiration for me when I was going through a rough patch. I count myself extremely fortunate to have you as a colleague at IITG during my PG and Ph.D. All of the technical discussions we engaged in during our spare time aided my quest for knowledge. I am very well aware that the opportunity to pursue B.Tech., M.Tech., and Ph.D. together at the same institutes is a rare blessing.

Finally, my utmost gratitude should go to Mother Nature for giving us somewhere to live and something to study. All the philosophers, free thinkers, and writers who taught me to see the beauty of nature and life, giving me the breath of life, should be honored.



*This thesis is dedicated to
my beloved family.*



Abstract

The economic and social transformation has led to rapid growth in mobility demand leading to a growth in vehicle usage, which has been further aggravated by ever increasing urban sprawl and new commuting needs. Subsequently, several negative externalities of transportation systems, including congestion, road crashes, and pollution arise, which are considered as fundamental challenges of modern times. According to a recent OECD report [1], the gross annual economic loss from traffic fatalities and injuries is projected to be about \$US 450 billion, or around 2% GDP of OECD countries. Researches have been conducted across the world to achieve zero fatalities or minimal severity on roads. Earlier studies have primarily depended on the historical crash reports for the road safety assessment. However, they were not successful in capturing the actual crash risk due to various limitations of the crash data. The proactive safety analyses, using an indicator that captures easily detectable and frequent unsafe events as a surrogate to the crash, is a better alternative for safety assessment.

Surrogate Safety Measures (SSMs) are widely used for the proactive assessment of potential crash risk. The majority of the existing SSMs are fundamentally designed to capture the rear-end collision risk. However, in reality, the traffic dynamics involve the simultaneous interaction of multiple vehicles on a 2-dimensional (2D) surface, which results in a broad spectrum of collision patterns, such as head-on, side-swipe, rear-end, angled, and run-off-road collisions. To capture the overall risk, it is essential to reckon all crash occurrences, irrespective of their patterns. The proactive safety analyses, with an indicator that can measure the transport facilities' overall safety, could allow effectively devising the road crash mitigation strategies to accomplish the zero fatality goal.

Extensive but accurate trajectory data is a vital input for the proactive safety assessment. The recent advances in computer vision and drone technology have inspired many researchers to collect traffic videos to understand naturalistic driving behavior. However, the transformation of the traffic videos into useful traffic data is challenging. Although several computer vision-based data extraction tools are available, many involve manual processing, or costly, or restricted to certain field conditions or road geometries. The present study focuses on remediating the current challenges of video-based microscopic traffic data extraction by developing a microscopic traffic data extraction tool named Semi-Automated VEHICLE TRAjectory eXtractor (SAVETRAX). This tool would enable the development of a vast trajectory database corresponding to various traffic facilities.

Nevertheless, deriving useful information from the observed trajectory data requires certain pre-

processing to handle the noises added at several stages of the data collection. The present study proposes an adaptive and data-driven path smoothing technique that can be universally applied to the video-based path data obtained from any camera platform (either drone or fixed camera setup). The proposed reconstruction framework works in three stages; in the first stage, the data is prepared by resampling any missing data. Also, any outlier in the observed path data is identified and removed at this stage. In the second stage, a Recursively Ensembled Low-Pass (RELP) filter is proposed to handle the ‘heavily tailed’ noise found in the video-processing-based trajectory data, particularly for drone videos. A robust adaptive Gaussian kernel smoothing is applied in the third stage to have a localized reconstruction. For Kernel smoothing, the smoothing parameters such as the optimal bandwidth and polynomial order are estimated using the proposed grid-search algorithm. The parameter estimation process ensures that the bias and variance are perfectly traded-off to achieve a smooth vehicle path. Since the noise pattern is different for the vehicle path’s longitudinal and lateral components, the proposed method applies separate smoothing for these components. Notably, the proposed approach does not cause any over/under-smoothing. Therefore, the critical traffic incidents are not misinterpreted, which is crucial for proactive safety assessments.

An extensive and reliable trajectory dataset would be beneficial for a proactive safety assessment at various road facilities. This study proposes a novel 2D SSM called Anticipated Collision Time (ACT) to capture the overall risk. This study also introduces another safety indicator derived from the ACT profile called Time of Evasive Action (TEA). TEA primarily captures the time at which a vehicle commences to respond, in terms of deceleration, when it encountered an unsafe situation. Such a measure helps to understand the response pattern of different vehicles/drivers to a potential collision. We also derived Time Exposed ACT (TE-ACT) and Time Integrated ACT (TI-ACT) from the ACT profile to capture the crash exposure and severity, respectively.

The potential of the proposed tool and measures, developed as a part of this research, was investigated through four case studies. First, the safety during a lane change operation was estimated. To this end, we developed a methodology to identify and classify interrupted and uninterrupted lane changes. We found that these two types of lane change operations are characteristically different. Further, the Powered Two Wheelers (PTWs) safety in an urban environment was analyzed with trajectory data collected from a busy urban midblock section. The analyses have shown that the likelihood of PTWs involving in a crash is double that of any other vehicle type. PTWs are found more vulnerable to side-swipe collisions. The analysis of vehicles’ safety at horizontal curves superimposed with vertical alignments revealed that the vehicles are more vulnerable to run-off-road crashes at such geometries; thus, installing adequate crash barriers or warning systems is necessary to reduce fatalities. The correlation analysis of different crash types with various road geometric variables shows that the design parameters exhibit a contradicting correlation when modeling different crash types. Further, a network-level proactive safety assessment was performed. The results indicate that it is essential to properly integrate the crash frequency, exposure, and severity for identifying a hot-spot on a transportation network.

The results emphasize the capabilities of ACT and the other derived indicators to perform a more comprehensive proactive safety assessment. It is apparent from the analyses that all the tools

and measures developed in this research opens a new avenue for a comprehensive investigation of the safety of various transport facilities, irrespective of the geometry and the traffic scenario.

This study has certain limitations that can be investigated in future studies. First, ACT's ability to quantify crash risk at various transportation network components is not validated. This needs comprehensive trajectory data at various transportation network components, which would soon be available thanks to technological advances. Second, the impact of trajectory data accuracy on the crash risk statistics is unknown, which needs to be further investigated. Third, the trajectory data sampling frequency might influence the derived quantities such as TE-ACT, TI-ACT, and TEA, which need to be examined. Fourth, the compatibility of ACT with simulation packages and collision warning systems has to be researched further. The outcomes of this study may serve as a springboard for future CAV technologies to improve their safety and efficiency. A detailed investigation is required to validate and adjust the proposed technique to make it deployable, which is the future scope of this project.



Contents

Certificate	i
Acknowledgements	ii
Abstract	vi
List of Figures	xiv
List of Tables	xix
List of Abbreviations	xxi
List of Symbols	xxiii
1 Introduction	1
1.1 General	1
1.2 Objectives of the study	3
1.3 Research Scope	4
1.4 Organization of the Thesis	4
2 Literature Review	6
2.1 Importance of Naturalistic Driving Data in Traffic Engineering	6
2.2 Image-Based Trajectory Data Collection: History, Challenges, and the Current Practices	7
2.2.1 Challenges of Video based Traffic Data Collection	10
2.2.2 Existing Tools for Image-Based Trajectory Data Extraction	12

2.3	Trajectory Data Reconstruction	13
2.3.1	Properties of a Meaningful Vehicle Trajectory	13
2.3.2	Various Trajectory Reconstruction Techniques Generally Practiced in Traffic Engineering	15
2.3.3	Importance of a Separate Smoothing for Lateral and Longitudinal Components of a Vehicle's Path	21
2.4	Trajectory-Based Proactive Safety Analysis	21
2.4.1	Review of Existing Surrogate Safety Indicators	22
2.5	Scenario-Specific Proactive Safety Analysis Using Trajectory Data	29
2.5.1	Lane Change Dynamics and the Associated Risk	30
2.5.2	Proactive Safety Assessment of Powered Two Wheeler (PTW)	32
2.5.3	Safety of Vehicles Traversing the 3D road Geometries	33
2.5.4	Transportation Network-Level Safety Analysis	35
2.6	Summary	36
3	SAVETRAX: A Tool for Trajectory Data Extraction	38
3.1	Introduction	38
3.2	Semi Automated Vehicle Trajectory Extractor (<i>SAVETRAX</i>)	39
3.2.1	Components of <i>SAVETRAX</i>	39
3.2.1.1	Pre-Processor	39
3.2.1.2	The Tracker Module	42
3.2.1.3	The Analyzer Module	44
3.3	Performance Evaluation	47
3.4	Summary	49
4	Vehicle Path Reconstruction using Recursively Ensembled Low-pass Filter (REL_P) and Adaptive Tri-Cubic Kernel Smoother	52
4.1	Introduction	52
4.2	Three Stage Method of Vehicle Path Reconstruction	53
4.2.1	Re-sampling the Time-Series Path Data	54
4.2.2	Recursively Ensembled Low-Pass Filtering (REL _P)	54

4.2.3	Adaptive Tri-Cubic Kernel Smoothing	60
4.2.3.1	Grid-Search Algorithm for Estimating the Optimal Kernel Parameters	62
4.3	Performance Evaluation of the Proposed Reconstruction Technique	64
4.3.1	Retrieval of Missing Data	64
4.3.2	Reconstruction of Observed Vehicle Path	65
4.3.3	Residual Analysis	69
4.3.4	Consistency Analyses	72
4.4	Summary	74
5	Anticipated Collision Time (ACT)	77
5.1	Introduction	77
5.2	The Anticipated Collision Time	78
5.2.1	Procedure for Orienting Vehicles, Distinguishing the Vehicles on a Collision Course, and Identifying the Collision Type	80
5.2.1.1	Orienting the Vehicle based on the Heading Angle	80
5.2.1.2	Estimating the Shortest Distance Between the Vehicles	81
5.2.1.3	Checking whether the Vehicles are on a Collision Course	83
5.2.1.4	Determination of the Collision Type	85
5.2.2	Other Crash Surrogates Derived from ACT-Profile	85
5.2.2.1	Time of Evasive Action (TEA)	86
5.2.2.2	Time-Exposed and Time-Integrated ACT	87
5.2.3	Comparison of ACT with TTC	88
5.2.4	Comparison of Actual and Estimated Crash Frequencies	89
5.3	Summary and Conclusion	90
6	Application of Proposed Trajectory-Based Surrogate Safety Analysis at Various Traffic Scenarios	92
6.1	Microscopic Outlook of the Lane-Change Dynamics	92
6.1.1	Data Preparation and Filtering	93
6.1.2	Identification of the Time Window Corresponding to Lane Change Operation	94

6.1.3	Classification of LC Trajectories into ILC and ULC	96
6.1.4	Characteristics of Interrupted and Uninterrupted Lane-Changes	98
6.1.4.1	Risk and Discomfort Associated to ILC and ULC	98
6.1.4.2	Duration of ILC and ULC	102
6.1.5	Summary and Conclusion	107
6.2	Analysis of Powered Two Wheeler Safety in Urban Environment Using ACT . . .	109
6.2.1	Data	109
6.2.2	Results	109
6.2.2.1	Analysis of Time of Evasive Action (TEA)	111
6.2.2.2	Analyses of Crash Exposure and Severity with an Emphasis on the Importance of ACT*	113
6.2.2.3	Comparison of Various Crash Types	115
6.2.3	Summary and Conclusion	116
6.3	Proactive Safety Assessment at 3D Road Geometries	118
6.3.1	Data	118
6.3.2	Results	120
6.3.3	Summary	123
6.4	Network-Level Proactive Safety Assessment for Identification of High-Risk Spots .	125
6.4.1	Data	125
6.4.2	Results	126
6.4.3	Summary and Conclusions	127
7	Summary and Recommendations	129
7.1	Summary	129
7.2	Major Contributions	131
7.3	Future Scope	132
	References	133
	List of Publications	153

Curriculum Vitae**155**

List of Figures

1.1	Global status report of road safety (<i>Source: World Health Organization [2]</i>)	2
1.2	Thesis structure and relations between chapters	5
2.1	Review of trajectory based studies in the recent past; a) Data platform; b) Data extraction method	9
2.2	Speed and acceleration distributions estimated from the observed data (extracted using Traffic Data Extractor (TDE) software [3]); a) Frequency distribution of the speeds estimated from the observed trajectory data; b) Frequency distribution of the accelerations estimated from the observed trajectory data	14
2.3	Observed acceleration distribution from drone and fixed camera setup; a) Acceleration distribution for drone based traffic data collection; b) Acceleration distribution for fixed camera setup (NGSIM data)	20
2.4	Field observed lateral position, speed, and acceleration profiles corresponding to LC maneuver (NGSIM, US-101 data); a & d) Lateral position profiles; b & e) lateral speed profiles; c & f) lateral acceleration profiles; corresponding to vehicle IDs Car-242 and Car-1210, respectively	31
3.1	Screen short of SAVETRAX	40
3.2	Point-point correspondence between real world and image plane; a) plain terrain type; b) non-planar terrain type	40
3.3	Comparison of the original and the stabilized camera paths	46
3.4	Result of smoothing of tracked data	47
3.5	Experimental setup for validating the accuracy of SAVETRAX; a) test bed and the instrumented vehicle; b) V-Box	48
3.6	Comparison of speed and acceleration obtained from V-Box and SAVETRAX; a) speed; b) acceleration	48
3.7	The test road passing though non-planar terrain	49

3.8	Comparison of the speed obtained from V-Box and from SAVETRAX corresponding to single cell (C) homography and multiple homography conditions; a, b, & c) Comparison of the speed obtained using single homography (at different locations on the curve) with the V-Box data; d) Comparison of the speed obtained using multiple homography and from V-Box	50
3.9	Performance of SAVETRAX at various traffic, camera, and lighting conditions; a) at different traffic states; b) with glare; c) bad weather; d) low light condition; e) bird-eye view (a-d corresponds to various camera angles)	51
4.1	Algorithm flow chart for the proposed vehicle path reconstruction technique	53
4.2	Algorithm flow chart of missing data retrieval process	54
4.3	Illustration of S-G filter; • denotes the input samples; ● denotes the filtered sample; ○ denotes the effective impulse response sample (weighting constant); and * denoted the convolution operation	56
4.4	Algorithm flow chart of RELP	59
4.5	Graphical representation of the Grid-search algorithm	64
4.6	Performance evaluation of the missing data retrieval process; a) Manually extracted NGSIM data reported by Coifman et al. [4], with the synthesized occlusion conditions; b) Comparison of different interpolation techniques for the missing data retrieval	65
4.7	Performance evaluation of the proposed reconstruction technique using the data from a controlled experiment; a) Comparison of the reconstructed path and the observed path; b) Comparison of reconstructed speed with the benchmark speed	66
4.8	Physical consistency analysis of the estimated acceleration; a) Distribution of acceleration after RELP (ranges between $-50 m/s^2$ to $50 m/s^2$); b) Distribution of acceleration after Kernel smoothing (ranges between $-3.5 m/s^2$ to $2.5 m/s^2$); c) Cumulative Frequency (in %) distribution after RELP and Kernel smoothing	66
4.9	Increasing benefit of the proposed reconstruction technique, a) Performance of RELP in denoising the gradient of vehicle path b) Improvement in speed after each stage of reconstruction, c) Improvement in acceleration after each stage of reconstruction	67
4.10	The performance of the proposed method on the NGSIM I-80 dataset presented on a sample trajectory (Vehicle ID 1929) and a comparison with the method proposed by Montanino et al. [5] (a) Trajectory, (b) Speed profile, and (c) acceleration profile.	67
4.11	A 90 degree near-crash, a) Vehicles involved in the near-crash (circled in the figure), b) Piled-up vehicles due to the bottleneck created by the near-crash (highlighted in the figure)	68

4.12	Position, speed, and acceleration profiles of the vehicle involved in near-crash (PTW 1) and the piled-up vehicles (Car 1, Car 2, Car 3, and Car 4), (a,b,c) Observed position, speed, and acceleration profiles, (d,e,f) Reconstructed position, speed, and acceleration profiles.	68
4.13	User interface of MATLAB application developed for vehicle path smoothing	69
4.14	Distribution of residuals and a fitted Gaussian function	70
4.15	a) Time-series of mean error per frame; b) Partial auto-correlation analysis	72
4.16	Internal consistency analysis of the proposed method and comparison with some of the off-the-shelf smoothing techniques	74
4.17	Physical consistency and Frequency spectrum analysis of acceleration	75
5.1	Simplified illustration of the concept of Anticipated Time to Collision (ACT); a) Closing-in of two vehicles till collision; b) Factors influencing the closing-in rate of two vehicles	78
5.2	The state of a vehicle i at a time instance t in GCS	80
5.3	Procedure for orienting the vehicles in the 2D space according to the heading angle; a) Base orientation; b) Rotation; c) Translation	81
5.4	Estimation of shortest distance between vehicles; a) a general orientation of vehicles; b) Algorithm flowchart of closest point estimation	82
5.5	State configuration of vehicles involved in angled collision scenario; a) Vehicles are closing-in; b) Vehicles are shying away	83
5.6	State configuration of vehicles in rear-end collision scenario; a) Conflict scenario of overlapping vehicles with parallel center-line; b) Conflict scenario for non-overlapping vehicles with angled center-line; c) Non-conflict condition for overlapping vehicles with angled center-line; d) Criteria for rear-end conflict	84
5.7	Estimation of TEA from ACT and acceleration profiles; a) ACT profile with safe and unsafe regions separated by ACT^* ; b) Acceleration profile showing acceleration and deceleration regions	87
5.8	Estimation of TE-ACT and TI-ACT from ACT profile	88
5.9	Comparison of distributions of crash exposure and severity estimated from ACT and TTC; a) Distribution of duration crash exposure derived from ACT and TTC; b) Distribution of crash severity derived from ACT and TTC	89
5.10	Estimation of the ACT-threshold using the simultaneous optimization; a) Error minimization; b) Comparison of the actual and estimated conflicts using the optimal ACT-thresholds.	90

6.1	Locations of the trajectory data collection	94
6.2	Method of lateral speed segmentation for LC window identification	95
6.3	a) An example of LC window detected using the proposed method; b) Normalized lateral position profiles of a few vehicles within the LC window (ULC (black), ILC (red))	96
6.4	Schematic representation LC interruptions reflected in the lateral speed profile	96
6.5	Lane change trajectories corresponding to; a) ULC; b) ILC	97
6.6	Histogram of a) TE-ACT; b) TI-ACT; and c) Acceleration Noise	98
6.7	a) Average number of sign changes in the lateral and longitudinal acceleration profiles; b) Mode of number of sign changes in the lateral and longitudinal acceleration profiles	101
6.8	LC duration distribution; a) ILC; b) ULC; c) probability plot of Log-Normal distribution of ILC duration; d) probability plot of Log-Normal distribution of ULC duration	102
6.9	Schematic diagram of a lane-changing process	103
6.10	Sensitivity analysis of explanatory variables of LC duration models	107
6.11	The number of fatalities among Vulnerable Road Users around the world (<i>Source: World Health Organization [6]</i>)	110
6.12	Study area	110
6.13	ACT-histogram of PTWs and other vehicles	111
6.14	ACT profile of PTWs (black line). The thick black line portions in the profile indicate the unsafe region. The red line shows the ACT threshold;	112
6.15	Comparison of ACT, speed, and acceleration profiles; (a) and (b) are the profiles corresponding to different PTWs	112
6.16	Statistical comparison of TEA for cars and PTWs using two sample t-test	114
6.17	Comparison of collision risk of PTWs and other vehicle classes; a) Duration of exposition to safety critic situation (TE-ACT); b) severity of the collision (TI-ACT)	114
6.18	The impact of threshold value (ACT*) on TE-ACT and TI-ACT	115
6.19	Analysis of the occurrence of different crash types for PTWs; a) Frequency of different crash types; b) Exposure time to different crash types; c) Severity of different crash types	116
6.20	Some of the curve sections where data is collected	119

6.21	Details of the reported crash events for the study area from December 2013 to September 2018 corresponding to curves; a) Relative frequency of the occurrence of different crash types; b) Severity of crashes; c) Vehicles involved in crashes	120
6.22	Cumulative probability density crash exposure ad crash severity for different crash types; a) TET; b) TIT	120
6.23	Correlation of TE-ACT and TI-ACT with geometric variables; a) Angled collision, TE-ACT; b) Angled collision, TI-ACT; c) Rear-end, TE-ACT; d) Rear-end, TI-ACT; e) Side-swipe, TE-ACT; f) Side-swipe, TE-ACT; g) ROR, TE-ACT; h) ROR, TI-ACT	121
6.25	Comparison of relative likelihood of observed and estimated crashes	123
6.24	Spatial distribution of the crash frequency along the curve; first and second rows corresponding to horizontal curve superimposed with a constant grade and third and fourth corresponding to horizontal curve superimposed with vertical curve; Each column represents a crash type, such as the angled collision, rear-end collision, side-swipe, and ROR, respectively	124
6.26	Study area and the drone coverage	125
6.27	pNEUMA Trajectory data	126
6.28	Spatial distribution of conflict frequency	127
6.29	Spatial distribution of exposure and severity of potential crashes	128

List of Tables

2.1	Review of the image-processing based trajectory extraction tools	12
2.2	Review of trajectory smoothing techniques employed in the transportation literature	16
2.3	List of surrogate safety indicators that measure temporal-proximity of vehicles . . .	24
2.4	List of surrogate safety indicators that measure spatial-proximity of vehicles	26
2.5	List of surrogate safety indicators based on the deceleration behavior	27
2.6	Parameters used to identify LC window	31
2.7	LC duration reported in various studies	33
4.1	Results of t-test to verify the zero mean of residuals	71
4.2	Residual statistics and the Normality test results	71
4.3	Test for equal variance of residuals in longitudinal and lateral directions	73
5.1	Criteria for classification of conflict types	86
5.2	Statistical comparison of duration of crash exposure and crash severity estimated using TTC and ACT using two sample t-test	89
6.1	Probability Distribution of various LC attributes capturing safety and comfort . . .	100
6.2	The Kruskal-Wallis ANOVA test statistics	101
6.3	Descriptive statistics of variables influencing LC duration for ULC and ILC	104
6.4	Statistics of LC duration models	105
6.5	Descriptive statistics of TEA estimated for cars and PTWs	113
6.6	Statistical comparison of TEA corresponding to PTWs and Cars using two sample t-test	114
6.7	Descriptive statistics of the crash exposure and severity estimated using TTC and ACT	116

6.8 Geometry of the curves considered in the study	119
--	-----



List of Abbreviations

1D	1-Dimensional
2D	2-Dimensional
3D	3-Dimensional
ACT	Anticipated Collision Time
AME	Absolute Mean Error
AN	Acceleration Noise
DLT	Direct Linear Transformation
ECMT	European Conference of Ministers of Transport
ECO	Efficient Correlation Operators
GDP	Gross Domestic Product
GOI	Grid of Interest
GPS	Global Positioning Systems
HSID	Hot Spot Identification
ILC	Interrupted Lane Change
IP	Image Processing
ITS	Intelligent Transportation Systems
LC	Lane Changing
LCV	Light Commercial Vehicle
MSE	Mean Squared Error
ND	Naturalistic Driving
OECD	Organisation for Economic Co-operation and Development
PTW	Powered Two Wheeler
RELP	Recursively Ensembled Low-Pass filter
RMSE	Root Mean Squared Error

ROR	Run-Off-Road
SAVETRAX	Semi-Automated Vehicle TRAjectory eXtractor
SSM	Surrogate Safety Measures
TEA	Time of Evasive Action
TE-ACT	Time Extended ACT
TET	Time Exposed Time-to-Collision
TI-ACT	Time Integrated ACT
TTC	Time-to-Collision
UAV	Unmanned Aerial Vehicle
ULC	Uninterrupted Lane Change
VRU	Vulnerable Road User
WHO	World Health Organization

List of Symbols

ϵ	Error term
$f(\cdot)$	Density
$\mathbf{H}$	Homography matrix
$\mathbf{A}$	Coefficient Matrix of least squared regression
J_d	Continuous T-periodic function of the interpolated feature layer
b_d	Interpolation kernel
$J\{x\}$	Interpolated feature map
f	Continuous T-periodic multichannel convolution filter
d	Feature layer
$p_{d,c}$	Set of learned coefficients
$J\{x\}(t)$	Feature vector
$\hat{z}^d[k]$	Fourier coefficient of the interpolated feature map
λ	Weight parameter
α_j	Exponentially decaying weight
γ	Learning rate
h_{SG}	Bandwidth
f_s	Sampling frequency
ϵ_N	Mean square approximation error
a	Vector of polynomial coefficients
$\tilde{a}$	Approximation of coefficient vector for the unit impulse input $\mathbf{d}$
$\mathbf{d}$	$(2M + 1) \times 1$ Column vector impulse
S_{in}	Input signals
S_{out}	Output signals
G_x	Lateral components of the gradient
G_y	Longitudinal components of the gradient
$\mathcal{X}$	Mtrix of vehicle position coordinates
$f(t_0)$	Regression function
p	Polynomial order
K	Kernel function assigning weight to each data point
$\hat{\beta}_j$	Solution vector of the least square problem
Ψ	Design matrix
$\mathbf{W}$	$n \times n$ Diagonal matrix of weights

α	Solution vector in the feasible solution space Ω
Φ_i	Objective to be minimized
N	Size of the data
h_{opt}	Optimal bandwidth
p_{opt}	Polynomial order
$\hat{\beta}$	Coefficients of local polynomial function
δ	Shortest distance between the approaching vehicles at a time instance t_1
$\frac{\partial \delta}{\partial t}$	Rate at which the vehicles approach each other (closing-in rate)
$\mathcal{R}el$	Operator that takes the vector sum of the quantities
δ	Estimate the shortest distance between the vehicles.
$\mathcal{X}$	Vehicle state data
t	Time
x	Longitudinal position
y	Lateral position
v_x	Longitudinal speeds
v_y	Lateral speeds
a_x	Longitudinal acceleration
a_y	Lateral acceleration
θ	Heading angle
$\dot{\theta}$	Yaw rate
$\mathbf{R}_z$	Rotation matrix
l	Length
b	Width
$\mathbf{P}$	Vector of boundary points of vehicle i
w_l	Width of the leader vehicles
w_f	Width of the follower vehicles
N_j	Length of the trajectory
Δy_j	Maximum lateral displacement for vehicle j ;
w_j	Width of vehicle j ;
l^j	Vector of lane occupancy indicator
Φ	Function that returns the size of the unique values in l^j
v_y	Lateral speed profile
s_l	Start time of the segment
s_m	End time of the segment
v_y^{max}	Maximum lateral speed
τ^*	Time
t_{LC}^{start}	Starting time of the LC operation
t_{LC}^{end}	Ending time of the LC operation
Δt_{LC}^{acc}	Acceleration phase
Δt_{LC}^{dec}	Deceleration phase
Υ_{LC}	Level of interruption

Δt_{LC}	Lane change duration
$\mathbf{X}$	Vector of explanatory variables
Δx_{lead}^{targ}	Gap with the leader vehicle in the target lane
ΔG^{lead}	Gap between the lead vehicle in the current and target lanes



Chapter 1

Introduction

1.1 General

Transportation plays a significant role in the day-to-day life of human beings. A high-quality transportation system allows communities to access essential services, facilities, and employment opportunities, thereby enhances life quality. Efficient transportation is a critical component of every economy since it enables transporting resources from one place to another, allowing people and societies to generate welfare. Despite the benefits, the transportation sector also faces some unprecedented challenges. Some of them include loss of life and property damage due to road crashes, noise and air pollution, and so on.

Road safety is one of the critical challenges faced by designers and policymakers, worldwide. Figure 1.1 shows the global status report of road safety, and it demonstrates the intensity of the road safety issues. Every year, almost 1.2 million people are killed in road crashes around the world. On the other hand, the number of people injured is as high as 50 million – the population of five major cities combined. This alarming number of road fatalities emphasizes that safety management is a crucial component that needs to be strengthened with vision, direction, coordination, management, funding, interventions, advocacy, monitoring, and evaluation. Although the key challenges vary in scope between regions, they are common everywhere.

Despite the fact that road safety has significantly improved in many countries over the last half-century, much more remains to be done to achieve a safer road environment. To bring the road safety issues under control, policymakers and designers must clearly understand the magnitude of the challenge. This should be achieved through assimilating the knowledge from past experiences and close inspection of the real-field data corresponding to individual vehicles. The existing practices, which are either based on the historical crash events or proactive methods that only apply to specific crash types, are inadequate to achieve such an understanding. In fact, studies have widely reported the limitations of the safety assessment using the historical crash data. Nevertheless, ECMT, WHO, and the World Bank have acknowledged that a significant shift in road safety thinking has taken place over the years. Some of them are summarized below [7–10]:

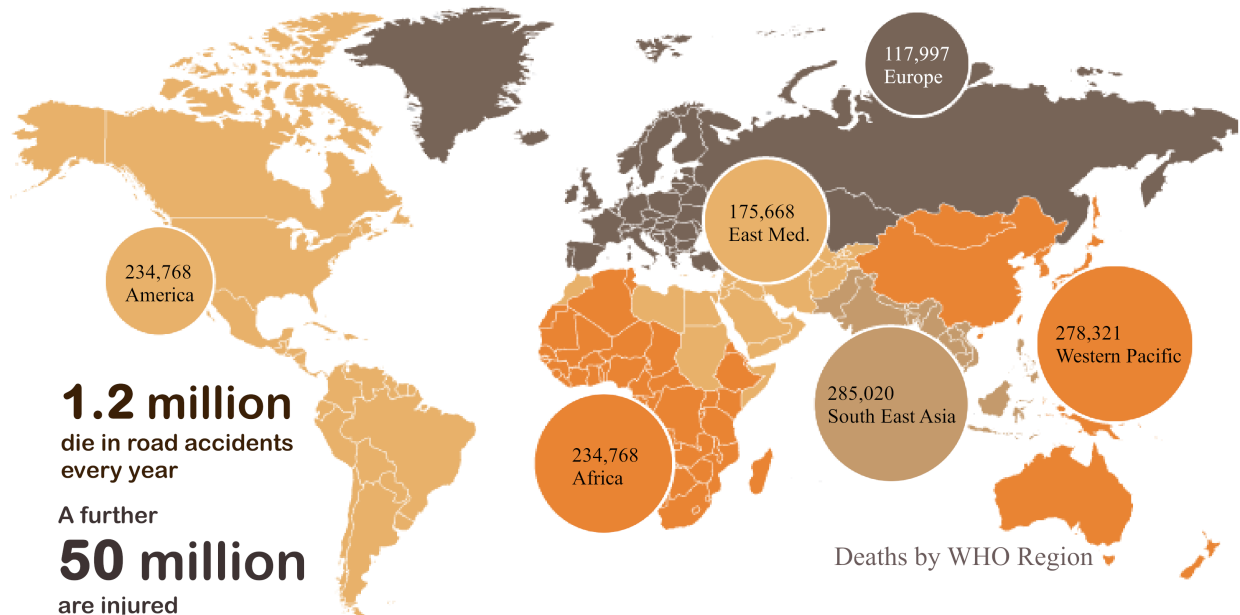


Figure 1.1: Global status report of road safety (Source: World Health Organization [2])

- i Although it is challenging to avoid road crashes entirely, traffic injuries can be predicted and minimized.
- ii Scientific approaches and reliable data are essential for rational analysis and effective remedial action.
- iii Traffic system providers should account the possible human errors while designing the traffic facilities. Also, they should consider the road user vulnerability and acknowledge the importance of speed management.
- iv Road traffic injury is a social equity issue – equal protection to all road users should be provided and avoid an unfair burden of injury and death for vulnerable road users, such as pedestrians and Powered Two Wheeler (PTW).
- v Technology transfer is most effective when it takes into account local issues, environments, and traffic patterns.

These points suggest the need for a data-centric, advanced, but straightforward technique for a strategic solution to road safety issues.

The ever-growing traffic injuries across the world indicate that the existing traffic injury prevention practices are not adequate to accommodate the safety issues on the road. Therefore, it is essential to have a deeper understanding of the mechanism behind the onset of a crash. Here, the challenge is twofold. First, there needs extensive data to understand vehicle-vehicle and vehicle-infrastructure interactions during safe and unsafe circumstances. Second, a more sophisticated safety indicator needs to be devised to capture the overall crash risk at a transportation facility.

As mentioned by several past studies, the safety assessment using the historical crash data known for its limitation. Therefore, researchers have used measures that proxy the actual crash

event - known as Surrogate Safety Measures (SSMs). The existing SSMs are primarily developed to capture the likelihood of specific crash category (mostly rear-end crash), or their application is limited to certain road geometry. Therefore, these measures are not suitable for an overall safety assessment at various road facilities. Due to this limitation, the existing proactive safety assessment can underestimate the actual crash risk at a transportation facility. Thus, this study emphasizes that the proactive safety evaluation necessitates a more sophisticated SSM that can work with different vehicle types, crash types, and facilities.

Vehicle trajectory data is a digital replica of the field dynamics, which helps extract all relevant information on vehicle kinematics and vehicle-to-vehicle interactions. In the context of the safety analysis, such data would help in a comprehensive understanding of the crash generation mechanism and the evasive actions taken by the vehicles. However, owing to the difficulties associated with the data collection, extraction, and standardization, achieving trajectory data for site-specific safety analysis is challenging. Though several trajectory data extraction frameworks are available, many of them are manually operated, costly, or erroneous. Thus, it is essential to have a cost-effective, user-friendly, and accurate data extraction platform for a comprehensive safety investigation at different locations.

One of the widely known issues with the field observed trajectory data is the embedded noise added at the several stages of the data collection and extraction. If not adequately accounted for, this noise would result in unrealistic vehicle kinematic characteristics, which significantly influences the reliability of the inferences made. The recent advances in drone and computer-vision techniques have attracted several researchers to collect the trajectory data using imaging techniques. However, the pattern of the noise in such data is unknown and necessitates special treatment. Therefore, there is a need for a more advanced trajectory reconstruction technique, which is independent of the camera platform and the pattern of the noise.

1.2 Objectives of the study

Motivated by all the facts discussed above, the goal of this study is to perform a proactive safety assessment of road traffic using trajectory data. With this goal, the objectives of the present research work are defined as follows:

1. Development and Standardization of Vehicle Trajectory Database
 - 1.1. Development of an image-based trajectory data extraction tool
 - 1.2. Development of a methodology for the reconstruction of vehicle trajectory data to remove embedded noise
2. Development of a 2-Dimensional surrogate safety indicator for general application
3. Proactive safety assessment at various elements of the transport network
 - 3.1. Lane change safety assessment

- 3.2. Safety of Powered Two Wheelers (PTWs) in a multi-class, weak lane-disciplined traffic stream
- 3.3. Vehicle safety at 3-Dimensional geometric elements
- 3.4. Network-level safety evaluation

1.3 Research Scope

The scope of this research work is as listed below:

1. The trajectory reconstruction method proposed in this study is tested only for the trajectory data collected using imaging techniques. However, the proposed methodology can be used with trajectory data from any source, but the applicability needs to be explored.
2. Safety issues addressed in this thesis are restricted to motion safety. Although the risk due to system failure and the governing factors are relevant aspects of a vehicle's safety, they are out of the scope of this work.
3. All the safety analyses performed in this research work are for motorized vehicles. The pedestrian/non-motorized vehicle's safety is not addressed, and they are outside the scope of this work. Nevertheless, all the contributions of this study are general and can be extended to pedestrian safety also.
4. During lane change analysis, vehicles were expected to follow the lane discipline. So, the LC identification method proposed in this study can only be used with lane discipline traffic conditions. For weak lane disciplined traffic conditions, additional constraints need to be added to identify the lane change events correctly, and it is outside the scope of this work.

1.4 Organization of the Thesis

Figure 1.2 graphically depicts the organization of this thesis. Each box represents a chapter, and the arrow between the boxes indicates the relationship between the chapters. A detailed description of the thesis organization is as follows.

Chapter 2 gives a critical review of the literature to highlight the gaps in the state-of-the-art. This chapter primarily has four sections. Section 2.2 reviews the image-based trajectory extraction tools. Section 2.3 reviews the importance of the noise reduction from the extracted trajectory data. Section 2.4 discusses the proactive safety analysis and the limitation of the existing surrogate safety measures. Section 2.5 presents the literature review corresponding to the four case studies conducted in this study. This includes the lane-change scenario, PTW safety, vehicles' safety at 3D road geometry, and the network-level safety assessment. Section 2.6 summarizes the literature survey and highlights the research gaps.

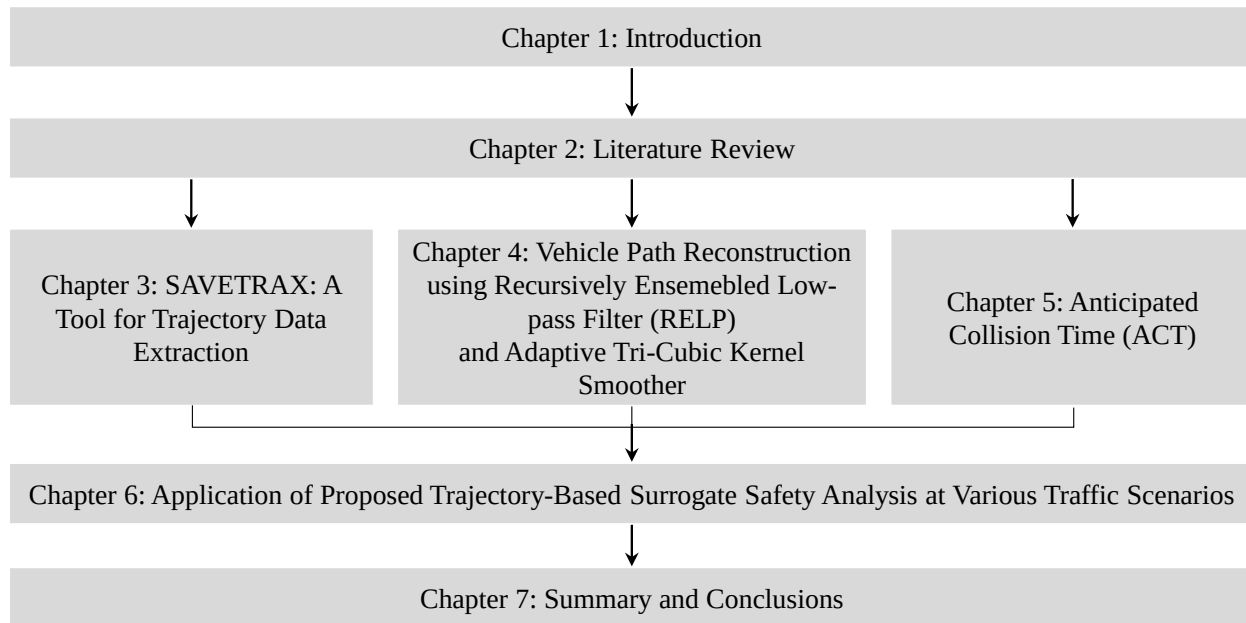


Figure 1.2: Thesis structure and relations between chapters

Chapter 3 presents the trajectory data extraction tool, the *SAVETRAX*, developed as part of this research work. Section 3.2 of this chapter explains the workflow of the developed application and its modules. Section 3.3 presents the evaluation of *SAVETRAX*, and shown the accuracy of trajectory data extracted using *SAVETRAX*. Section 3.4 summarizes the work.

Chapter 4 presents the trajectory reconstruction framework developed in this research work. Section 4.2 gives the details of the proposed reconstruction methodology. Section 4.3 evaluates the efficiency and reliability of the proposed framework. Section 4.4 summarizes the chapter.

Chapter 5 discusses the novel surrogate safety indicator developed as part of this research work. Section 5.2 describes the derivation of ACT and a detailed methodology to estimate ACT. This section also discusses the other SSMs derived from an ACT profile, such as Time Extended ACT (TE-ACT), Time Integrated (TI-ACT), and TEA. Section 5.3 summarizes and concludes the chapter.

Chapter 6 presents the application of the proposed method at various traffic scenarios. Section 6.1 presents the safety assessment of a lane change operation. This section emphasizes the need to consider interrupted and uninterrupted lane changes separately while modeling and analyzing lane change dynamics. Section 6.2 performs the safety evaluation of the PTWs in an urban environment. Section 6.3 presents the safety of vehicles traversing a horizontal curve superimposed with a vertical alignment. Section 6.4 presents a network-level safety evaluation.

Finally, Chapter 7 summarizes the thesis and recommends future research directions.

Chapter 2

Literature Review

This chapter presents a detailed literature review, with an aim to identify the research gaps corresponding to the objective of this thesis work. Section 2.1 highlights the importance of the naturalistic driving data in the empirical traffic flow studies. Section 2.2 reviews the image-based trajectory extraction tools. This section highlights the limitations of the existing trajectory data extraction tools and the need for a new tool. Besides, this section also discusses the desirable properties for a novel trajectory data extraction tool. Section 2.3 reviews the importance of the noise removal from the extracted trajectory data. This section also describes the properties of an accurate trajectory data. Further, different trajectory reconstruction techniques employed in the literature are presented, and highlighted the need for a novel trajectory reconstruction technique. Section 2.4 discusses the proactive safety analysis and the limitations of the existing surrogate safety measures. Section 2.5 presents the literature review corresponding to the four case studies conducted in this study. This includes the lane-change scenario, Powered-Two-Wheeler safety, vehicles' safety at 3D road geometry, and the hotspot identification for the transportation network. The last section summarizes the literature survey and highlights the research gaps.

2.1 Importance of Naturalistic Driving Data in Traffic Engineering

Rapid advances in vehicle position acquisition technologies such as Global Position Systems (GPS), Radio Frequency Identification (RFID), location estimation of 802.11, smartphone sensors, GSM beacons, infrared or ultrasonic systems, imaging techniques, and so on, have resulted in several open-source trajectory databases [11–13]. A trajectory is typically defined as a sequence of timestamped geographical locations of a moving object [12, 14]. A broad spectrum of transport applications can benefit from trajectory data mining, if the data corresponding to all traffic scenarios are readily available. Some of the applications of the trajectory data include, but not limited to, the scenario-based testing of autonomous vehicles [15, 16], developing and validating the traffic simulation models [17, 18], understanding the complex vehicle dynamics at various traffic facilities [19–21], and promoting safe and green transportation strategies [22–26]. Furthermore, the trajectory data is equally crucial for performing microscopic, mesoscopic, and macroscopic traffic

flow analyses [27].

Trajectory data opens up several research avenues but, achieving large-scale trajectory data is not a straightforward task [28]. The availability of an extensive trajectory database corresponding to all traffic scenarios is limited due to the difficulties associated with the trajectory data collection, extraction, and standardization. Accuracy and inhomogeneity vary among the trajectory databases, due to the differences in the traffic conditions, road layouts, and the spatial and temporal range of observations. Furthermore, the limited sharing of available data within the research community, restricted a comprehensive understanding of many aspects of traffic flow theory [29]. Some of the recent attempts of the trajectory data collection include, the Next Generation Simulation Models (NGSIM), T-Drive [30,31], Mixed Traffic Trajectory Data [32], MIT Trajectory Data Set [33], New York City Taxi Trip Data (2010-2013) [34], Mobile Century Data [35], GPS Trajectories Data Set [36], SHRP2 data [37], HighD data [38], and pNEUMA data [39]. Although the release of these trajectory datasets had led to a better analysis of various traffic dynamics, otherwise unknown previously, we are still facing a shortage of high-quality trajectory data [38].

In the recent past, the researchers are more concerned about acquiring Naturalistic Driving (ND) data. There have been several attempts to achieve ND data through various techniques; the two most prominent techniques are through vehicle instrumentation [37], and imaging techniques [39]. Some of the known limitations of data collection through vehicle instrumentation are;

1. This approach can attain only data corresponding to the instrumented vehicle
2. The driver is conscious that he/she is being monitored
3. Costly if need data at a higher penetration

The image-processing (IP) based technique can overcome most of these limitations, however, compromising the spatial range of the data. Nevertheless, being an extrusive data collection method, the imaging technique can capture naturalistic driving behavior without the driver's consciousness. Moreover, the data corresponding to all the interacting vehicles can be extracted, rather than only the instrumented vehicle. Although there are several benefits to using IP-based techniques for trajectory data collection, it is not straightforward. The following section briefly discusses the challenges of IP-based trajectory data collection and some of the existing practices.

2.2 Image-Based Trajectory Data Collection: History, Challenges, and the Current Practices

The IP based traffic data extraction dates back to the 1960s, and the studies during that time mostly relied on labor-intensive data extraction techniques. The IP based technique as a substitute to the presence detectors was first investigated by the Federal Highway Administration (FHWA) in the mid-1970s [40]. Later, Michalopoulos [41] has introduced a video-based traffic data collection

system, the *Autoscope*, which has been widely used thereafter. However, the function of the *Autoscope* was similar to that of a presence detector, not trajectory data extraction. Due to the higher cost and the laborious manual processing required in the office, IP based traffic data collection was not the primary mode of traffic data collection during that time. In recent years, with the advances in imaging technology, the availability of low-cost cameras, high-speed computers, and various image-processing techniques, the IP based traffic data collection has become more prominent among the other methods [42, 43]. The advantages of the IP-based traffic data collection include:

1. The ability to make site-specific adjustments
2. Portability
3. Visual reproduction of the traffic stream
4. Availability of unlimited data forms
5. Manual and automatic data processing capability

Researchers have utilized various IP techniques in a multitude of traffic applications such as the incident detection [44, 45], macroscopic traffic analysis [46–49], analysis of driver behavior at different scenarios [50–52], risk assessment [53–58] and so on. However, for several reasons, many of the above studies have reported that, transforming the video data into useful traffic data is a challenging process. Albeit, video-based traffic data collection is widely employed in the field of traffic engineering.

Achieving a better quantity and quality of ND data with sufficient information is undoubtedly demanding, and requires an appropriate data processing tool that can efficiently and accurately extract the data [59, 60]. Benefiting from the advances in computer vision and drone technologies, several attempts have recently been made to acquire ND data corresponding to various traffic scenarios. [38, 39, 49, 59, 61–63]. However, many of them have compromised the quantity/quality of the data, either to minimize the extent of human supervision required in the data extraction process or the cost. To avail better quality data, generally manual data extraction was followed [4, 32, 49, 64], which is, however, a backbreaking exercise [65, 66]. In the case of fully automated data extractors, some researchers have reported significant error in the extracted trajectory data [4]. A reliable trajectory extraction tool should provide an optimal trade-off between the data quantity, quality, extent of manual intervention, and the cost. It has already been reported that, achieving high-quality and detailed vehicle trajectory data with minimal cost and manual interventions is challenging even in the era of intelligent technologies and big-data, due to the lack of appropriate data collection technology and devices [27, 67]. Notwithstanding all the progress made in recent years, the trajectory data extraction problem still needs substantial improvement to achieve equitable access to better data for all researchers to participate in building efficient, safe, and green transportation infrastructures.

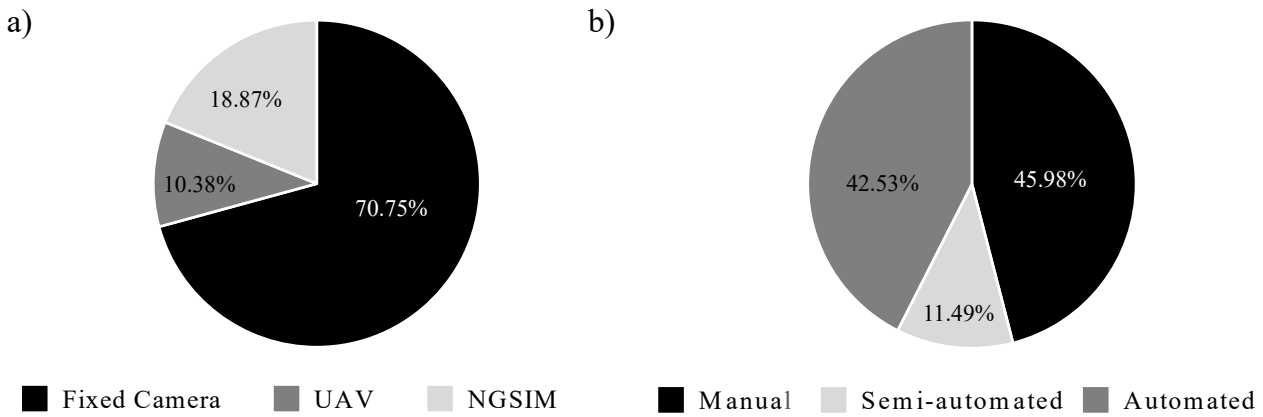


Figure 2.1: Review of trajectory based studies in the recent past; a) Data platform; b) Data extraction method

To understand the state-of-the-art of the IP based traffic flow analyses, the authors have reviewed one-hundred-and-fifty trajectory-based studies published in the top tier Transportation Engineering journals between 2017 and 2020. It can be seen from Fig. 2.1a that, 18.87% of the studies have used NGSIM data, even though Coifman et. al [4] have reported that the NGSIM data is highly erroneous, and the existing trajectory reconstruction techniques are inadequate to deal with the associated error. This indicates that the researchers depend on such data sources to minimize the complexity associated with the collection and extraction of a new set of traffic data or the limited availability of a publicly shared traffic database. In other words, there still lacks comprehensive, high-quality, and publicly available traffic data to perform the research in the field of traffic engineering. Further, 81.13% of the studies have used other than NGSIM data, out of which 70.75% have used a fixed camera setup, and 10.38% have used Unmanned Aerial Vehicles (UAVs) for the data collection. Though the studies that have utilized drones for traffic data collection are minimal, the recent trend indicates that drone-based traffic data collection will overrule the fixed camera-based traffic data collection [39, 68–74]. For transforming the videos into useful traffic data, studies have used various methods, which were broadly classified as manual, semi-automated, and automated. It is to be noted that, even in the era of technological innovations, nearly 45% of the researchers have employed manual data extraction, indicating the lack of accessibility to more advanced trajectory extraction tools.

Apparently, a cost-effective, reliable, and user-friendly tool that can incorporate the advances in computer-vision techniques is imperative for high-quality traffic data extraction. It is noteworthy that the absence of an appropriate tool to extract trajectories from a multi-class, disordered traffic stream was a major constraint for a detailed analysis of the multi-class traffic dynamics prevalent in most of the developing countries. Moreover, for the penetration of the autonomous driving systems in the market of developing countries, it is essential to get more and more data from corresponding traffic scenario. Similarly, the existing trajectory extraction tools are not suitable for extracting the trajectory data from complex road geometries such as horizontal curves superimposed with a vertical curve, as it violates the planar terrain assumption considered during coordinate transformation in many trajectory extraction tools.

2.2.1 Challenges of Video based Traffic Data Collection

The present-day computer-vision and drone technologies can extract high-quality vehicle trajectories from traffic videos with minimal supervision but are not easy to use for many researchers working in traffic engineering. The latest object tracking algorithms possess higher tracking accuracy and speed, and many are publicly shared. However, to utilize these tracking algorithms developed in different programming languages (mostly in python or MATLAB), the users need to make necessary modifications according to the study requirements. To make such modifications, the users should have a better programming language proficiency and an advanced knowledge of the computer vision algorithm flow. Nevertheless, the vehicle trajectory extraction involves multiple advanced processing stages; thus, only access to the tracking algorithm will not serve the purpose. The associated challenges in trajectory data extraction include the coordinate transformation, compensating the camera movements, initial detection, classification, defining tracking boundaries to terminate a completed track, detection and correction of the failed tracks, and compiling and exporting the data into a useful file format. As the computer-vision or advanced computer programming is not the area of expertise for a traffic engineer, it would be hard to make necessary modifications in the available code according to the requirements. Thus, only a limited studies could benefit from advances in object tracking techniques, and most researchers still depend on the manual data extraction frameworks [75]. Moreover, for a traffic engineer, the research interest may not be the extraction of the trajectory data but to infer the cause and effects of several traffic dynamics and incidents from the trajectory data through various analyses. Recently, there were several data extraction startups [76–81] that provide paid services for extracting the traffic data from videos, and is often used by researchers [39]. However, the cost of the data extraction/purchasing may not be affordable always for many of the researchers due to the lack of research funding or restricted access [82, 83]. Nevertheless, the data is indispensable; thus, there is a necessity for an accurate and user-friendly data extraction tool with minimal user-interventions and cost.

As discussed, numerous challenges arise at each level of the traffic data acquisition using imaging techniques. At the stage of video recording, the viewing angle of the camera affects the accuracy of the collected trajectory data when extracted using some of the current trajectory extraction tools. It is to be noted that, many of the current trajectory extraction tools demand a front facing camera-view, aligned with the center-line of the road [3, 84]. Such a consideration limits the traffic data collection at the locations where a road-related infrastructure such as the foot-over-bridge is available for mounting the camera [32, 49, 85, 86].

It was reported that the camera movement would add error to the trajectory data [13, 74]. From a controlled experiment using a small Unmanned Aerial Vehicles (UAVs or drones), Barmounakis et al. [68] have found that the accuracy of the speed data extracted from the drone video is highly dependent on the camera stability and the geo-referencing procedure. Therefore, it is essential to perform necessary corrections to the extracted trajectories that nullify the effect of camera movements. Since the vehicle trajectories were assumed to be on a Euclidean plane, minor devi-

ations such as scaling, rotation, or translation (in 2 dimensions) could lead to a significant error in the extracted trajectory data. Yet, only a few studies have considered the effect of the camera movement on the extracted data accuracy. Kaufmann et.al [74] have collected vehicle trajectories using a small-scale quadcopter camera, and they have reported a prolonged but permanent movement of the camera, which impacts the data accuracy. To correct this error, they have evaluated the camera position for each frame and applied a necessary adjustment to the tracked vehicle position based on the shifted camera position. Kaufmann et.al [74] have reported that the measurements in two neighboring video frames (with a frame rate of $25fps$) have caused a significant error in the speed that was about $1pixel/frame = 1m/0.04s = 90km/h$. Similarly, several researchers have discussed the importance of video stabilization during drone-based traffic data collection [68–70, 87, 88].

The detection and classification of a wide variety of vehicles in the traffic video is another challenge associated with trajectory data extraction. Many studies have reported that this stage is more prone to error compared to tracking. As the traffic composition and vehicle types differ from region to region, it is not easy to train a detection or classification module considering all the vehicle types. Moreover, to achieve better accuracy in vehicle detection and classification, the image should contain sufficient vehicle features. For automated detection and classification, it is essential to preserve the image quality to keep sufficient features of the object of interest. Studies that employed the automatic detection and tracking of the vehicles have limited the height of the camera/drone to maintain sufficient foreground information [38, 59]. It is apparent that, as the height of the flight goes up, the foreground information in the traffic video would depreciate. In such cases, the automatic processing fails more frequently with illumination changes, background clutter, partial occlusions, high-speed moving target, moving camera, target deformation, and so on [24, 59, 66]. On the other hand, when the height of the camera/drone is restricted, its spatial coverage would decrease, resulting in shorter trajectory data. It is to be noted that the issues that arise from an inferior quality video can be resolved with a little human supervision, rather than compromising the quantity of the data. Moreover, the automated classification of vehicles limits the application of such tools to the traffic stream that contains only the trained vehicle classes, and user does not have the freedom to add vehicle classes of his/her study interest.

The transformation of the vehicle coordinates in pixels into real world coordinates is another challenge related to IP based traffic data collection. Researchers have proposed various camera calibration techniques to transform the image coordinates into real world coordinates [89–93]. Fung [89] proposed a two-vanishing-point technique, considering road lane markings as the reference lines for calibration. The major drawback of this approach is the ill-conditioning effect that occurs due to near-parallel lines. He et al. [90] proposed a calibration method using a single vanishing point from the lane markings to overcome the errors caused by ill-conditioning. This method does not assume any extrinsic parameters to be known a priori. Direct linear transform (DLT) is another technique, which uses point to point correspondence for estimating the homography matrix [91]. However, the prime assumption of all the techniques mentioned above is the planarity of the real-world region. Since the existing tools are limited to the planar region, most of the image-processing

based trajectory extraction studies are restricted to the midblock sections falling in plain terrain or straight road with a constant grade [32, 63, 84, 94, 95].

While developing a trajectory data extraction tool, it is essential to consider the above-listed facts to achieve better quality and quantity of trajectory data. However, no method can eradicate all the potential errors that can occur during trajectory data extraction. Therefore, the extracted raw trajectory data should be post-processed to eliminate the noise. Section 2.3 addresses various noise reduction techniques employed in the existing studies. Some of the commonly used trajectory data extraction tools are presented in the following section.

2.2.2 Existing Tools for Image-Based Trajectory Data Extraction

The researchers have employed several trajectory extraction tools in the literature, and each one has its own advantages and disadvantages. Table 2.1 lists some of the tools used to extract the vehicle trajectory data from traffic videos, and their properties.

Table 2.1: Review of the image-processing based trajectory extraction tools

Author & Year	Tool	Manual	Remarks
Wei et al., (2005) [95]	Vehicle Video-Capture Data Collector (VEVID)	✓	User manually picks the vehicle position at each frame. VEVID outputs the trajectory data along with speed, acceleration, and the gap between the targeted vehicles.
Mallikarjuna et al., (2009) [84]	TRAZER	×	Automatically detect, track, and classify the vehicles; TRAZER outputs traffic characteristics such as the classified vehicle flows, average vehicle speeds, average occupancies, vehicle trajectories, lateral, and longitudinal spacing.
Munigety et al., (2014) [3]	Traffic Data Extractor (TDE)	✓	Data can be extracted with fidelity of 0.1 sec with a manual vehicle tracking at each frame; The real-world coordinates from the two-dimensional image coordinates were obtained after camera calibration using vanishing point-based technique.
Apeltauer (2015) [96]	Vehicles' Trajectory Extraction from Aerial Video Data Captured Using UAV	×	Detection is based on Boosting technique. To increase the robustness to changes in illumination and to accelerate the training phase, Multi-scale Block Local Binary Patterns (MB-LBP) was employed. A sequential particle filter has been adopted. A gradually updated rectangular template represented the target.
Lee et al., (2008) [94]	Trajectory Extractor	✓	The video images are displayed frame by frame according to the given processing interval; tracking is by mouse-click on the vehicle at each frame. The screen coordinates of the mouse-clicks are then converted to the real-world coordinates.

Babinec et al., (2015) [97]	Automatic Vehicle Trajectory Extraction from Aerial Video	×	Vehicles were detected using the AdaBoost classifier constructed of Multi-Scale Block Local Binary Patterns features. The vehicle tracking is done by multi-target tracker based on a set of intra-independent Bayesian bootstrap particle filters specialized in dealing with environmental occlusion, multi-target overlap, low resolution and feature salinity of targets and the changes in their appearance
--------------------------------	---	---	--

2.3 Trajectory Data Reconstruction

Raw trajectory data are complex and often contain noise and errors [5, 13, 27, 29, 64, 98–100]. Prior to using the trajectory data for any analysis, it is essential to filter out abnormal/erroneous data and standardize it. If not properly accounted for, the resulting speeds and accelerations would be inconsistent and may have significant implications on practical applications [29]. A trajectory reconstruction method should retrieve all the kinematic features of the vehicle motion conforming to the actual vehicle dynamics. Therefore, before developing any trajectory reconstruction technique, it is crucial to know the properties of a meaningful trajectory.

2.3.1 Properties of a Meaningful Vehicle Trajectory

Vehicle trajectories are the most detailed empirical data available to investigate the traffic dynamics [29]. The accuracy of the trajectory data is ensured through various consistency analyses such as internal consistency, platoon consistency, and physical consistency. Internal consistency deals with the kinematics of a single-vehicle, while platoon consistency verifies the coherence of inter-vehicle spacing of vehicles in car-following. The physical consistency ensures the realism of the macroscopic traffic dynamics, considering the distribution of the speed and acceleration of the vehicles. These properties are derived from various physical and dynamic constraints of the vehicle motion.

At any time instance, a moving vehicle will have a well-defined speed and acceleration, which are the first and second derivatives of the position with respect to time, respectively. That is,

$$\left. \frac{dx}{dt} \right|_{(t_i, x_i)} = v_i \quad \quad \quad \left. \frac{d^2x}{dt^2} \right|_{(t_i, x_i)} = a_i$$

Where x_i is the position of the vehicle at time t_i ; v_i is the instantaneous speed; a_i is the instantaneous acceleration. This property makes the trajectory data to be continuous as well as differentiable. Besides this, vehicle trajectories should give meaningful speed and acceleration values that satisfy the equations of motion [101].

$$x(t_i) = x(t_0) + \int_{t_0}^{t_i} v(t) dt \quad (2.1)$$

$$v(t_i) = v(t_0) + \int_{t_0}^{t_i} a(t)dt \quad (2.2)$$

Equations 2.1 and 2.2 can be used to check the *internal consistency* of a trajectory [102], and any reconstructed trajectory should satisfy these conditions. Treiber et. al [103] have stated that the internal consistency conditions are the primary kinematic constraints of the vehicle trajectories. The internal consistency ensures the agreement between the position, speed, and the acceleration values [32]. Montanino et al. [5] have considered the internal consistency as a constraint in the non-linear optimization problem, for the local reconstruction of the vehicle trajectory. Punzo et al. [29] have found that the internal consistency considerations alone may not ensure the exactness of inter-vehicle spacing in case of falsely tracked vehicles.

Platoon consistency is an important aspect to be considered if the data acquisition method leads to false detection of the vehicles. False detection is one of the major issues associated with the NG-SIM data, and this leads to unrealistic acceleration or deceleration of the vehicles [5]. The concept of platoon consistency was introduced by Punzo et al. [101] to ensure non-overlapping vehicle trajectories. Punzo et al. [104] claimed that this analysis is necessary since the trajectory data are intended not only for studying the individual vehicle dynamics, but also for understanding microscopic traffic dynamics. Trajectory data collected on several vehicles should satisfy the platoon consistency which deals with the integrity of the inter-vehicle spacing [29]. This consistency evaluation is derived from the car-following behavior and ensures that the trajectory data do not result in unrealistic headways. Kanagaraj et al. [32] have also mentioned that, ensuring the platoon consistency avoids the overlapping of the estimated trajectories. In case of the semi-automated trackers, the possible noise in the trajectory data could be the White Gaussian. The possibility of false tracking rarely occurs with semi-automated trackers because of the direct human inputs. Hence, the issue of platoon inconsistency may not arise. Even in case of the false detection, the equations of platoon consistency, as proposed by Punzo et al. [104], are not applicable for the heterogeneous traffic streams due to the multiple leader-follower combinations and parallel movements. Within the existing setup of the platoon consistency, it is not possible to incorporate such microscopic constraints to the trajectories obtained from heterogeneous, no lane-disciplined traffic stream.

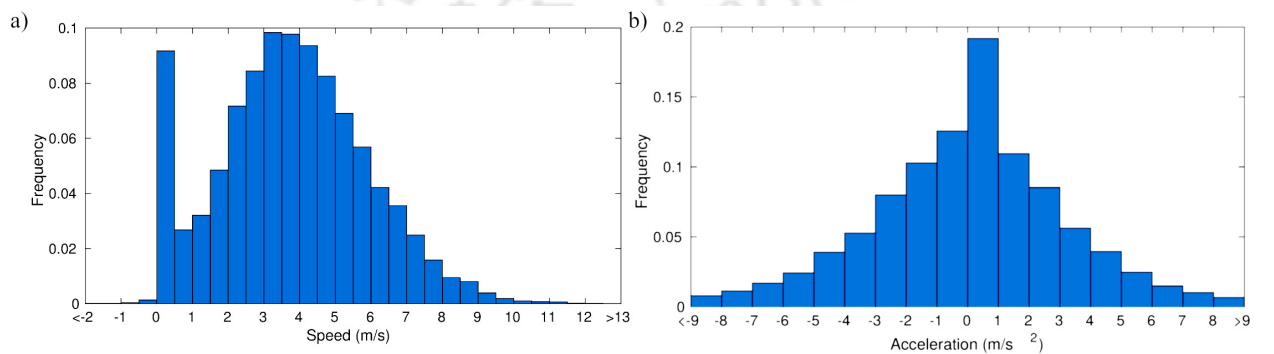


Figure 2.2: Speed and acceleration distributions estimated from the observed data (extracted using Traffic Data Extractor (TDE) software [3]); a) Frequency distribution of the speeds estimated from the observed trajectory data; b) Frequency distribution of the accelerations estimated from the observed trajectory data

Physical consistency can be applied to the entire set of trajectories, and it verifies whether the speed and acceleration distributions are within the realistic range. Figure 2.2a shows the frequency distribution of the speed data of about 155 vehicles observed over a 100-meter road stretch during a 120 second period. It can be seen from the figure that there are negative speeds. Also, it is evident from the Figure 2.2b that, there are several unrealistic acceleration values exceeding $3m/s^2$ and the deceleration exceeding $5m/s^2$ [5].

Another property that can be conveniently assumed is that the vehicle trajectories are a combination of piece-wise lower order polynomials. The real world driving patterns result from the instantaneous decisions taken over the physical driving environment [105] and also depends on the safety, convenience, and comfort aspects. According to Hamdar et al. [106], driving is not anymore a discrete process, but a story divided into multiple episodes. This continuous story is characterized by the probability of terminating a given episode based on the driving experience encountered. Various driving patterns of a driver can be considered as, stopped condition, constant speed motion, acceleration, and deceleration. Acceleration and deceleration processes result in the non-linear portions of the trajectory, whereas the other two actions result in the linear portions. The degree of non-linearity depends on the rate of acceleration/deceleration. The rate of change of acceleration is termed as the *jerk*, which is proportional to the rate of change of force [107]. It can be an apparent behavior of any driver that he/she always tries to minimize the rate of change of force to keep them on a safe and comfortable side while executing the acceleration, deceleration, and lane change maneuvers. Hence, these non-linear portions can be conveniently represented with the lower degree polynomials.

The above-discussed characteristics of the vehicle trajectories need to be considered in the reconstruction framework. Trajectory data are collected in different ways, and the corresponding errors are entirely different. As mentioned earlier, the reconstructed trajectories must exhibit internal, platoon, and the physical consistency. Some of the approaches consider these properties as constraints in the reconstruction methodology [5]. Many researchers have not included these properties in the reconstruction methodology but used for the validation of the reconstruction methodology [108, 109]. Any primary approach used for smoothing need to consider the internal consistency implicitly in the smoothing approach. The representation of trajectories with continuous, piece-wise lower order polynomial functions will always satisfy the internal consistency conditions.

2.3.2 Various Trajectory Reconstruction Techniques Generally Practiced in Traffic Engineering

With the advent of NGSIM dataset [63] where a vast amount of trajectory data made publicly available, a series of trajectory reconstruction techniques were developed [5, 29, 99, 100, 106, 108–111]. Except a few, many of the trajectory reconstruction techniques reported in the transportation literature were mainly meant for and customized to NGSIM data. Without checking the adequacy, some of the studies have employed off-the-shelf smoothing techniques with an arbitrarily chosen

smoothing parameters, and invariantly applied across different vehicle trajectories. It is apparent that the nature of noise for each trajectory would be different due to several reasons as discussed in the later part of this chapter. In the case of NGSIM trajectories, a major emphasis was on the error resulted due to the tracker malfunctioning, owing to the low quality of the NGSIM videos (resolution 640×480 pixels), and the extreme angular view of the camera [5]. Nevertheless, in the modern era of traffic data collection, the researchers could collect high-quality traffic videos (e.g., at 4k (4096×2160) resolution) at a bird-eye view, thanks to the drone technology. By virtue of the advances in computer-vision techniques, the process of trajectory extraction was also revolutionized, thereby achieved an accuracy level comparable to the manual trajectory extraction. Unlike the error pattern observed with the fixed camera setup, a new sort of error pattern is found in the trajectory data extracted from drone videos, which is unknown and cannot be sufficiently attenuated using the existing filtering techniques.

The contemporary traffic data collection using the drone videography is more complicated compared to the stationary or fixed camera system [112]. Moreover, the information on the accuracy and reliability of the extracted data from drone video is very minimal [68]. In order to handle various errors generated while collecting the trajectory data, researchers have employed several filtering/smoothing techniques. Table 2.2 lists some of the recent studies based on trajectory data and the corresponding filtering strategy. It is apparent from the table that the trajectory smoothing problem is highly disparate and offhand in terms of its implementation. Though it is well-known that the smoothing results would significantly affect the study output, many studies have not given significant importance to the efficacy of the smoothing techniques. It is to be noted that the selection of smoothing parameters has an evidential impact on the accuracy of the reconstructed path and the reliability of all the derived quantities including, speed, acceleration, Time-to-Collision (TTC), headway, and gap. However, the smoothing parameters considered in most of the studies were arbitrarily chosen and were invariant across different vehicles' trajectories. Many of the recent studies have employed the moving average method, which is known for the lag produced in the output data when a larger filter window is used [113]. A few studies have applied a correction to the imposed lag through correlation analysis; however, majority studies have not addressed this issue. It is apparent that, such a lag can significantly affect the result of the time-series analyses or derivation of time-dependent attributes, such as TTC. Moreover, the use of inappropriate smoothing parameters would lead to over/under-smoothing of the data, which may cause a misrepresentation of the actual vehicle kinematics. The over smoothing may subside the effect of the harsh events on the kinematic quantities, such as the acceleration and jerk, and would lead to the under-representation of potential crash events. The under-representation of the possible crash events is, however, a problem for vulnerable road users [114, 115]. On the other hand, the under-smoothing of the data would lead to jerky vehicle actions that may underrate the quality of the service. Thus, an appropriate estimation of smoothing parameters to achieve adequate smoothing is crucial in the case of traffic studies.

Table 2.2: Review of trajectory smoothing techniques employed in the transportation literature

Author	Data	Smoothing/ Filtering Technique	Remarks
Barmounakis et al. [39]	Data collected using a swarm of UAVs (10 UAVs) from an urban road network covering 1.3km.	<i>Kalman filter</i>	Mentioned that the filter was able to filter out the noise in the measurement upto a level of 3.3 cm. No information on smoothing parameters is provided.
Li et al. [116]	NGSIM I-80 and I-101	<i>Moving average filter</i>	No information on the filter width. A negative vehicle movement was reported due to the moving average filtering.
Durrani et al. [117]	NGSIM (US-101)	<i>Symmetric exponential moving average filter</i>	Chosen a smoothing window of 0.5 second for position, 1 second for speed, and 4 seconds for acceleration.
Gu et al. [71]	Data was collected using UAV at a merging area of Maqun interchange in Nanjing, China.	<i>Moving average filter</i>	A window size of 2 seconds, and a sampling rate of 4Hz was considered for filtering. All trajectories were filtered with the same filter configuration. A lag is clearly visible with the smoothed data shown in Fig. 5 of Gu et al. [71], but no measure has been taken.
Mohammed et al. [118]	Trajectories are extracted from the video data collected at Brooklyn bridge, USA, for two hours.	<i>S-G filter</i>	Used S-G filter with seven time-step window (7/15 sec) for all trajectories. The method is validated using the average speed measured using the known distance and time-spend.
Tian et al. [119]	25 car platoon experiment data	<i>Moving average filter</i>	Used a constant window size of 1 second, for all trajectories. Time delay is corrected using correlation analysis.
Wang et al. [56]	Data from ten urban signalized intersections collected using a drone. A total of 4hr data.	<i>Cubic spline</i>	No further information on the implementation of smoothing technique.
Wei et al. [120]	Video based trajectory data collected at a 4-legged intersection	<i>Kalman smoothing</i>	No information on smoothing parameters. The boundary effect is visible in the presented acceleration and jerk profile (Ref. Fig. 9 in Wei et al. [120])

Wu et al. [22]	Data from a controlled experiment with 19-22 cars	<i>Spline smoothing</i>	Each trajectories were reconstructed using a fourth-order B-spline. The choice of fourth-order spline was to achieve smoothness in acceleration. The weights are chosen such that the acceleration estimates obtained from the second derivative of B-spline are within the physically reasonable bound. Additionally, a parameter called tolerance was used to control the length of every single fitting window.
Xing et al. [57]	Trajectory data collected from a toll plaza using UAV. The total duration of data was 1.5hr, and the length of the road stretch covered was 400m.	<i>Moving average filter</i>	Reported that the UAV flying in the air would have jitter due to the impact of wind, which causes an error in the extracted trajectory. No information on the smoothing parameters used.
Krajewski et al. [38]	HighD Dataset	<i>Rauch-Tung-Striebel (RTS) Smoothing</i>	RTS was basically developed for state estimation of linear dynamic systems in the presence of additive Gaussian noise.
van Beinum et al. [61,72,73]	Used empirical trajectories from 14 sites in the Netherlands. Trajectories are collected using a camera mounted on a helicopter. The length of the road covered was 1200-1500m	<i>Polynomial regression filter</i>	No information on the polynomial order and window size used for smoothing.
Fard et al. [109]	NGSIM (I-80)	A two-step approach based on wavelet analysis	Step 1: outliers are identified locally and modified through applying the Gaussian kernel regression. Step 2: raw trajectory data is decomposed into three levels using DWT and Haar wavelet as the basis function. The denoising process is then conducted by defining a threshold for the detail coefficients. At last, the vehicle trajectory data is reconstructed by employing an inverse DWT. The smoothing technique is difficult to implement or not available for direct use.
Saifuzzaman et al. [121]	NGSIM (US-101)	<i>Weighted moving average filter</i>	Used a constant window size of 1 second, for all trajectories. An artificial time lag was introduced, thus used a two-sided weight function.

Zhou et al. [122]	NGSIM (I-80)	<i>Moving average filter</i>	Used a constant window size of 0.8 second, for all trajectories. Negative vehicle movements due to MA filtering were set to zero. Further refined the acceleration to a range of (-3.41,3.41).
Montanino et al. [5]	NGSIM	A four-stage reconstruction process	Step 1: outlier removal using a natural <i>cubic spline</i> with ten reference points before and after the sequence of outliers; Step 2: noise removal using a <i>Butterworth filter</i> of order one and cut-off frequency of 1 Hz; Step 3: Local weighted reconstruction using a fifth-degree <i>polynomial function</i> ; Step 4: residual noise removal by applying same lowpass filter in step two. This method is customized to NGSIM trajectories and hard to implement.
Zaki et al. [123]	Trajectory data were collected from two locations in North America and the Middle East using a fixed camera setup near junctions.	<i>Savitzky-Golay (S-G) filter</i>	Utilized a second-order S-G filter for all the trajectories. No information available on the filter size used.
Ahn et al. [124]	NGSIM (I-80, I-101)	<i>S-G filter</i>	The filter window was 1 second for all the trajectories.
Bagdadi et al. [125]	200 passenger car data	<i>Weighted exponential smoothing</i>	Weighted averaging with a 3 point moving window, and a weighting constant of 0.4 for all the vehicles.
Thiemann et al. [99]	NGSIM	<i>Symmetric exponential moving average filter</i>	Chosen a smoothing window of 0.5 seconds for the position, 1 second for speed, and 4 seconds for acceleration based on the visual comparison of the smoothing results with different window sizes

Regardless of the availability of a customized smoothing technique for NGSIM data [5, 100, 109], researchers are still using the basic, off-the-shelf filtering techniques, without further verifying the smoothing accuracy. Such a practice could be due to the difficulty in implementing the multi-stage algorithms developed for NGSIM data compared to the off-the-shelf smoothing techniques. However, for enhancing the quality of transportation researches using vehicle trajectory data, it is desirable to develop an appropriate smoothing technique that can be easy to use for the researchers.

A preliminary analysis of the data collected using the fixed camera system and UAV clearly distinguish the difference in the error pattern (see Fig. 2.3). The extracted trajectories from aerial videos have shown a large and unrealistic variance with heavy tails for acceleration distribution

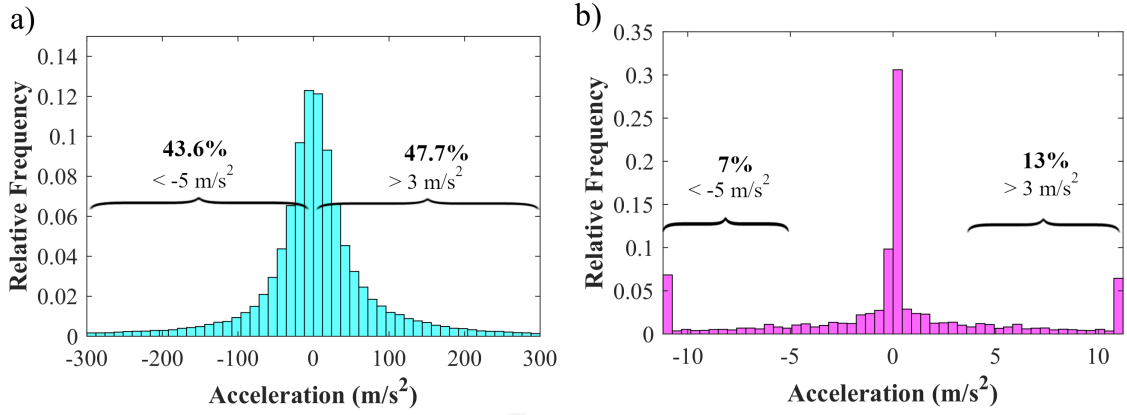


Figure 2.3: Observed acceleration distribution from drone and fixed camera setup; a) Acceleration distribution for drone based traffic data collection; b) Acceleration distribution for fixed camera setup (NGSIM data)

(Fig. 2.3a), compared to the fixed camera system (Fig. 2.3b). Thus, it is apparent that the observed trajectory data from drone videos would contain the error that has a large variance, and follows a heavy-tailed distribution. "Heavy tails" means that for large values of error ϵ , the density, $f(\epsilon)$, approaches zero more slowly than the Gaussian [126]. One of the possible causes of heavy tailing in the error distribution could be the camera shakiness. Generally, the camera shakiness is induced either from the mechanical vibrations, or the drift of drone due to the prevalent weather conditions and an associated repositioning of drone to the original position using the embedded positioning technology (e.g., Vision or GPS positioning), which is available in most of the high-end drones (e.g., DJI-Inspire, DJI-Phantom). It is to be noted that the listed models of the drone are commonly used for traffic data collection [38, 39, 71], thus, a heavy tailed error pattern could be expected in the extracted trajectory data. Though the researchers have employed several video stabilization techniques, all of them can only minimize the camera movements or smooth out the camera transitions. No stabilization technique can altogether remove the camera movements; thus, certain camera movements, mostly Gaussian, are expected to remain. Yet, a sub-pixel level movement of the camera can add a significant error to the extracted vehicle path in the real-world coordinate system, which might also vary with the height of the flight, and video resolution. Apart from the camera shakiness, the jump of the bounding box around the subject vehicle during the tracking process can also lead to heavy tailing in the error distribution. It is to be noted that, during the time of camera jitter or jump of the bounding box, the noise in the extracted trajectory data would be non-Gaussian and has much heavier tails than the Gaussian. The treatment of heavily tailed error is a challenging task since an inappropriate reconstruction of such data would lead to local artifacts, thereby unrealistic estimation of kinematic quantities. Thus, it is indispensable to have a novel trajectory reconstruction technique to address the heavily tailed error pattern, associated with the current trajectory data collection methods.

2.3.3 Importance of a Separate Smoothing for Lateral and Longitudinal Components of a Vehicle's Path

In the literature, the vehicle dynamics are represented using one-dimensional space trajectories due to the prominence of one-dimensional traffic flow theories [127–130], and the lateral dynamics were studied and modeled separately [131–136]. However, the state-of-the-art studies are mainly focusing on the two-dimensional vehicle dynamics due to the prominence of autonomous driving technology, and also for a more realistic explanation of traffic dynamics [137–143]. In fact, understanding the lateral dynamics such as lane changing, overtaking, in association with the longitudinal dynamics, or the vehicle dynamics at intersection, roundabouts, or curves, the two-dimensional trajectories are inevitable. To clearly distinguish the dimensional difference, hereafter, we would use the time-series of two-dimensional position as the vehicle path. The existing reconstruction techniques mainly discuss the trajectory reconstruction [5, 29, 64, 100, 102, 109]; and, the path smoothing is underrated. The challenges in smoothing the path data include the added non-linearity due to steering, apart from the non-linearity due to acceleration/deceleration and the dissimilar error pattern in the longitudinal and lateral direction of the path.

As presented in Table 2.2, most of the existing trajectory smoothing techniques have considered invariant smoothing parameters for all the vehicle trajectories with an assumption that the noise is similar in every case. We conjecture that the level of noise in the extracted trajectories would vary from vehicle to vehicle and can have local artifacts in the form of tiny spikes with a wide range of amplitudes. Such a noise can be due to the nonuniform jittering of drone video, and a variable tracking accuracy across the vehicles based on the lighting condition, the number of features available, background clutter, etc. Furthermore, the existing trajectory smoothing techniques have either neglected the reconstruction of the lateral position of a vehicle or applied the same level of smoothing for both the longitudinal and lateral positions. By saying ‘the same level of smoothing’ we mean that the smoothing parameters are invariant for longitudinal and lateral directions. But, it is very likely to have a dissimilar error pattern for the longitudinal and lateral positions irrespective of the camera platform. The difference in the pixel dimension along and across the road, perspective distortion, the difference between optical flow and the pixel directions, the difference in the magnitude of camera vibration in each direction, vehicle localization error while tracking, are a few reasons to name. From the foregoing discussion, it is evident that the nature of noise is different for different vehicle paths and also different for their lateral and longitudinal components. To capture such an extensive heterogeneity of error in the observed path data, an adaptive smoothing needs to be performed for each vehicle path and each direction according to the level of noise.

2.4 Trajectory-Based Proactive Safety Analysis

It is estimated that more than 90% of road crashes are due to human error [144]. As such, the observation and generalization of hazardous events is a real challenge. In the past, the steps taken to improve road safety were mostly dependent on the crash data analyses and modeling. However,

such steps were not fully successful in improving road safety due to several issues related to the crash data. The primary issues concerning the crash data are [145–148]: (i) traffic crashes are quite exceptional compared to other traffic incidents as they are the product of a sequence of unpleasant realizations of many small probabilities; (ii) a long period of data collection time is often needed to obtain adequate crash data for statistical modeling; (iii) not all traffic crashes are recorded, and the degree of under-reporting depends on the crash severity and the categories of road users involved; (iv) information on the behavioral aspects preceding the crash is rarely accessible. Besides, the available traffic crash reports typically describe injury levels as discrete variables [149]. Also, the biases caused by missing data in crash database with incomplete records frequently hampered the estimation of crash contributors' effect on crash severity and the prediction of crash severity outcomes [150].

The proactive safety assessment has been reported as a powerful tool to assess the crash risk probabilities on various transport infrastructures [24, 54]. An extensive set of accurate trajectory data would allow to proactively assess the safety at various transportation facilities, thereby implementing appropriate mitigation strategies. In the last 40 or more years, a large body of literature has accumulated to identify and develop measures relating to 'crash proximity,' typically called crash surrogates. The idea is to identify the circumstances that may lead to a traffic crash and assess the likelihood and severity. As the road crashes are rare events, the primary role of safety surrogates is to indicate the opportunities for enhancing road safety without focusing on individual crash occurrences. It also helps to recognize places that have not undergone historical crashes but can predict potential instances of future crashes [151]. Thus, several researchers have postulated that road safety analyses would surely benefit from reliable measures that use detectable non-crash events as a surrogate to crashes [152, 153]. The following section presents a detailed review of the existing SSMs.

2.4.1 Review of Existing Surrogate Safety Indicators

Owing to the well-known drawbacks of the reported crash data, researchers have realized the need to use some surrogate safety measures (SSMs) that can capture the crash risk. The SSM would allow one to:

1. Assess road traffic safety more effectively and in a shorter time;
2. Elaborate the link between design elements and risk;
3. Fully understand the relationships between driving behavior and the associated risk; and,
4. Gain a better understanding of the mechanisms that characterize regular traffic and critical events such as crashes.

Past studies have introduced numerous surrogate safety measures. In search of existing surrogate safety indicators, the present study have identified 17 indicators that are widely discussed.

Broadly, these surrogate safety measures can be classified into three categories: temporal, spatial, and acceleration-based SSMs. Table 2.3, 2.4, & 2.5 present some of these indicators. The tables also present the advantages, disadvantages, and the suitability of SSMs in evaluating different crash types. It is evident from the table that there is no single measure that can capture all crash risks, irrespective of the road geometry, crash type, and traffic condition.

The most common types of crashes on the freeway are the rear-end and sideswipe collisions [154]. When compared to the number of studies on rear-end collisions, research on sideswipe crashes is rather limited [155–157]. While several safety indicators are available to quantify the crash risk, most of them primarily focus only on the rear-end collision risk. Furthermore, the applicability of these safety indicators towards addressing overall crash risk during driving, or at a road facility is questionable [146, 153, 158]. Presently, no safety indicator is available to capture the crash risk corresponding to all possible crash occurrences. It is a major gap of the surrogate safety analysis reported by several studies [115, 159, 160].

Traffic dynamics at crossings, roundabouts, curves, and vehicle operations such as lane-changing, overtaking, and merging are unique and involve interaction between multiple vehicles on a 2-dimensional (2D) surface. Under these conditions, a broad spectrum of collision patterns, such as head-on, side-swipe, rear-end, and angled collisions, are possible. An SSM that could comprehend these wide-ranging collision patterns, and include them under a common severity hierarchy, needs to be developed for a more reliable proactive safety analysis. Since many of the existing SSMs are 1-dimensional (1D), as per definition, and assumes the kinematic quantity of vehicle motion to be a scalar, they cannot be used explicitly in all the traffic scenarios [161]. In reality, the vehicle states are vector quantities, and their usage in a scalar proximity calculation is non-trivial [26, 161].

It is to be noted that, among the existing SSMs, the temporal proximity indicators are the most popular measures since they integrate spatial proximity and speed. In the generation of temporal proximity measures, the Time-to-Collision (TTC) is a widely used surrogate safety indicator with several extensions [160, 162, 163]. More generally, TTC can be defined as the time left for a potential crash if two vehicles continue their states [164]. The conventional TTC measure assumes a car-following situation where the collision course is conformed with the lateral overlap of the vehicles. Notwithstanding the potential benefits of TTC in practice, this measure has several limitations, and some of them are [160, 162, 165]: (i) highly customized to rear-end conflict type, therefore, other conflict sequences are not captured; (ii) the conflicting vehicle's movement is restricted to either right-angled or parallel to each other; (iii) since collision course detection is conditional, TTC estimation limited only to specified cases; (iv) it is difficult to achieve a common severity hierarchy.

Table 2.3: List of surrogate safety indicators that measure temporal-proximity of vehicles

Indicator	Definition	Limitations	Advantages	Suitability for collision type
Time-to-Collision (TTC) or Time-Measured-to-Collision (TMTC)	TTC is the time remaining for two vehicles to collide if they continued on their course at the same rates.	TTC ignores several potential conflicts due to acceleration or deceleration discrepancies; TTC can provide the magnitude of crashes but not their severity;	TTC is used in practice than other SSMS; TTC is more informative than PET; Many automobile collision avoidance systems or driver assistance systems use TTC.	Rear-end, turning/weaving, hit objects/parked vehicle, crossing and hit pedestrian.
Time Exposed Time-to-Collision (TET)	Summation of all moments that a driver approaches a front vehicle with a TTC-value below the threshold (TTC*).	It does not provide information on the severity of a crash.	TET can be estimated for different vehicle classes separately. It gives an idea, on average, how long a vehicle is exposed to crash risk.	Same as TTC.
Time Integrated Time-to-Collision (TIT)	Integral of the TTC-profile during the time it is below the threshold (TTC*).	It is difficult to interpret; Estimation is not straightforward;	Severity levels of a vehicle collision can be understood; Allow to perform safety of transport facilities in terms of probable collision severity.	Same as TTC.
Modified Time-to-Collision (MTTC)	A modified TTC version that considers all the potential longitudinal conflict scenarios due to the acceleration or deceleration discrepancies.	Not fit for angled or side-swipe collisions; Does not reflect the severity of a collision.	More general than TTC; Consider driving discrepancies; Severity of the collision could be weighted using CI indicators.	same as TTC

Table 2.3 continued from previous page

Indicator	Definition	Limitations	Advantages	Suitability for collision type
Crash Index (CI)	Influence of the speed on kinetic energy involved in collisions.	CI describes only the safety information about two vehicles at a particular time and place. Not fit for side-swipe or angled collision.	CI quantifies the severity of a potential crash; Severity and the likelihood of a potential conflict could be interpreted.	Same as MTTC.
Headway (H)	Time elapsed between the successive arrival of two vehicles	Mainly applicable to read-end collision in lane-based traffic environment; Not applicable for other crash types.	Easy to measure; Level of safety could be distinguished.	Rear-end collision
Time-to-Accident (TA)	Time-to-Accident (TA) is the time that remains to an accident from the moment one vehicle starts an evasive action if they continue their speed and directions.	Often criticized for relying heavily on the subjective judgment of speed and distance. Mainly rely on vehicle's evasive action.	Widely used; Easy to measure; Can be done by both manually or by Video analysis, Couple of manuals have been developed in different countries.	Same as TTC.
Post-Encroachment Time (PET)	The time between the moment that a vehicle leaves the location of a potential collision and the other road user arrives at that location.	Only useful in the case of transversal (i.e., crossing) trajectories (Right angle collision); Cannot reflect changes with the dynamics of safety-critical events over a larger area; Levels of severity as well as the impact of a conflict are not taken into account.	PET is more appropriate than TTC for intersecting conflicts; PET can be easily extracted; Can be easily estimated using photo-metric analysis in video or simulated environment; PET represents the driver behaviors.	Mainly for right angle or crossing crash. Merging/diverging, head on (to a certain extent).

Table 2.4: List of surrogate safety indicators that measure spatial-proximity of vehicles

Indicator	Definition	Limitations	Advantages	Suitability for collision type
Potential Index for Collision with Urgent Deceleration (PICUD)	Distance between the two vehicles when they completely stopped state.	Mainly applicable in lane changing condition when leading vehicle apply emergency break; Threshold value needs to be set manually; Do not take into account lateral conflicts.	PICUD is more suitable to evaluate the collision risk of consecutive vehicles of similar speeds. PICUD can detect the change in traffic condition and conflicts more sensitively than TTC.	Same as TTC.
Proportion of Stopping Distance (PSD)	Ratio between the distance remaining to a potential collision point and the minimum acceptable stopping distance.	Estimated based on evasive actions; PSD provide higher percentage of vehicles interaction and time exposure to conflict than TTC and DRAC.	Single vehicle conflict with fixed or unfixed objects can be evaluated; Easy for observation and calculation.	Hit object (on road or road side), overturning.
Margin to Collision (MTC)	Ratio of summation of the inter-vehicular distance and the stopping distance of the preceding vehicle and the stopping distance of the following vehicle.	Same as Stopping distance. In addition, it does not consider a response delay of the following vehicle. A non-dimensional parameter.	It provides the possibility of conflict when the preceding and following vehicles abruptly decelerate at the same time.	Same as Stopping distance.
Unsafe Density (UD)	Level of “unsafe” in the relation between two consecutive vehicles on the road for a determined simulation step.	UD could be used only for comparison purposes; fit for only rear-end collision analysis; Applicable only for a lane based traffic situation.	Gives more accurate information than typical micro-simulation outputs; Comparative study between road network links can be done.	Rear-end, merging and diverging or lane changing.

Table 2.4 continued from previous page

Indicator	Definition	Limitations	Advantages	Suitability for collision type
Difference of Space distance and Stopping distance (DSS)	Difference of the spacing and stopping distance.	Provide information on the number of unsafe vehicles but cannot provide the degree of danger and the duration.	The calculation formula and dangerous threshold value are simple and clear.	Rear-end, hit object and turning.
Time Integrated DSS (TIDSS)	Total time integral of gap between DSS and the dangerous threshold value.	Mainly suitable for rear-end conflict.	Consider the degree and the duration of danger.	Same as DSS

Table 2.5: List of surrogate safety indicators based on the deceleration behavior

Indicator	Definition	Limitations	Advantages	Suitability for collision type
Deceleration Rate to Avoid the Crash (DRAC)	Differential speed between a following/response vehicle and its corresponding subject/lead vehicle (SV) divided by their closing time.	Fails to accurately identify the potential traffic conflict situation; Not suitable for lateral movement.	Explicitly considers the role of differential speeds and decelerations in traffic flow.	Rear-end, Hit object/parked vehicle, Hit pedestrian, Merging and diverging manoeuvres.

Table 2.5 continued from previous page

Indicator	Definition	Limitations	Advantages	Suitability for collision type
Crash Potential Index (CPI)	Probability that a given vehicle DRAC exceeds its maximum available deceleration rate (MADR) during a given time interval.	Not suitable for lateral movement; mainly applicable at intersection.	Address some of the issues found in DRAC like vehicle braking capability for prevailing road and traffic conditions.	Same as DRAC
Criticality Index Function (CIF)	Multiplication of vehicle speed with the required deceleration	Like TTC, it considers constant speed of consecutive vehicle; Further evaluation is needed using additional field data for validation.	Chance of occurrence and severity could be measured.	Crashes at turning, Right angle crashes

Furthermore, Generalizing the TTC to the planar (2D) case is surprisingly challenging [161]. Till recently, for situations other than car-following, there were no safety indicators available analogous to TTC. Ward et. al., [161] proposed an Extended Time-to-collision (ETTC) formulation for unconstrained/planar motion of vehicles. They applied a unique 'Loom-gating' method to filter out the collision courses. 'Loom-gating' is based on a parameter called 'loom-rate', which takes into account the rotational behavior of the possible conflicting vehicles to check if the collision could take place or not. In the work of Xing et. al., [57], a simpler version of the above ETTC formula was used to study conflicts upstream of a Toll Plaza, with many lane-changing maneuvers in a restricted space and time. However, we found that the available 2-Dimensional extensions of safety indicators are overestimating the risk. It is primarily because of the overestimation of vehicles' closing-in rate due to the improper projections of the velocity vectors.

The proactive safety metric that can be extended to all traffic situations and road geometries will have a wide range of applications. Some common applications, but not limited to, would be safety critical assessment of transport network operation, VRU safety analysis, and real-time crash risk estimation for CAVs. Several researchers have reported that vehicle-to-vehicle and vehicle-to-infrastructure communication with the help of the latest sensor technologies enable the exchange of state information to improve the safety and efficiency of transport networks [161, 166–168]. Also, the real-time crash risk monitoring aid in understanding the crash precursors, thereby implementing proactive traffic safety management strategies to mitigate unsafe traffic conditions [24, 169, 170]. At present, predicting crash risk in real-time, irrespective of road geometry and traffic scenario, is still limited within an idea that is not yet ready for deployment [170]. The literature also reveals that, though several traffic conflict severity indicators have been developed over the past decades [152, 171, 172], they are not suitable for a network-level safety assessment as the road geometry, conflict types, and traffic scenarios are heterogeneous on a transportation network.

2.5 Scenario-Specific Proactive Safety Analysis Using Trajectory Data

It has been widely reported that surrogate safety indicators could be the most suitable alternatives for a proactive assessment of risk at various transport components and traffic facilities. In this study, we have analyzed the crash risk corresponding to four traffic scenarios, which are the most highlighted areas of research in the transportation literature. The scenarios selected for safety assessment in this study are:

1. Lane-changing (LC) maneuver
2. Powered Two-Wheelers in an urban environment
3. Vehicles traversing 3-Dimensional road geometry
4. Identification of hotspot on a road transportation network

The following sections give a comprehensive review of the literature corresponding to each scenario.

2.5.1 Lane Change Dynamics and the Associated Risk

LC is one of the complicated driving tasks as it comprises multiple vehicle interactions, tactical decision-making, and low-level trajectory planning [173,174]. During an LC, a driver goes through several stages: decision-making, preparation, execution, and interaction with other vehicles [19,175]. Any mistake made during any of these stages may put the driver in danger or cause discomfort. It has been reported that the LC maneuver during unsafe traffic conditions develops conflict scenarios with different surrounding vehicles, thereby increasing the risk of a potential crash [19,176]. The LC activities has a significant impact on the traffic system's safety and capacity [177,178].

The crashes during LC maneuver account for about 4% to 10% of all traffic crashes [179]. Although this value is low, the delay due to LC crashes accounts for 10% of the total delay due to all traffic crashes [180]. Nonetheless, any crashes must be avoided because they result in the loss of life or impose a financial burden. To ensure safety, potentially risky LCs must be identified in advance [176,179]. Furthermore, due to the negative consequences of LC crashes, researches on developing control/warning strategies for dangerous LC behavior has received much attention [181,182]. Although LC safety has been researched for decades, most studies were focused on determining the overall safety during LC using historical data [154]. Such studies have ignored how the interaction between the LC vehicle and its surrounding vehicles impacts the crash likelihood and severity [176,183,184]. Therefore, a comprehensive LC safety research is still lacking, which is vital in preventing LC crashes, reducing property damage and life loss [176].

The instantaneous risk estimation of lane change is a complex problem because lane change safety varies greatly depending on vehicle velocity and traffic conditions [176]. The current LC models, which are predominantly deterministic and rule-based [185–188], can capture presumed cognitive driving behaviors, but they do not necessarily correspond to reality [189]. In reality, the LC dynamics could be highly stochastic and involve multiple vehicle interactions both in longitudinal and lateral directions, simultaneously. Until recently, studies have only considered the impact of the leader vehicle in the current and the target lane to evaluate the LC risk [180]. Many studies have already reported that, during an LC, the lag vehicle in the target lane poses a greater risk than the leader vehicles in the current or target lane [179,180,190–192].

The interruptions due to the surrounding vehicles during LC can cause significant differences in the LC behavior compared to the uninterrupted LC [143]. Moreover, to the knowledge of the authors, only a few studies differentiated the Uninterrupted LC (ULC) and interrupted LC (ILC) for a comprehensive understanding of LC dynamics [193]. Due to the absence of relevant literature on ILC, this review primarily focuses on the microscopic aspects of LC operation, which have been used to characterize ILC and ULC. To understand and model different aspects of the LC dynamics, researchers have used various attributes of LC. For example, comfort and safety are

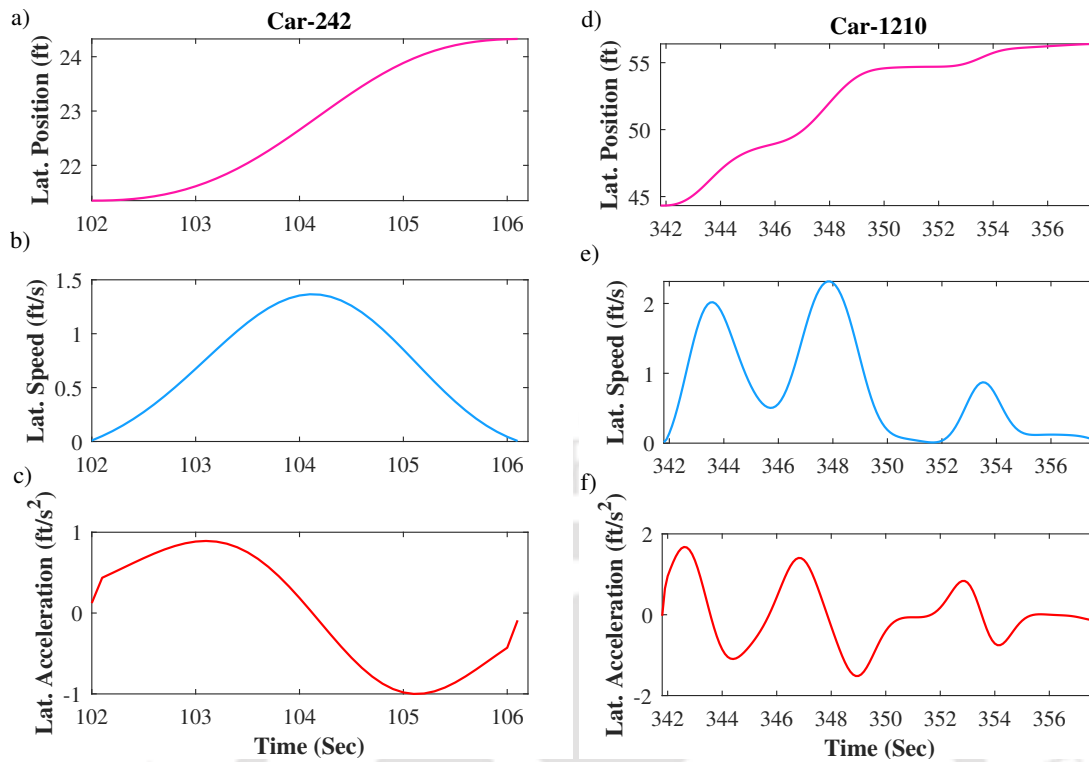


Figure 2.4: Field observed lateral position, speed, and acceleration profiles corresponding to LC maneuver (NGSIM, US-101 data); a & d) Lateral position profiles; b & e) lateral speed profiles; c & f) lateral acceleration profiles; corresponding to vehicle IDs Car-242 and Car-1210, respectively

the main attributes considered while developing a dynamic lane change path planning algorithm [143]. The LC duration is an important input for LC modeling [194]. To extract various attributes from LC trajectories, the time window of LC operation needs to be identified accurately. Table 2.6 shows some of the parameters used to identify LC window from the trajectory data. It is to be noted that, many past works [194, 195] followed a manual detection of LC window, which could be erroneous due to human error. Evidently, the method used to determine the LC window has a significant impact on the LC analyses, especially on the LC duration (Table 2.7). Thiemann et al. [99] have stated that *"since the real beginning of a lane change—the decision for making the lane change—is impossible to measure, and the physical beginning—the moment at which the driver starts to turn the wheel—is very difficult if not impossible to measure..."*. The difficulty associated with the LC window estimation is evident from this statement.

Table 2.6: Parameters used to identify LC window

Author & Year	Parameter
Jin et al., (2010) [135]; Toledo et al., (2007) [194]	The curvature of lateral position profile of LC vehicle
Balal et al., (2014) [195]	Time at which LC vehicle touches the lane marker
van Winsum et al., (1999) [196]	Steering angle
Wang et al., (2014) [197]; Balal et al., (2016) [198]	Lateral speed profile

Most studies treat LC as an instantaneous process or consider a constant LC duration across all vehicles for simplicity [19, 197]. In reality, the LC duration varies from one LC to another due to vehicle interactions and driver preferences. Since the LC duration is a critical variable in various LC-related studies, including safety studies, operational analysis [194, 199, 200], trajectory planning [143], decision-making [201], and microscopic simulation, it is vital to determine the necessity of separate duration modeling for ILC and ULC.

2.5.2 Proactive Safety Assessment of Powered Two Wheeler (PTW)

Powered two-wheelers (PTWs) are becoming more popular worldwide as a viable alternative to cars, especially in congested traffic situations [206, 207]. PTWs occupy less space and offer alternate means of transport for low-income commuters or where a less established transit system is in operation [208]. The key attractions of PTWs are their flexibility, lightweight, affordability, and the comfort to park. Despite these benefits, the PTW riders are much more sensitive to the changes in the surrounding traffic conditions than the rest of the vehicles, and they frequently move and avail the available inter-vehicle spacing for a faster movement. Owing to their narrower width, PTW trajectories and riding behavior vary significantly from those of other vehicles in a multimodal urban environment. This behavior of PTWs is more prominent during overtaking, lane-sharing, and tailgating [207, 209–212]. On the other hand, these unpredictable navigation patterns often cause inhomogeneity in the traffic stream, altering traffic flow characteristics at both the microscopic and macroscopic levels [207, 213].

Studies have already mentioned that the PTWs pose a significant threat to road safety [2, 206, 207, 214]. Each year, more than 286,000 PTW users die worldwide, accounting for roughly 23% of all road traffic fatalities [214]. This alarming number of potentially avoidable deaths emphasizes the importance of paying greater attention to PTWs and their usage in road safety policy. It is noteworthy that the use of PTWs increases year after year worldwide and is expected to increase in the future also [207, 215, 216]. However, despite substantial traffic safety changes for all road users, motorcycle riders' exposure to road risk has increased. As per reports, global road fatalities reduction from 2001 to 2016 for other road users was 53%, whereas for PTW, fatalities reduction was 37% [206]. Though PTW fatalities concern societies almost everywhere, the high percentages and absolute numbers of PTW fatalities in Southeast Asia, China, and India necessitate more efforts to improve PTW safety [217]. To effectively reduce PTW injuries in a specific region, it is vital to identify the most frequent collision patterns and crash configurations in that region [217, 218].

Considerable research has been performed on PTW crashes [216, 219–223]; however, there has always been a debate among practitioners regarding the types of crashes that should be targeted first and how resources for preventive measures should be prioritized. To prepare and execute highway safety improvement programs, it is essential to recognize factors that influence crash injury severity and their causal relationship. The number of studies on crash severity prediction is growing with time [224–227], showing how important this topic is to the scientific community.

Table 2.7: LC duration reported in various studies

Author & Year	Duration Statistics
Toledo et al., (2003) [202]	LC duration ranges from 1.0 to 13.3 sec, with a mean of 4.6 sec.
Thieman et al., (2008) [99]	5 to 6 sec as an average LC duration
Ahmed et al., (2019) [203]	Average LC duration of 5.5sec.
Olsen et al., (2002) [204]	Average LC duration of 6.28sec with a standard deviation of 2.0sec
Wang et al., (2020) [205]	LC duration ranges from 1.6 to 20sec, with a median of 6.6sec, a mean of 7.0sec, and a standard deviation of 2.1sec
Yang et al., (2019) [19]	LC duration varied from 0.7sec to 16.1sec

There is also a need to recognize the PTW crashes that pose a higher safety risk to riders based on the frequency and severity of the crashes [216]. However, the crash categories used in crash databases provide an oversimplified description of crashes. More specifically, the majority of the existing crash reports only offer details of the crash severity based on a predefined severity level (e.g., head-on crashes are the most severe and rear-end crashes are the least severe), which is not adequate for a comprehensive understanding of the actual crash scenario [216]. Consequently, there is almost no information on how a crash-generating process can be handled. A confluence of factors causes a significant proportion of conflicts, and most of the time, more than one factor leads to a crash [228, 229]. Moreover, all these factors play a significant role during various stages of crash progression and influence the frequency or severity of crashes [229]. Studies have shown that the PTW riders filtering on urban roads are at risk of crashes almost four times higher (3.94 times higher) than riders on general traffic lanes [230, 231]. Passenger cars were reported to be the most common collision partners for a PTW, with the most common configuration being a car front colliding with the PTW's side [217].

Several studies have already reported the limitations of current proactive safety approaches in studying PTW safety due to the infinite conflict configurations that arise from the higher flexibility of PTWs [115]. Nevertheless, a detailed understanding of the risk-taking behavior of PTWs is essential to enhance the overall safety of a traffic stream [206, 232]. It is also essential to recognize the PTW crashes that pose a higher safety risk to riders based on the frequency and severity of the crashes [216].

2.5.3 Safety of Vehicles Traversing the 3D road Geometries

Highways passing through hilly terrain typically have complex geometry features, such as steep gradients or sharp curves, resulting in a significantly different driver behavior and vehicle performance than non-mountainous highways [150]. It has already been reported that the hilly terrain

poses significant challenges for drivers while controlling a vehicle, based on the perceived geometry [233, 234]. Though there were studies that estimated the crash likelihood at curve sections, most studies have combined all the crash types during the safety assessment. Bagdadi et. al., [235] has stated that the hazardous situations for a vehicle can be classified into several types based on the severity of the situation, the type of evasive action, and whether the other road users are involved or not.

Crash occurrence and its severity are often the result of complex interactions between factors such as driver characteristics, vehicle characteristics, roadway geometry, and environmental conditions [236]. According to EU statistics, rural roads account for most of the crashes (54%), despite having a significantly lower traffic volumes than urban areas [237]. More than 60% of all crashes on rural roads are single-vehicle collisions, with approximately 32% occurring in horizontal curves [238]. Studies have shown that, up to 72% of crashes at the curves are caused by improper speed selection or failure to maintain a proper lateral position [239–241]. Surprisingly, similar figures can be seen in different countries' crash reports; thus, it is a worldwide scenario indicating the importance of research on the vehicle's safety at 3D geometries.

Contemporary road safety strategies include proactive approaches to improve road infrastructure design, vehicle safety systems, legislation, and traffic participants' awareness [242]. The "self-explaining roads" is one of the most well-known concepts related to road infrastructure that encourages the drivers to adopt a behavior consistent with the road design naturally [243]. The concept aims to communicate information about the upcoming situation to drivers easily and intuitively, using various measures, such as road markings and traffic signs [244, 245]. Misperceptions of speed and curvature and attentional failures can compound drivers' inability to maintain proper lane position through the curve, resulting in a loss-of-control, head-on, or other types of crashes [239]. Wider lanes and shoulders have been suggested to make curves more forgiving, but this can also increase drivers' speeds, negating overall safety gains [239, 246]. Typically, the warning signs are placed ahead of a horizontal curve to alert drivers about the upcoming changes in the alignment, thereby eliminating the surprise element. Unfortunately, the impact of curve warning signs on drivers' risk perceptions was minimal and could achieve only a 6% increase in overall safety [247]. Shinar et al. [248] have installed curve warning signs on two high-crash prone curves but failed to change drivers' entry speed significantly. It has been suggested that one of the reasons for their limited effectiveness is their overuse, especially in low-risk situations [247].

Several studies have already reported the dependency of curve geometry on the frequency and severity of crashes [249–251]. Different crash types are possible at horizontal curves, which are: angled, rear-end, side-swipe, and Run-Off-Road (ROR) collisions. Nevertheless, limited studies have investigated the likelihood of different crash types on horizontal curves owing to the lack of a detailed crash data. It is also to be noted that, lack of adequate trajectory data extraction tools for complex geometry along with effective surrogate safety measures has restricted the proactive safety assessment at complex road geometries. To date, most studies have depended on speed based safety and geometric design consistency evaluation, which may not be helpful to quantify the risk exposure and severity of crashes. Given these considerations, it is necessary to comprehend the

significant contributors to crash likelihood and severity to improve mountainous freeway design, operation, and management, more effectively [150].

2.5.4 Transportation Network-Level Safety Analysis

The road network safety assessment consists of four essential interconnected components: identification of sites needing safety examination, diagnosis of safety problems, adoption of feasible treatments for potential treatment candidates, and prioritization of treatments, given limited budget [252]. In general, the transportation network components such as road segments, intersections, ramps, interchanges do not perform consistently in terms of safety. In fact, there is heterogeneity in crash frequencies and rates due to differences in the driving population, roadside features, weather, traffic conditions, and design considerations [253]. Road safety professionals have attempted to identify and improve the "high risk" locations through various techniques based on the crash history, to provide a safe driving environment. However, the inability to distinguish between sites that are truly risky and sites that have experienced random "up" fluctuations in crashes during a period of observation, is a challenge [253–255].

Identification of high risk locations, also known as the sites with promise, hotspots, black spots, crash-prone locations, or priority investigation locations, is an essential and routine exercise for enhancing roadway networks' overall safety [253, 256]. It has been achieved through various network screening techniques. The primary goal of network screening is to evaluate the roadway network by ranking and recognizing locations with the highest crash likelihood for implementing countermeasures [257]. There is a vast literature that concentrated on developing techniques for hotspot identification (HSID) [253, 256]. It is very important that the methods used to recognize sites requiring safety investigation should be effective. Otherwise, resources could be wasted on sites incorrectly identified as unsafe, and truly unsafe sites could go untreated if they are missed in this process [252, 255].

Although federal road agencies and transport departments seek to minimize crash counts across their road networks, safety enhancement programs focus mainly on preventing fatal and severe injury crashes as the cost associated with the fatal and severe injury crashes is significantly higher than that of minor and non-injury crashes [258]. Moreover, society's aggregate social cost of crashes is significantly higher than individual social costs [259]. Consequently, the high-risk sites must be identified considering both crash severity and crash frequency [260]. Previous research has shown that approaches based on the simplistic ranking of crash counts or crash rates generate a significant number of false positives due to the random fluctuation of crashes from year to year, which resulted in safety improvement implementation at relatively safe places [253].

Identifying sites that require investigation for possible safety treatment is one of the most critical aspects of infrastructure safety management. Extensive research has been conducted on this process to improve the efficiency of black spot detection techniques [252]. As most safety enhancement programs concentrate on avoiding fatal and severe injury crashes, recognizing high-risk sites must ideally involve estimating the severity and frequency of crashes [260, 261]. Previous stud-

ies have reported that the current hotspot identification methods are inconsistent in most cases [257]. Particularly, the traditional reactive approaches were ineffective in a low-volume road network with sparse crash occurrences [262]. To choose candidate locations for safety improvements, most highway agencies use some form of statistical analysis [263]. The quality control and crash severity methods, which involve actual or observed crash frequency, are the most commonly used methods for this purpose [261]. It was reported that the overall network safety funding criteria are sensitive to the method used for identifying hotspots besides the economic assessment criterion [261]. Highway agencies are expected to seek safety-need estimation more frequently in the near future, based on a more rational and accountable basis, such as maximizing some system-wide utility (such as average crashes saved per dollar), rather than the current approach based on historical spending levels [261].

In recent years, road network evaluation systems were introduced to track road safety standards as well as to demonstrate the need for improvements. Such systems follow predictive ratings that examine the safety of various road network and retrospective components, including risk mapping and performance monitoring. Road geometry data and historical crash data were mainly used to rate possible risk elements of a transport network. Nevertheless, due to the shortcomings of the existing safety evaluation systems, the transport network safety administrators are eager to switch to proactive network safety management approaches [264, 265]. It is mainly because the current approaches have to wait for crashes or need extensive real crash data for the safety assessment [262]. Therefore, constructive safety evaluation of transportation networks, leveraging appropriate surrogate safety metrics, will give more insight into the predictive rating.

2.6 Summary

It is evident from the review that there exist several gaps in the state-of-the-art corresponding to trajectory data collection and standardization, as well as in the area of trajectory-based proactive safety analyses. The major gaps identified in the literature by virtue of this review are:

1. There is a need for an image-based trajectory data extraction tool that can work irrespective of camera platform, terrain type, traffic condition. Besides, such a tool should provide highly accurate data with minimal manual interventions.
2. The observed trajectory data contains inevitable noise, added at several stages of data collection and extraction. This noise needs to be adequately removed for an accurate assessment of traffic dynamics. The existing trajectory reconstruction techniques pose certain limitations. Therefore, we need a new method that can accurately reconstruct vehicle trajectories.
3. The existing surrogate safety indicators are not suitable for assessing the overall safety of a road transportation system. Many of them could be applied only to specific scenarios. For a comprehensive and overall safety analysis at various transportation facilities, a new indicator needs to be developed, irrespective of crash type, traffic facilities, and traffic conditions.

4. To avoid unsafe situations, the vehicles perform evasive actions before encountering an unsafe situation by perceiving the risk in advance. There are no indicators that can capture the vehicles' evasive actions, and such knowledge is essential during collision warning systems development.
5. The existing studies on lane change safety do not consider the difference between interrupted and uninterrupted LCs. It is essential to know their characteristic difference for better modeling of LC. However, there is no method available to identify and classify the interrupted lane changes.
6. A comprehensive, proactive safety analysis is lacking, which would give an overall understanding about the safety of vulnerable road users (e.g., PTWs), or safety at various traffic facilities (e.g., horizontal curve superimposed with a vertical curve, transportation network).

This research work addresses the gaps listed above. The following chapters discuss contributions of this research work that fills the gaps in the state-of-the-art.

Chapter 3

SAVETRAX: A Tool for Trajectory Data Extraction

This chapter describes the trajectory data extraction tool, SAVETRAX, which has been developed as part of this research work. SAVETRAX is an image-processing-based, semi-automated vehicle trajectory extractor, which can be used for extracting trajectories from traffic video collected using any camera platform and from any traffic scenario. Section 3.2 provides a comprehensive discussion on each component of SAVETRAX. Section 3.3 presents the details of the controlled experiment carried out to test the reliability of SAVETRAX and the major findings. Section 3.4 summarizes the important finding corresponding to the proposed tool.

3.1 Introduction

This chapter deals with the first objective of this thesis work, which is the development of a *Semi-Automated VEHICLE TRAjectory eXtractor (SAVETRAX)*, and to create an extensive, publicly available trajectory database by sharing this tool with prospective users. SAVETRAX is a reliable trajectory data extraction tool, which is independent of road geometry, traffic scenario, require minimal human interventions. The SAVETRAX comprises three modules; namely the *Pre-processor*, *Tracker*, and the *Analyzer*. A detailed description of each module is given in Section 3.2. In order to evaluate the reliability and performance of SAVETRAX, a controlled experiment with an instrumented vehicle and a drone was performed. The SAVETRAX was tested for its reliability to extract trajectory data from the videos collected from both planar and non-planar terrain types. The vehicle kinematic data extracted from a high-end GPS device attached to a test vehicle was compared with the data for the same vehicle, extracted from the drone video, using SAVETRAX. The comparative analyses signify the potential of SAVETRAX in collecting accurate traffic data for various transportation research applications.

3.2 Semi Automated Vehicle Trajectory Extractor (SAVETRAX)

Considering the limitations of the existing image-based trajectory extraction framework, a novel framework that incorporates separate modules to resolve the above limitations was developed. The tool was developed and compiled as a MATLAB application, and named as **Semi-Automated VEHICLE TRAjectory eXtractor (SAVETRAX)**. Unlike the existing trajectory data extraction frameworks, *SAVETRAX* has several merits; *i*) it is unaffected by the camera angle; *ii*) independent of terrain type (e.g., plain terrain, hilly terrain); *iii*) independent of road geometry (e.g., mid-block section, horizontal/vertical curves, intersections, round-about); *iv*) user can define the vehicle classes apart from the pre-defined vehicle classes such as Auto Rickshaw, PTW, Bus, Car, Light Commercial Vehicle (LCV), Multi-Axle vehicles, Non-Motorized Vehicles (NMV), Truck, and Pedestrian; *v*) it works under all traffic conditions (e.g., free-flow, congested); *vi*) the lighting and weather conditions have minimal impact on the tracking accuracy. Moreover, unlike the existing trajectory extraction tools, *SAVETRAX* has a well-designed user-feedback process about the possible errors while extracting the trajectory data, which expedites the trajectory extraction task for the prospective users. Figure 3.1 shows the screenshot of the user-interface of *SAVETRAX*.

3.2.1 Components of SAVETRAX

SAVETRAX has three modules; namely, *Pre-processor*, *Tracker*, and *Analyzer*. Each module is designed to perform specific tasks, which are of utmost importance to achieve accurate trajectory data. The design of *SAVETRAX* and its features were decided after interacting with several researchers who work on image-based traffic analysis, besides the authors' experiences. A detailed description of the functionality of each module is described in the following subsections.

3.2.1.1 Pre-Processor

The primary functions of this module are to define tracking boundary and homography estimation for coordinate transformation. Many tracking algorithms were developed for long-term tracking applications, which will indefinitely search for the reappearance of the tracking object even if it goes out of the camera view, without terminating the tracking task. However, the trajectory data extraction from traffic videos needs to terminate a tracking task once the vehicle departed from the region of interest. In *SAVETRAX*, it is achieved by defining a tracking boundary by drawing a polygon on the video-frame. Adding the tracking boundary to the tracking algorithm helps to automatically stop the vehicle from tracking once it exits the predefined tracking boundary. Such a consideration would prevent tracker failure due to the departure of the vehicle from the camera view. The camera calibration was also performed in this module by mapping 2D point correspondence between the image plane and the real-world plane. A transformation/homography matrix is estimated at this stage using a minimum of four known points in the field and the corresponding pixel coordinates. The present study considers the Direct Linear Transformation (DLT) to estimate the point-point correspondence. In order to extend the application of *SAVETRAX* to a

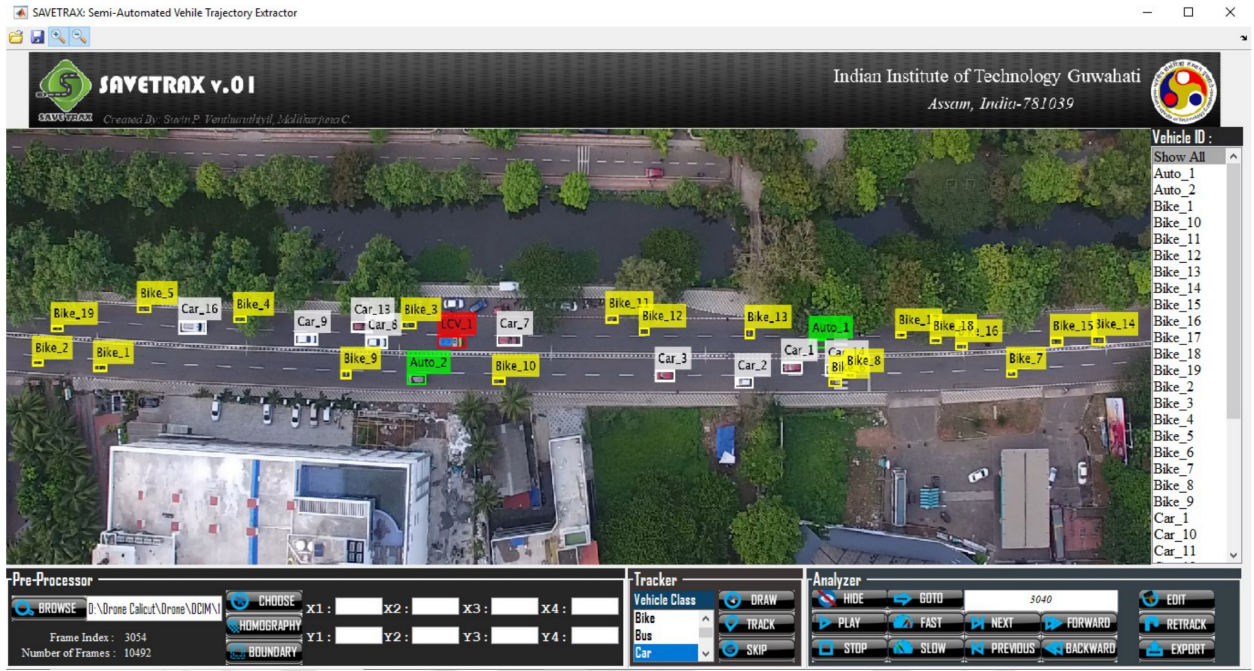


Figure 3.1: Screen short of SAVETRAX

non-planar terrain, multiple homographies were estimated after considering the non-planar terrain to piece-wise planar grids. Later, the homography matrix would be used to transform the vehicle trajectories in pixel coordinates to real-world coordinates.

Homography Estimation: Figure 3.2a shows a graphical representation of the real-world plane and image plane from a drone perspective. It is to be noted that the coordinates in the real-world system $\mathbf{x}_i$ are projected as $\mathbf{x}'_i$ in the image/pixel coordinate system. Between these homogeneous coordinate systems, a linear transformation matrix could be estimated as a non-singular 3×3 matrix using the relation $\mathbf{x}'_i \sim \mathbf{H}\mathbf{x}_i$, where $\sim$ indicates equality up to scale. This equation may be expressed in terms of the vector cross product as $\mathbf{x}'_i \times \mathbf{H}\mathbf{x}_i = 0$, which would enable to derive a simple linear solution for $\mathbf{H}$ as a minimization problem.

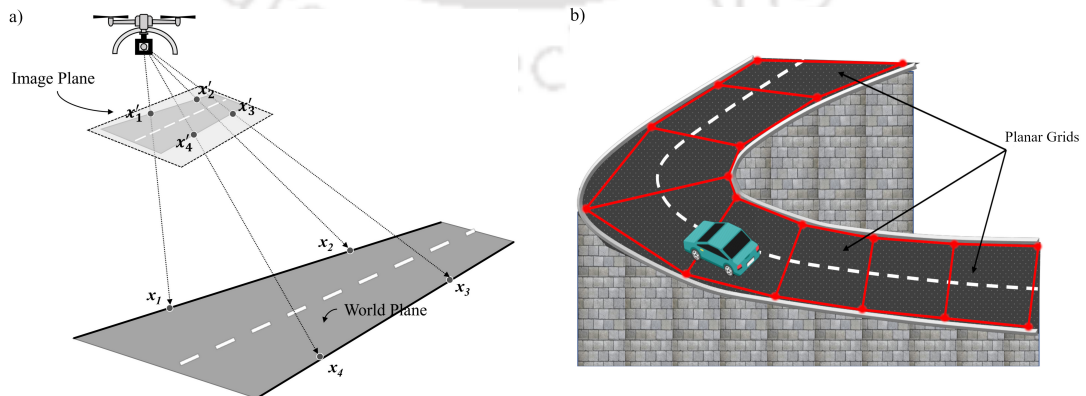


Figure 3.2: Point-point correspondence between real world and image plane; a) plain terrain type; b) non-planar terrain type

If the j^{th} row of $\mathbf{H}$ is denoted by $\mathbf{h}^{jT}$, and by writing $\mathbf{x}'_i = \begin{pmatrix} x'_i & y'_i & w'_i \end{pmatrix}^T$, the cross product may be given explicitly as;

$$\mathbf{x}'_i \times \mathbf{H}\mathbf{x}_i = \begin{pmatrix} y'_i \mathbf{h}^{3T} \mathbf{x}_i - w'_i \mathbf{h}^{2T} \mathbf{x}_i \\ w'_i \mathbf{h}^{1T} \mathbf{x}_i - x'_i \mathbf{h}^{3T} \mathbf{x}_i \\ x'_i \mathbf{h}^{2T} \mathbf{x}_i - y'_i \mathbf{h}^{1T} \mathbf{x}_i \end{pmatrix} \quad (3.1)$$

Since $\mathbf{h}^{jT} \mathbf{x}_i = \mathbf{x}_i^T \mathbf{h}^j$ for $j = 1, 2, 3$, this gives a set of three equations, which may be written in the following form

$$\begin{bmatrix} 0^T & -w'_i x_i^T & y'_i x_i^T \\ w'_i x_i^T & 0^T & -x'_i x_i^T \\ -y'_i x_i^T & x'_i x_i^T & 0^T \end{bmatrix} \begin{pmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{pmatrix} = 0 \quad (3.2)$$

These equations have the form $\mathbf{a}_i \mathbf{h} = 0$, where $\mathbf{a}_i$ is a 3×9 matrix, and $\mathbf{h}$ is a nine-element vector made up of entries of the matrix $\mathbf{H}$. In Equation (3.2), only two rows are linearly independent, thus it is common to omit the third equation while solving for $\mathbf{H}$ [266]. Then, the Equation (3.2) can be represented as;

$$\begin{bmatrix} 0^T & -w'_i x_i^T & y'_i x_i^T \\ w'_i x_i^T & 0^T & -x'_i x_i^T \end{bmatrix} \begin{pmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{pmatrix} = 0 \quad (3.3)$$

In this form, $\mathbf{a}_i$ is a 2×9 matrix. Given an estimate $\mathbf{h}$, the quantity $\|\mathbf{a}_i \mathbf{h}\|$ is the algebraic error of the i^{th} datum. Thus the homography matrix $\mathbf{H}$ can be estimated by minimizing the sum of squared algebraic errors as,

$$\hat{\mathbf{h}} = \underset{\mathbf{h}}{\operatorname{argmin}} \sum_{i=1}^N \|\mathbf{a}_i \mathbf{h}\|^2 = \underset{\mathbf{h}}{\operatorname{argmin}} \|\mathbf{A} \mathbf{h}\|^2 \text{ s.t. } \|\mathbf{h}\| = 1 \quad (3.4)$$

Where $\mathbf{A} \in \mathbb{R}^{2N \times 9}$ is obtained by stacking $\mathbf{a}_i$ vertically corresponding to all i . A norm constraint is added to prevent trivial solutions since the homographic transformation has only 8 degrees of freedom. Here the solution is the least significant right singular vector of $\mathbf{A}$ and is obtained through Singular Value Decomposition (SVD). Using the homography matrix $\mathbf{H}$ that is reshaped from $\mathbf{h}$, the vehicle position $\mathbf{x}'$ in pixel coordinate in the trajectory data could be transformed to real-world coordinates $\mathbf{x}$ by

$$\mathbf{x} = \mathbf{H} \mathbf{x}' \quad (3.5)$$

As the DLT based homography estimation is valid only for a planar region, it is essential to consider certain modifications to make it work for non-planar terrain. One straightforward way is to convert the non-planar region into piece-wise planar grids and estimate homography for each grid separately [267]. A non-planar road section can be effectively divided into piece-wise planar regions, as shown in Figure 3.2b. Since each planar region is sharing common vertexes, a smooth transition of the homography is achieved. Based on the position of the vehicle, the homography matrix can be chosen for coordinate transformation. SAVETRAX performs this task by automatically identifying the grid occupied by the vehicle.

3.2.1.2 The Tracker Module

The main tasks performed in this module include initial detection of the vehicle, classification, and tracking. The initial detection of the vehicle was done by drawing a bounding box around the subject vehicle when it first appeared in the frame. While constructing the bounding box, the vehicle class is also defined either by selecting from the drop-down menu in the SAVETRAX interface or by pressing the keyboard short-keys assigned to each vehicle class. Once the bounding box is drawn, a vehicle identification number is automatically generated to distinguish the vehicles. The user can fine-tune or delete the bounding box if it is not correctly defined.

The tracking algorithm considered in SAVETRAX is the ‘Efficient Correlation Operators (ECO) for Tracking’ [268]. The selection of ECO as the tracking algorithm was after considering the following benefits; i) it works on both CPU and GPU; ii) Tracking accuracy and speed are adequate for traffic data extraction applications; iii) Minimal system requirements compared to the other recent tracking algorithms. A detailed description of the ECO algorithm is given in the following section.

Efficient Correlation Operators (ECO) for Tracking: The ECO tracker was developed from the foundations of Continuous Convolution Operator Tracker (C-COT) [268]. In this formulation, a correlation filter was designed to learn from available negative data by including all shifted versions of the training sample. The convolution filter was discriminately learned from a collection of M training samples, $\{x_j\}_1^M \subset X$. Each feature layer $\{x_j^d\} \in \mathbb{R}^{N_d}$ has an independent resolution, N_d . Considering an interpolation model, the feature map is projected to a continuous spatial domain $t \in [0, T]$ by introducing an operator, J_d , as,

$$J_d \{x^d\} (t) = \sum_{n=0}^{N_d-1} x^d[n] b_d \left(t - \frac{T}{N_d} n \right) \quad (3.6)$$

Where, b_d is an interpolation kernel with a period $T > 0$. The resulting $J_d \{x^d\}$ is the continuous T-periodic function of the interpolated feature layer. The entire interpolated feature map is denoted as $J \{x\}$, where $J \{x\} (t) \in \mathbb{R}^D$. In the C-COT formulation, a continuous T-periodic multichannel

convolution filter $f = \{f^1, f^2, f^3, \dots, f^D\}$ was trained to predict the detection score.

$$S_f\{x\} = f * J\{x\} = \sum_{d=1}^D f^d * J_d\{x^d\} \quad (3.7)$$

Danelljan et al. [268] have observed that many of the filters, f^d , have negligible energy, which is apparent for high dimensional deep features. The contribution of such filters towards target localization is minimal and adversely affect the training time. To overcome the above, a factorized convolution operator, on a small set of basis filters $f^1, \dots, f^C$, where $C < D$, was considered instead of learning separate filter for each feature channel d . The filter corresponding to the feature layer d was then constructed as a linear combination of the filters f^c and a set of learned coefficients $p_{d,c}$ as,

$$\sum_{c=1}^C p_{d,c} f^c \quad (3.8)$$

The coefficients can be represented compactly as a $D \times C$ matrix, $P = (p_{d,c})$. The new multichannel filter is nothing but the matrix-vector product, Pf . Then the factorized convolution operator is,

$$S_{Pf}\{x\} = Pf * J\{x\} = \sum_{c,d} p_{d,c} f^c * J_d\{x^d\} = f * P^T J\{x\} \quad (3.9)$$

The factorized convolution (Eq.3.9) could alternatively be viewed as a two-step operation, where the feature vector $J\{x\}(t)$ is first multiplied with the matrix P^T at each location t . The resulting feature map with dimension C is then convolved with the filter f . The matrix P^T acts as a linear dimensionality reduction operator where the filter f and the matrix P jointly learned in a discriminative fashion, by minimizing the classification error (Eq.3.10) of the factorized operator (Eq.3.9).

$$E(f) = \sum_{j=1}^M \alpha_j \|S_{Pf}\{x_j\} - y_j\|_{L^2}^2 + \sum_{d=1}^D \|w f^d\|_{L^2}^2 \quad (3.10)$$

The corresponding loss in the Fourier domain for considering the factorized operator (3.9) is,

$$E(f, P) = \left\| \hat{z}^T P \hat{f} - \hat{y} \right\|_{l^2}^2 + \sum_{c=1}^C \left\| \hat{w} * \hat{f}^c \right\|_{l^2}^2 + \lambda \|P\|_F^2 \quad (3.11)$$

Where the $\hat{z}^d[k] = X^d[k] \hat{b}_d[k]$ is the Fourier coefficient of the interpolated feature map, $z = J\{x\}$. The Frobenius norm of P in Eq.3.11 acts as a regularization term, controlled by the weight parameter λ . Here the loss function (Eq.3.11) is a non-linear least squares problem. Due to the bi-linearity of $\hat{z}^T P \hat{f}$, the loss function is similar to a matrix factorization problem. Here, the parameter optimization is based on the Gauss-Newton method, and the quadratic sub-problems are optimized using the Conjugate Gradient method. The Gauss-newton method was derived by linearizing the residuals in Eq.3.11 using a first-order Taylor series expansion.

Generally, the DCF trackers add one training sample x_j in each frame j with an exponentially

decaying weight $\alpha_j \sim (1 - \gamma)^{(M-j)}$ controlled by the learning rate γ . The samples with the smallest weights are replaced when the number of samples reaches the maximum limit M_{max} . A larger sample limit needs to be set to obtain a representative sample-set. Moreover, collecting a new sample at each frame may lead to massive redundancies in the sample sets. The problem of sample redundancies due to continuous sampling strategy was tackled by considering a probabilistic generative model. For a given joint probability distribution $p(x, y)$ of the sample feature-map (x) and corresponding desired output score (y), a filter that minimizes the expected correlation error was identified. A predetermined Gaussian function was considered for the desired correlation output y from the sample x . This formulation has reduced the model complexity by reducing the number of training samples.

Usually, the DCF based trackers update after every addition of the features. In ECO, a sparser model updating scheme was considered instead of updating the filter at every frame, which is generally considered in the non-DCF trackers. In the ECO formulation, instead of explicitly searching for the objective change, the model was updated by performing optimization at every N_s^{th} frame. Danelljan et al. [268] have reported that, a moderately infrequent update of the model with $N_s \approx 5$ resulted in a significant improvement in the tracking result.

3.2.1.3 The Analyzer Module

This module performs mainly the following four tasks.

1. Identification and correction of faulty tracks
2. Camera motion compensation
3. Reconstruction of vehicle path data
4. Estimation of kinematic variables

Identification and Correction of Faulty Tracks: No vision tracking algorithm can perform the tracking task flawlessly. It is essential to identify and rectify such instances where the tracker fails and leave error in the extracted trajectories. The *Analyzer* module is an avenue to identify and correct such error in the trajectories. This module is capable of automatically identifying erroneous trajectory and inform the users for possible corrections through re-tracking. However, the users can also inspect the tracks manually to identify the tracking failure by playing the video with the annotations. For any situation where the annotated bounding box deviates from the vehicle could be re-positioned and re-tracked.

For all tracked vehicles, different consistency evaluation are performed and those trajectories not satisfying any of the conditions would be listed as a faulty track. The validation criteria considered are;

1. *Failed by meeting maximum track duration (τ_{max}):* The tracking algorithm was constrained to track maximum for a user-defined duration, which can be estimated using the average traffic speed (v_{avg}), maximum traveled distance (l_{max}), and maximum delay (w_{max}).

$$\tau_{max} = \alpha \times \left(\frac{v_{avg}}{l_{max}} + w_{max} \right)$$

where α is a factor greater than one. The consideration of maximum track duration was to restrict the continuation of an unexpectedly failed track. All the tracks that were active till τ_{max} would be considered as failed track. Users can later verify the track and perform necessary action.

2. *Internal consistency analysis:* It is already known that the vehicle kinematics should satisfy the laws of physics. Once there is a tracking failure, there could be an unrealistic acceleration event. This study assumes that all the tracks that show an acceleration/deceleration value outside the 97 percentile acceleration/deceleration value of all the vehicles could be corresponding to the tracker failure. All such vehicles are reported to the user for manual inspection.
3. *Trajectory overlap:* This is similar to the platoon consistency analysis performed by Montanino et al. [5] in their trajectory reconstruction procedure. The platoon consistency concept was proposed to tackle the scenario of false tracking that leads to trajectory overlapping, indicating a potential crash occurrence that might not occur in reality. However, the platoon consistency only looks for longitudinal overlap of vehicles. In our case, we extended the platoon consistency into a two-dimensional space and look for any overlap of trajectories both in longitudinal and lateral directions. The longitudinal and lateral spacing between vehicles was evaluated and verified for overlap by considering a minimum spacing threshold of 0.5m and 0.3m, respectively, for longitudinal and lateral direction.

Camera Motion Compensation: Unlike the video stabilization provided by different packages, the stabilization demanded by the trajectory data extraction problem is different. The camera movement would add a systematic bias to the extracted trajectories due to the superimposition of the camera trajectory with the vehicle/pedestrian trajectory. We have compensated the camera movements by projecting all the frames in the video to the orientation of a reference frame where the homography for coordinate transformation was estimated, using a global homography/projection matrix estimated using DLT. A planar projective transformation is a linear transformation on homogeneous vectors represented by a non-singular 3×3 matrix:

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (3.12)$$

The 3×3 matrix in Equation (3.12) represents the homography matrix used to transform the image views. The Direct Linear Transformation (DLT) method, already discussed in Section 3.2.1.1, is used for the estimation of global transformation/homography matrix [269, 270]. The SIFT key points were detected and matched for a pair of input images using the VLFeat library [271]. Mismatches among the point pairs $\{\mathbf{x}_i, \mathbf{x}'_i\}_{i=1}^N$ were removed using the Random Sample Consensus (RANSAC) algorithm [272]. The DLT approach has been used to estimate homography between the reference frame and all other frames considering the matched inlier point features. All the tracked positions in a frame are transformed to the orientation of the reference frame considering the estimated homography. Figure 3.3 shows the path of a moving camera and the path obtained after motion compensation.

Trajectory Reconstruction: It is already known that the observed trajectory data would always be embedded with certain noise. Therefore, it is essential to remove the noise for the final applications. Reconstructing the observed vehicle trajectories to a continuous, uniformly sampled, and noise free form is essential for several microscopic analyses. Several trajectory reconstruction methods are available in the literature [5, 64, 102, 109], but many are not suitable for a general application. In SAVETRAX, we incorporated one of the latest trajectory reconstruction method. The reconstruction method considered here is an adaptive, data-driven path smoothing technique that is independent of the camera platform (either drone or fixed camera setup), traffic state, or vehicle dynamics. A detailed description of the smoothing technique is provided in Chapter 4

It is a three-phase reconstruction process, and at every phase, the estimation of smoothing parameters is subject to the noise level in the observed path data. In the first stage, a resampling of path data is performed to retrieve any missing information. It was reported that the spline interpolation gives better results compared to the other interpolation techniques for retrieving missing data of vehicle trajectories (more details given in Chapter 3). It is also evident that as the size of occlusion increases, the prediction accuracy gets depreciated. Thus it is advisable to avoid occlusion of more than $10m$ while collecting the data to keep the trajectories accurate up to $5cm$. The outliers were also identified and removed at this phase. In the second phase, a Recursively Ensembled Low-Pass

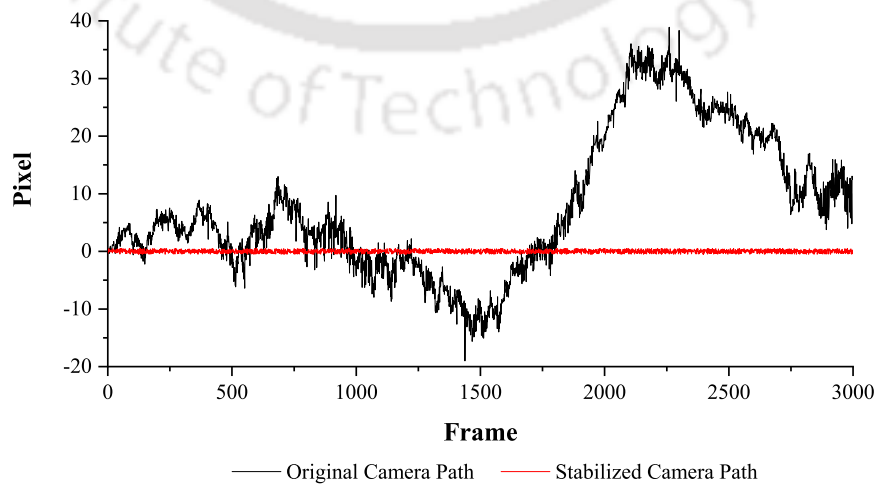


Figure 3.3: Comparison of the original and the stabilized camera paths

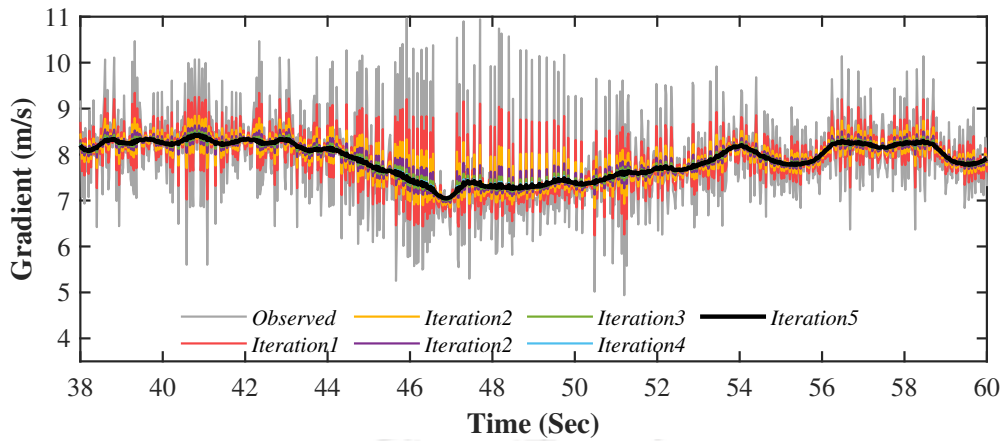


Figure 3.4: Result of smoothing of tracked data

filter (RELP) was used to curtail the heavy-tailed error, which adaptively denoise the gradient of the path. After this stage, the vehicle path would fulfill the internal consistency requirements to meet reasonable vehicle kinematics. The third stage performs the local reconstruction of the vehicle path using a tri-cubic kernel smoother. The optimal smoothing parameters, such as bandwidth and polynomial order, for local reconstruction was estimated through a grid-search algorithm. Figure 3.4 shows the results of the smoothing of the erroneous path data using the method considered in SAVETRAX.

Estimation of Microscopic Variables: In addition to the position data, SAVETRAX allows the users to export a wide variety of kinematic features from the traffic video. It includes the speed and its component in the longitudinal and lateral directions, acceleration and its component in longitudinal and lateral directions, and heading angle. The speed and acceleration are estimated using the numerical differentiation of the reconstructed position data. After estimating all the microscopic variables, the data can be exported as a .csv file.

3.3 Performance Evaluation

The accuracy of the trajectory data extracted using SAVETRAX was evaluated using a controlled experiment. A vehicle instrumented with V-Box was driven on a test bed, and monitored parallelly using a drone that has good gimbal support (DJI-Inspire 1). The height of flight of drone was 120m. The test setup is shown in Figure 3.5. The kinematic features of the test vehicle such as the speed and acceleration are extracted from the V-Box, using the RACELOGIC 2.13.5 (Build 211) 2003 software. In order to perform the comparison of the data from GPS logger and SAVETRAX, it is essential to synchronize the data. To synchronize the logged data, we have performed a correlation analysis to find the lag between the data. After estimating the time lag, one of the data was time-shifted for the synchronization. The synchronized speed and acceleration data obtained from both V-Box and SAVETRAX are presented in Figure.3.6. It is evident that the speed and acceleration obtained from SAVETRAX is closely matching with those obtained from high-end GPS device,

which indicates the higher accuracy of SAVETRAX.

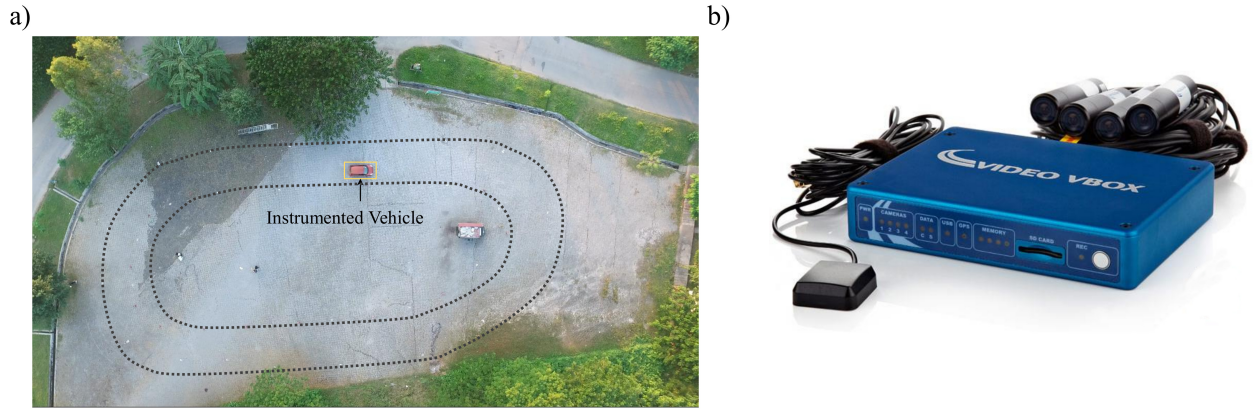


Figure 3.5: Experimental setup for validating the accuracy of SAVETRAX; a) test bed and the instrumented vehicle; b) V-Box

Further, we have quantified the accuracy of SAVETRAX, considering the data from an high-end GPS device as bench-mark. As sampling rate for V-Box (10Hz) and SAVETRAX (25Hz) were different, the error estimation was not straightforward. Therefore, the SAVETRAX data was downsampled to 10Hz, and estimated the mean deviation. It has been found that the mean difference of the reconstructed and the benchmark speed was observed as 0.495 m/s with a variance of $0.161 \text{ m}^2/\text{s}^2$.

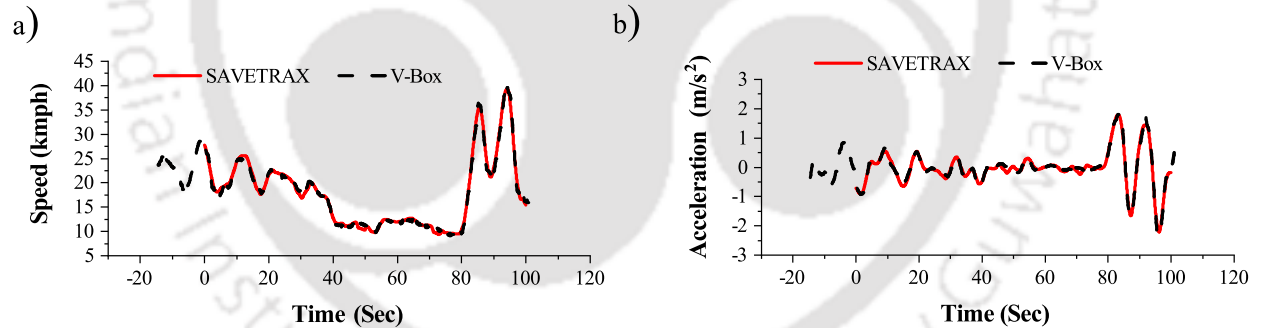


Figure 3.6: Comparison of speed and acceleration obtained from V-Box and SAVETRAX; a) speed; b) acceleration

We have performed another experiment to highlight the applicability of SAVETRAX for 3D geometries. The test vehicle was driven on a 3D geometry shown in Figure 3.7. For an accurate projection of the 3D geometry, the multiple homography was estimated as discussed earlier. For this purpose, the road stretch was divided into multiple cells and measured the corner coordinates using Total Station. The video was captured using a drone and extracted the trajectory of the test vehicle using SAVETRAX. In order to highlight the potential of the multiple homography approach, we have presented the estimated trajectories using single homography at different cell positions. It is evident from Figure 3.8 that the multiple homography approach estimates vehicle speed accurately. The comparison with the speed profile obtained from V-Box device indicates that the mean error in speed is less than 0.44kmph. However, the consideration of single homography

led to a higher error, and it increases as the vehicle moves away from the cell used for homography estimation. For example, it is evident from Figure 3.8 that the speed obtained from V-Box and that from SAVETRAX matches inside the cell considered for homography estimation (the cell boundaries are represented using dotted vertical lines in Figure 3.8a, b, & c, and a letter 'C' to represent the cell interior).

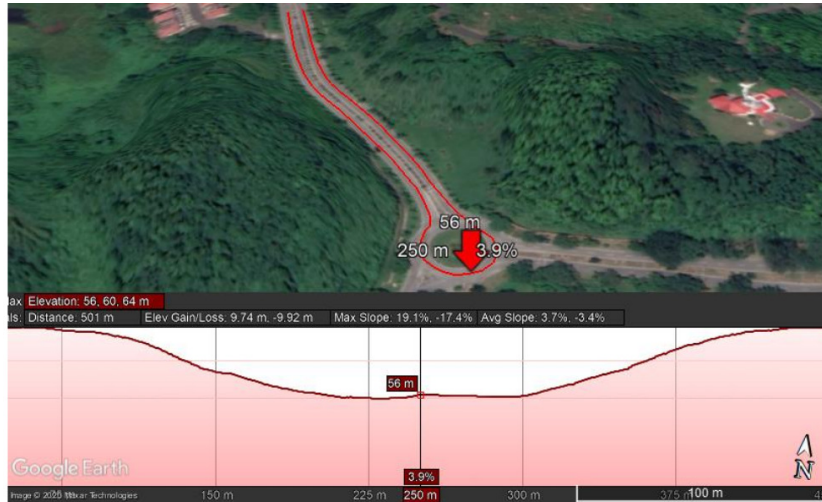


Figure 3.7: The test road passing through non-planar terrain

In addition, we tested SAVETRAX's performance in a variety of traffic conditions, camera angles, and illumination settings. The scenario and the retrieved trajectory are shown in Figure 3.9. The results show that SAVETRAX can extract vehicle trajectories with reasonable accuracy in a wide range of conditions.

3.4 Summary

The present study has developed a semi-automated, image-processing based vehicle trajectory extraction tool, the SAVETRAX, to achieve a better quantity and quality data with minimal manual efforts. This study has outlined several technical challenges associated with image-based trajectory data collection, devised necessary solutions, and integrated them into SAVETRAX. SAVETRAX consists of three modules, namely, the *Pre-Processor*, *Tracker*, and the *Analyser*. In the *Pre-processor*, camera calibration can be performed apart from defining the tracking boundary. The vehicle detection, classification, and tracking are performed in the *Tracker* module. Various corrections to the tracked vehicle path can be applied in the *Analyser* module, such as identification and correction of faulty tracks, camera movement compensation, reconstruction of the vehicle path to remove embedded noise, and estimation of various vehicle kinematic variables. One important feature of this module is the automatic identification of incorrect tracks based on the internal consistency of the vehicle's path. Such a consideration would ensure the extracted trajectory data is error-free. Some of the notable merits of SAVETRAX include; *i*) it is unaffected by the camera angle; *ii*) independent of terrain type (e.g., plain terrain, hilly terrain); *iii*) independent of road geometry (e.g., mid-block section, horizontal/vertical curves, intersections, round-about); *iv*) user can

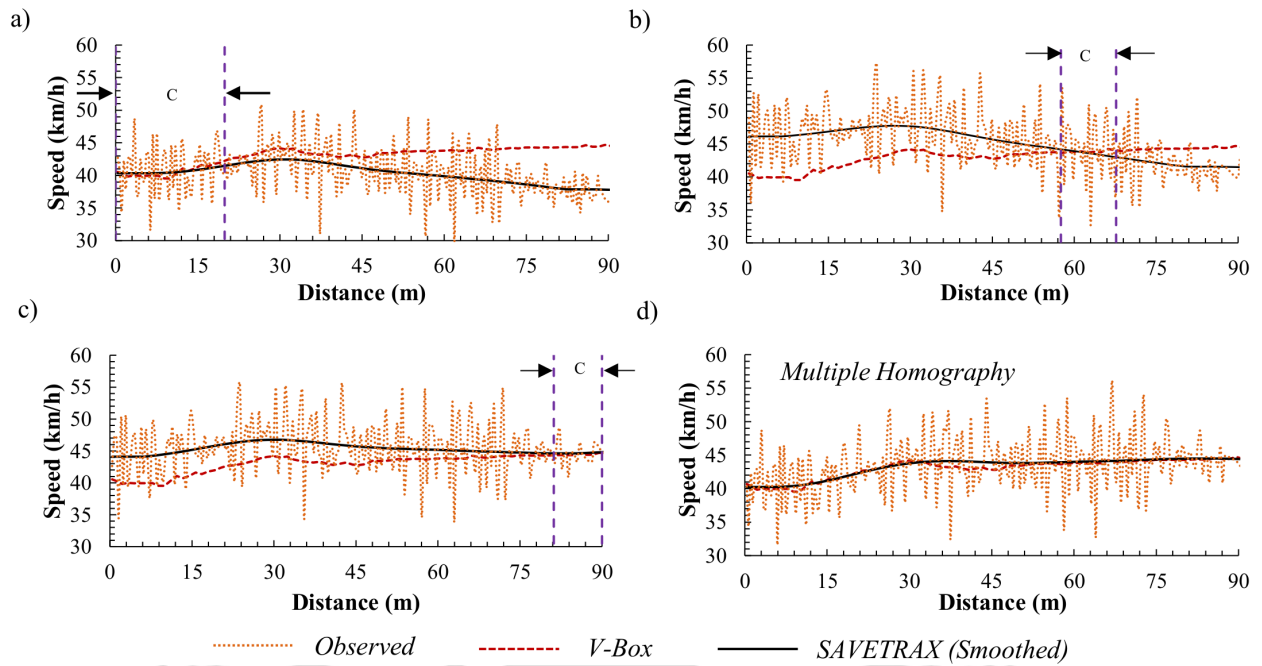


Figure 3.8: Comparison of the speed obtained from V-Box and from SAVETRAX corresponding to single cell (C) homography and multiple homography conditions; *a, b, & c*) Comparison of the speed obtained using single homography (at different locations on the curve) with the V-Box data; *d*) Comparison of the speed obtained using multiple homography and from V-Box

define the vehicle classes apart from the pre-defined vehicle classes such as Auto Rickshaw, Bike, Bus, Car, Light Commercial Vehicle (LCV), Multi-Axle, Non-Motorized Vehicles (NMV), Truck, and Pedestrian; *v*) it works under all traffic conditions (e.g., free-flow, congested); *vi*) the lighting and weather conditions have minimal impact on the tracking accuracy. Moreover, the optimal and compact design of the SAVETRAX, making the traffic data extraction process simple and reliable.

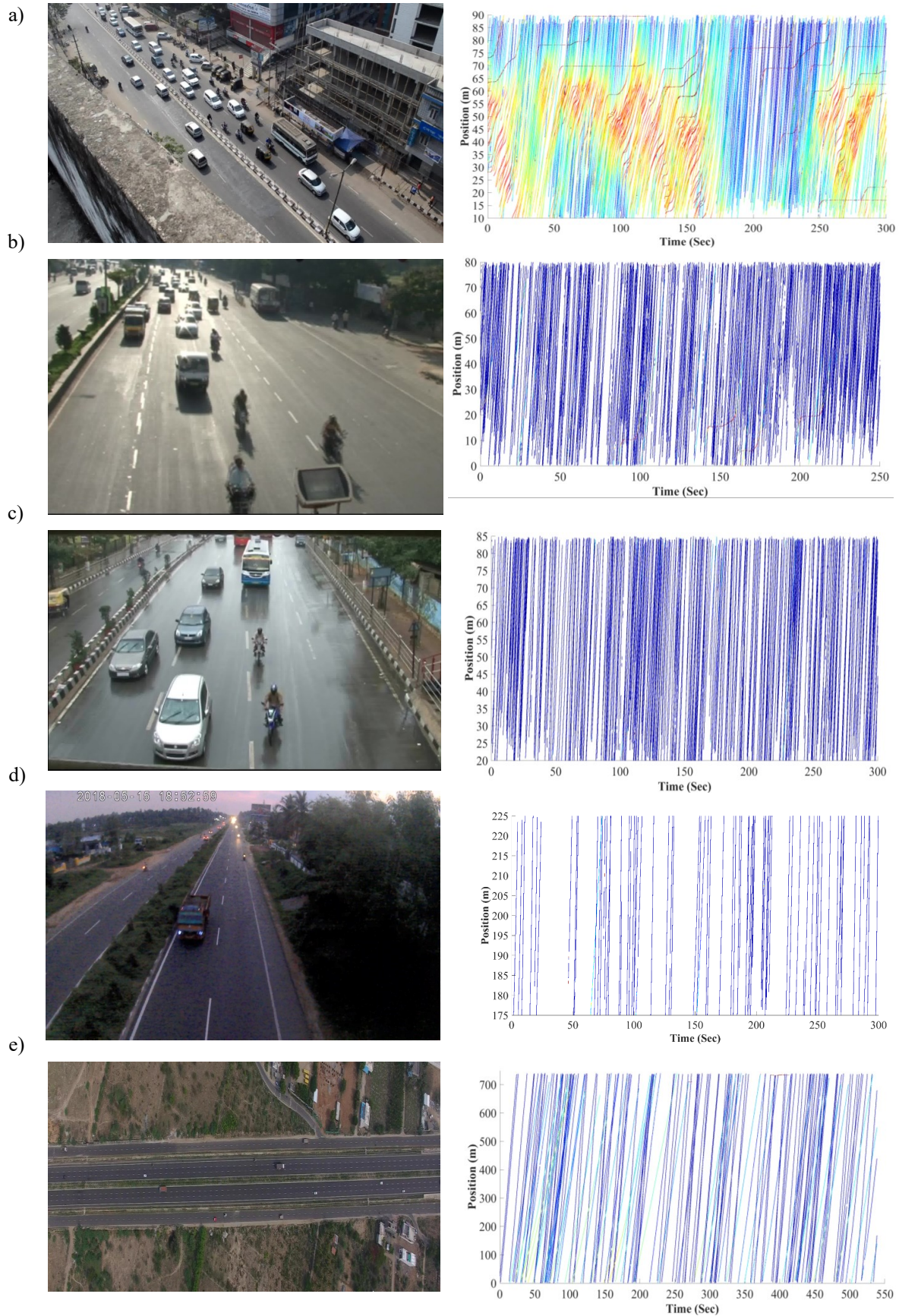


Figure 3.9: Performance of SAVETRAX at various traffic, camera, and lighting conditions; a) at different traffic states; b) with glare; c) bad weather; d) low light condition; e) bird-eye view (a-d corresponds to various camera angles)

Chapter 4

Vehicle Path Reconstruction using Recursively Ensembled Low-pass Filter (RELP) and Adaptive Tri-Cubic Kernel Smoother

Accurate information of paths followed by vehicles is essential for improving traffic systems' safety and efficiency. Recently, the availability of such data is plenty, thanks to the advancements in drone and computer-vision technologies. Nevertheless, extracting useful information from the observed data requires certain pre-processing to handle the noises added at several stages of the data collection. This chapter presents an adaptive and data-driven path smoothing technique that can be universally applied to the video-based path data obtained from any camera platform (either drone or fixed camera setup). Section 4.1 introduce the research problem. Section 4.2 discusses the detailed methodology of the reconstruction technique. Section 4.3 highlights the potential of the reconstruction technique through a performance evaluation. Finally, Section 4.4 summarizes the chapter.

4.1 Introduction

Trajectory data is an inevitable ingredient for developing and validating various Intelligent Transportation Systems (ITS) applications. However, the accuracy of such data significantly influences the performance of the models or tools developed/validated using such data. Benefiting from the advancements in computer vision and drone technologies, researchers have collected abundance of high fidelity vehicle path information, in fact at a broader spatiotemporal region [38,39,61,72,73]. Subsequently, access to a vast amount of trajectory data has unlatched a new avenue of research where various previously unknown vehicle dynamics and interaction patterns, at diverse driving conditions, could be explored. Nevertheless, extracting useful information from the observed data,

from any camera platform, requires certain pre-processing that can handle the noises added at several stages of the data collection. The direct use of the noisy trajectory data for deriving the attributes of vehicle kinematics and dynamics, such as speed, acceleration, jerk, or gap, would lead to unrealistic values, thereby leading to a misinterpretation of traffic dynamics.

Numerous methods have been developed to handle the noise added at several stages of trajectory data extraction, of which pros and cons are already discussed in Chapter 2. In this study, a novel reconstruction technique was developed to handle the unique feature of the error pattern found in the trajectory data collected using the latest technologies. The proposed reconstruction framework is an adaptive, data-driven, and camera platform independent approach. In this framework, the smoothing parameters' estimation is subject to the noise level in the observed path data. It is a three-stage reconstruction method; in the first stage, the time-series path data is resampled by retrieving any missing information or removing any repeated observations. Any outlier in the observed path data is also removed at this stage. The second stage curtails the heavy-tailed error function using the proposed Recursively Ensembled Low-Pass filter (RELP) technique, which adaptively denoise the gradient of the path. This stage ensures the internal consistency of the vehicle's path, thereby a reasonable vehicle kinematics. The third stage performs the local reconstruction of the vehicle path using a tri-cubic kernel smoother. A grid-search algorithm was also proposed to estimate the optimal smoothing parameters such as bandwidth and polynomial order. The performance and reliability of the proposed reconstruction framework were evaluated using different datasets. Further, the efficiency of the proposed framework was evaluated and reported through the consistency analyses.

4.2 Three Stage Method of Vehicle Path Reconstruction

We propose a three-stage framework for vehicle path reconstruction from the observed time-series of position, and an overview of the proposed reconstruction algorithm is shown in Fig. 4.1. The details of each stage of the reconstruction technique are described in the following sections.

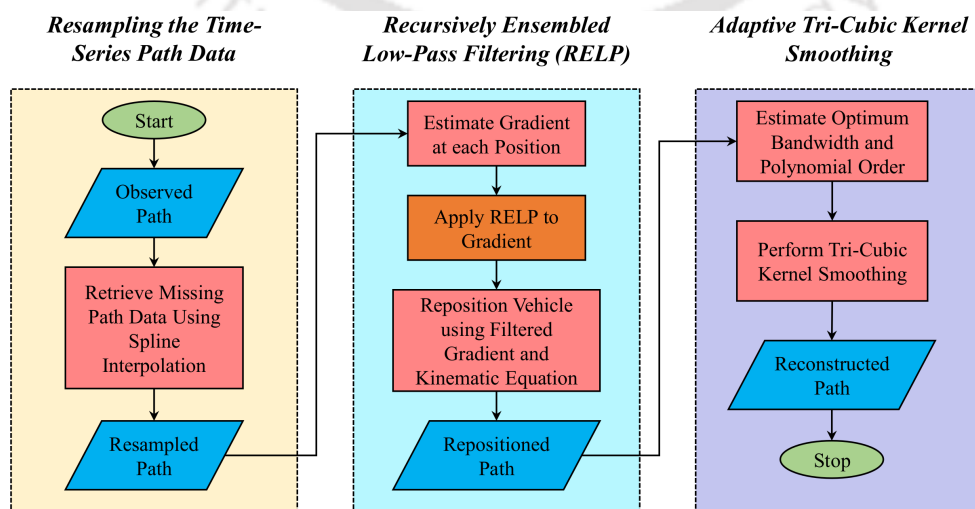


Figure 4.1: Algorithm flow chart for the proposed vehicle path reconstruction technique

4.2.1 Re-sampling the Time-Series Path Data

Sensor malfunctioning or occlusion during the vehicle path data collection would leave missing information in the observed path. Nevertheless, a uniformly sampled, but high fidelity data is essential for the microscopic car-following or lane-changing model calibration, training an autonomous car, and also for the data denoising using a low-pass filter. Thus, the missing data needs to be retrieved to make the observed data into a useful form. The present study uses the spline interpolation method to restore any missing information in the observed path data. The choice of interpolation method was made after a comparative assessment of different interpolation techniques, as presented in Fig. 4.6. Moreover, the spline interpolation is often preferred over the other polynomial interpolation methods due to the least interpolation error, even with a low degree polynomial spline [273]. It is also worth noting that the use of splines for the interpolation can avoid the so-called Runge's phenomenon, which leads to the oscillation between the points when higher degree splines are used for interpolation. Since our focus in this study is primarily to denoise the vehicle path data, we are limiting the further discussion on spline interpolation. However, more details about the spline interpolation technique can be found in the study of de Boor et al. [274].

Fig. 4.2 shows the algorithm flow chart of the missing data retrieval process considered in the present study. The performance of the retrieval process was evaluated using the manually extracted trajectory data reported by Coifman et al. [4], and the same is discussed in Section 4.3. After retrieving all the missing data, if any, the outliers in the observed path were removed by applying a median filter with a 1-second window. Further, RELP was used to achieve a kinematically consistent vehicle path. A detailed description of RELP is given in the following section.

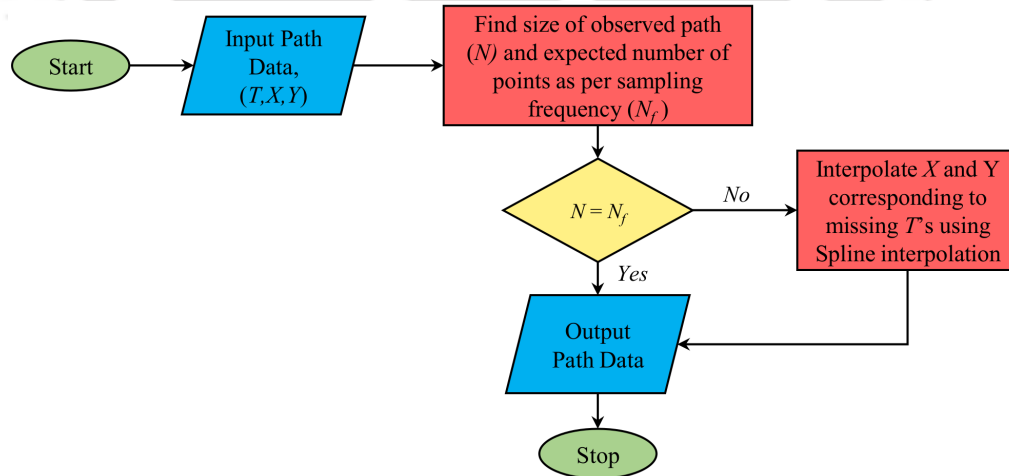


Figure 4.2: Algorithm flow chart of missing data retrieval process

4.2.2 Recursively Ensembled Low-Pass Filtering (RELP)

The observed vehicle path data are generally superimposed upon the indistinguishable errors, which are characterized as noises regardless of their source. The superimposed noise would alter vehicles' kinematic characteristics, making it inconsistent with the fundamental equations of mo-

tion. Nevertheless, the reconstructed vehicle path needs to follow the kinematic equations, which is often referred to as internal consistency. It is imperative to the traffic engineers that the removal of noise from the trajectories should be as much as possible, at the same time, retain the whole of the underlying information. We have already seen that the existing smoothing techniques consider a fixed window size, which may lead to over-smoothing or under-smoothing based on the window size. It is known that a larger window size leads to over-smoothing and deflections at boundaries, thereby the loss of important information on vehicle dynamics. However, the smaller window size can not handle errors of higher amplitude that are most common in drone-based trajectory data collection. Thus, it is essential to achieve a path smoothing, which can attenuate the localized, high amplitude errors, without losing the vital information of vehicle dynamics. To overcome this, we propose a Recursively Ensembled Low-Pass Filtering (RELP) framework, which recursively attenuates the noise in the observed vehicle path through a backward rectification of position error using the kinematic equation of motion (Eq.4.1a), and filtered path gradient.

$$x(t) = x(t_0) + \int_{t_0}^t v(t) dt \quad (4.1a)$$

$$v(t) = v(t_0) + \int_{t_0}^t a(t) dt \quad (4.1b)$$

The RELP is primarily a low-pass filter; however, it has certain additional features. A low-pass filter (LPF) is generally designed to attenuate frequencies in the input signal higher than a user-defined cut-off frequency, at the same time, to preserve the frequencies lower than the cutoff frequency. The frequency response of the filter depends on the filter design. In the path reconstruction problem, as discussed before, the filter needs to be designed for each vehicle separately based on the noise characteristics. As, either the error frequency or frequency of the signal-of-interest is unknown, the design of filter with appropriate cut-off frequency for different vehicles is cumbersome. We resolve this issue through recursive ensembling strategy added in RELP. The Savitsky-Golay (SG) low-pass filter was considered as the base filter for RELP, with a bandwidth (h_{SG}), which is the reciprocal of sampling frequency (f_s).

$$h_{SG} = \frac{1}{f_s}$$

The selection of one-second window for the base filter was based on the assumption that in a one-second window, the vehicle dynamics can be modeled using a lower order polynomial.

SG filter is the most frequently used digital filters in spectrometry due to its wide range of mathematical properties [275]. SG filter works on the method of least squares, however, it may be used without additional computation complexity, and with considerable improvement in the information obtained [276]. The main attraction of SG filter, as shown by Savitzky et al. [276] is that, fitting a polynomial to a set of input samples and then evaluating the resulting polynomial at

a single point within the approximation is equivalent to discrete convolution with a fixed impulse response, which is depicted in Figure 4.3.

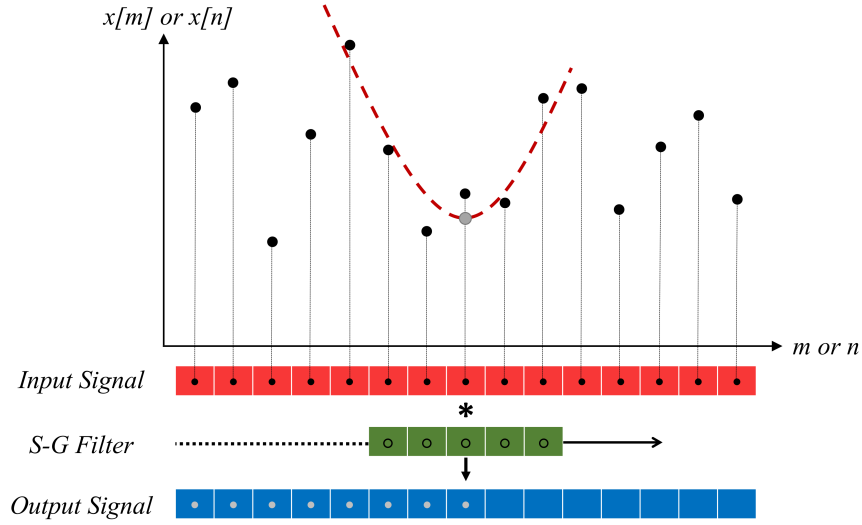


Figure 4.3: Illustration of S-G filter; $\bullet$ denotes the input samples; $\circ$ denotes the filtered sample; $\circ$ denotes the effective impulse response sample (weighting constant); and $*$ denoted the convolution operation

Before endeavoring into the formulation of RELP, let us begin with the conceptualization of an SG filter. Let $p(n)$ be the N^{th} order polynomial fitted to $2M + 1$ samples centered at $n = 0$ from a sequence of samples $x[n]$, then,

$$p(n) = \sum_{k=0}^N a_k n^k \quad (4.2)$$

Where a_k , $k = 0, 1, \dots, N$, are the coefficients of the polynomial $p(n)$. The coefficients are generally estimated as a minimization problem, where the mean square approximation error (ϵ_N) is minimized,

$$\epsilon_N = \sum_{n=-M}^M \left(\sum_{k=0}^N a_k n^k - x[n] \right)^2 \quad (4.3)$$

From the fitted polynomial function, the filtered point is evaluated corresponding to $x[n = 0]$ (the central point), which is nothing but $p(0)$,

$$\hat{x}[0] = p(0) = a_0$$

This process is repeated to all the sample points by shifting the filter window, to produce $\hat{x}[n]$.

The original paper by Savitzky et al. [276] observed that the output sample can be computed by discrete convolution of the form,

$$\hat{x}[n] = \sum_{m=-M}^M h[m]x[n-m] = \sum_{m=n-M}^{n+M} h[n-m]x[m] \quad (4.4)$$

where, Eq.4.4 replaces the polynomial fitting at each sample described in Eq.4.2 with a single

evaluation. The equivalency of least square estimation and single convolution of impulse response is proved in Theorem 1.

Theorem 1. For discrete observations $x[n]$, the least square polynomial estimation of smoothed version $\hat{x}[n]$ using a least-square polynomial approach with polynomial order N and window size $2M + 1$, where $N \leq 2M$, is equivalent to estimating a single finite impulse response (FIR).

Proof. To prove, let us begin with the least square polynomial estimation. The optimal coefficient of polynomial (Eq.4.2) could be determined by differentiating ϵ_N in Eq.4.3, with respect to each of the unknown coefficients and setting the corresponding coefficient to zero. This yield for $i = 0, 1, \dots, N$,

$$\frac{\partial \epsilon_N}{\partial a_i} = \sum_{n=-M}^M 2n^i \left(\sum_{k=0}^N a_k n^k - x[n] \right) \quad (4.5)$$

by interchanging the order of summation, Eq.4.5 becomes a set of $N + 1$ equations with $N + 1$ unknowns;

$$\sum_{k=0}^N \left(\sum_{n=-M}^M n^{i+k} \right) a_k = \sum_{n=-M}^M n^i x[n], \quad i = 0, 1, \dots, N \quad (4.6)$$

Eq.4.6 is known as the normal equation in the least square approximation problem.

For additional insight, let us represent Eq.4.6 in the matrix form. Let us define a design matrix $\mathbf{A}$ of dimension $(2M + 1) \times (N + 1)$ as,

$$\mathbf{A} = \begin{bmatrix} (-M)^0 & (-M)^1 & \dots & (-M)^N \\ (-M+1)^0 & (-M+1)^1 & \dots & (-M+1)^N \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ (M-1)^0 & (M-1)^1 & \dots & (M-1)^N \\ (M)^0 & (M)^1 & \dots & (M)^N \end{bmatrix}$$

which is nothing but the Vandermonde matrix. Let the vector of input sample be, $\mathbf{x} = [x[-M], \dots, x[-1], x[0], x[1], \dots, x[m]]^T$, and define $\mathbf{a} = [a_0, a_1, \dots, a_N]^T$ as the vector of polynomial coefficients. The Eq.4.6 in matrix form would be,

$$\mathbf{A}^T \mathbf{A} \mathbf{a} = \mathbf{A}^T \mathbf{x} \quad (4.7)$$

then the solution of the polynomial coefficients is,

$$\mathbf{a} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{x} = \mathbf{H} \mathbf{x} \quad (4.8)$$

where $\mathbf{H} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$. Now, note that the output for the group of samples centered at $n = 0$ is $y[0] = a_0$; i.e., we only need to obtain the coefficient a_0 . Which means that, we are only interested in the 0^{th} row of the $(N + 1) \times (2M + 1)$ matrix $\mathbf{H}$. By the definition of matrix multiplication, a_0 is nothing but the linear combination of $(2M + 1)$ elements of the $(2M + 1) \times 1$ column vector of $\mathbf{x}$.

$$y[0] = a_0 = \sum_{m=-M}^M h_{0,m} x[m]$$

Where $h_{i,n}$ denotes the elements of the $(N + 1) \times (2M + 1)$ matrix $\mathbf{H}$ and, $h_{0,m}$ is an element of the 0^{th} row. Consequently, comparing this equation to the second term of Eq.4.4 with $n = 0$, we observe that,

$$h[-m] = h_{0,m}, \quad -M \leq m \leq M$$

It is noteworthy that $\mathbf{H}$ depends only on M and N , and independent of the input samples. Therefore, the same weighting coefficients will be obtained at each group of $2M + 1$ input samples, hence we can think of least square smoothing as a shift-invariant discrete convolution process. $\square$

Since the weighting matrix $h_{0,m}$ is independent of the input sample $\mathbf{x}$, it can be estimated corresponding to a unit impulse centered at $-M \leq m \leq M$, and solve for all the coefficients of corresponding polynomial approximation.

Let $\tilde{\mathbf{a}}$ be the approximation of coefficient vector for the unit impulse input $\mathbf{d}$, then,

$$\tilde{\mathbf{a}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{d} \quad (4.9)$$

where $\mathbf{d} = [0, \dots, 0, 1, 0, \dots, 0]^T$ is a $(2M + 1) \times 1$ column vector impulse. Then, for the impulse input $\mathbf{A}^T \mathbf{d}$ is a $(N + 1) \times 1$ column vector. Let,

$$(\mathbf{A}^T \mathbf{A})^{-1} = \begin{bmatrix} \tilde{a}_0 & \tilde{a}_1 & \dots & \tilde{a}_N \\ \tilde{a}_1 & \bullet & \dots & \bullet \\ \vdots & \vdots & \dots & \vdots \\ \tilde{a}_N & \bullet & \dots & \bullet \end{bmatrix}$$

where matrix entries denoting $\bullet$ do not enter into the computation of $\tilde{\mathbf{a}}$. From the definition of matrix multiplication, the 0^{th} row of the matrix $\mathbf{H} = (\mathbf{A}^T \mathbf{A})^{-1}$ is,

$$[h_{0,-M}, h_{0,-M+1}, \dots, h_{0,0}, \dots, h_{0,M}] = [\tilde{p}(-M), \tilde{p}(-M+1), \dots, \tilde{p}(0), \dots, \tilde{p}(M)]$$

where $\tilde{p}(n)$ is the polynomial fit to the unit impulse,

$$\tilde{p}(n) = \sum_{k=0}^N \tilde{a}_k n^k, \quad -M \leq m \leq M \quad (4.10)$$

Then, the impulse response of the SG filter is,

$$h[-n] = h_{0,n} = \tilde{p}(n) \quad (4.11)$$

Equation 4.11 gives $h[-n]$ since the impulse response is flipped around $n = 0$ while evaluating the discrete convolution. The convolution formulation of SG filter could be easily implemented in MATLAB. In our formulation, we implemented Eq.4.4, as the impulse response formulation in Eq.4.11, recursively. At every iteration, the resulting $\hat{x}[n]$ ensembled with the input $x[n]$ as,

$$\hat{x}^i[n] = \sum_{m=n-m}^{n+m} h[n-m] \left(\frac{\hat{x}^{i-1}[m] + \hat{x}^{i-2}[m]}{2} \right), \quad i = 1, 2, \dots \quad (4.12)$$

with $\hat{x}^i[n] = x[n]$ if $(i-1) \leq 0$ or $(i-2) \leq 0$, where i is the index of iteration. The recursion is continued till reaching the stopping criteria, which is; $|S_{in} - S_{out}| = \frac{f_s}{5 \times L}$, where S_{in} and S_{out} are the input and output signals, respectively. An algorithm flowchart of the proposed RELP is shown in figure 4.4.

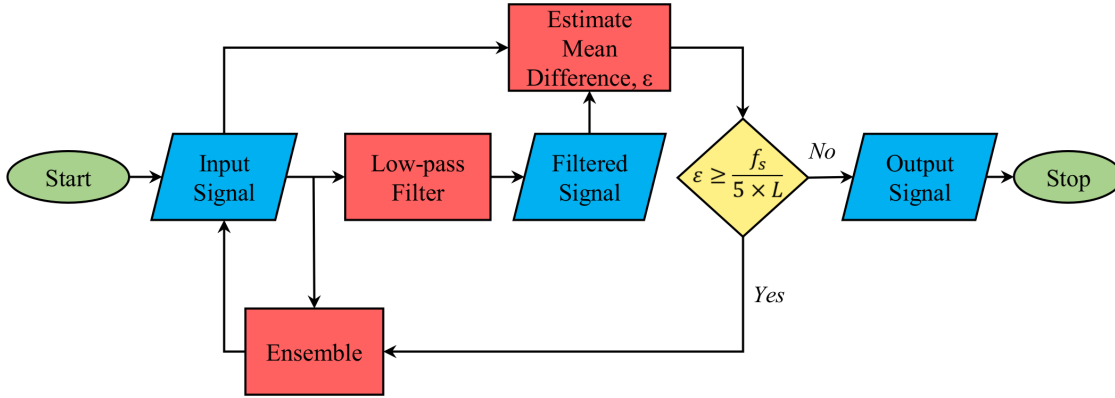


Figure 4.4: Algorithm flow chart of RELP

To address the dissimilar error in the longitudinal and lateral directions, the RELP filter is applied individually to the longitudinal and lateral gradient of the vehicle path. The gradient is estimated from the time-series position data through numerical differentiation. Let, (T, X, Y) be the time-series of vehicles' path data, then the lateral and longitudinal components of the gradient are,

$$G_x = \frac{dX(t)}{dt}, \quad G_y = \frac{dY(t)}{dt}$$

The noise-free gradient of the path is nothing but the speed of the vehicle (i.e., $G_x = v_x$, and $G_y = v_y$). It is obvious that an exact and physically pertinent vehicle path can only be obtained

through ascertaining a realistic and reasonable speed profile along the path. Thus it can be stated that the primary function of RELP is to achieve the speeds v_x and v_y from the noisy gradient data G_x and G_y . It is to be noted that, due to the applied filtering, there could be deformations at the boundaries of time-series data. In order to handle this, the boundaries are padded using the reflection method in the formulation of RELP. In the reflection methods, the boundaries are extended by replicating the values along the ends, i.e., $G(N+1), (G(N+2), \dots$ are replaced using $G(N), G(N-1), \dots$. This would help in minimizing the boundary effects. Using the speed obtained after RELP, the vehicle position is adjusted using the kinematic equation of motion given in Eq.4.1a, by which an internally consistent vehicle path is achieved. However, certain white Gaussian noise could remain in the path data, and the same is removed using an adaptive tri-cubic kernel smoothing. A detailed description of Kernel smoothing is provided in the following section.

4.2.3 Adaptive Tri-Cubic Kernel Smoothing

Even after the first two steps, which would be ensuring equispaced time-series of path and conditions of internal consistency, certain white Gaussian noise would be remaining in the path data. The data need to be free from such noises to achieve a smooth and continuously differentiable vehicle path. Kernel smoothing techniques are desirable to handle the white Gaussian noise after an appropriate selection of the smoothing parameters such as the bandwidth and polynomial order. Compared to the filtering based trajectory/path smoothing technique that suffers from severe boundary effects, the kernel-based technique could automatically maintain the same bias order at the interior and boundaries [277]. We consider a tri-cubic Kernel function in our method, due to its better shape-preserving property while smoothing, as reported by Cleveland et al. [64, 278]. However, the impact of the shape of the kernel on smoothing is minimal compared to the other parameters, such as the bandwidth and polynomial order. The choice of bandwidth and polynomial order, to a great extent, results in the under-smoothing or over-smoothing of the data. Thus, we propose a grid-based search algorithm for an appropriate choice of bandwidth and polynomial order by trading-off the bias and variance.

Consider a time series of position $(T, \mathcal{X})$ of a vehicle, where T is the vector of sampling time instances, and $\mathcal{X}$ is the matrix of vehicle position coordinates. We assume the underlying vehicle path is smooth and continuous, and can be represented using a regression function $f(t_0) = \mathbb{E}(\mathcal{X}|T = t_0)$ locally, and its derivatives $f'(t_0), f''(t_0), \dots, f^p(t_0)$ exist. Suppose that $(p+1)^{th}$ derivative of $f(t)$ at $t + t_0$ exists, we then approximate the unknown regression function $f(t)$ locally by a polynomial order p . A Taylor expansion for $f(t)$ in the neighborhood of t_0 is,

$$f(t) \approx f(t_0) + f'(t_0)(t - t_0) + \frac{f''(t_0)}{2!}(t - t_0)^2 + \dots + \frac{f^p(t_0)}{p!}(t - t_0)^p \quad (4.13)$$

This polynomial is fitted locally by a weighted least square regression problem that minimizes,

$$\min \sum_{i=1}^n \left[\mathcal{X}_i - \sum_{j=0}^p \beta_j (T_i - t_0)^j \right]^2 K_h(T_i - t_0) \quad (4.14)$$

where, h is the bandwidth controlling the size of the local neighborhood, and $K_h(\cdot/h)/h$ with K as a Kernel function assigning weight to each data point. We have considered a tri-cubic weight function (Eq.4.15) in our formulation due to its better local adaptive property required to capture the local path variations.

$$K(u) = \begin{cases} (1 - |u|^3)^3, & \text{if } |u| \leq 1. \\ 0, & \text{otherwise.} \end{cases} \quad (4.15)$$

Let $\hat{\beta}_j$ be the solution vector of the least square problem described in Eq.4.14. It is clear from the Taylor expansion in Eq.4.13 that $\hat{f}_p = p!\hat{\beta}_p$ is an estimator for $f^p(x_0)$, $p = 0, 1, \dots, P$. To estimate the entire function $f^{(p)}(\cdot)$ we solve the above weighted least square problem for all points x_0 in the domain of interest.

The design matrix of the problem described in Eq.4.14, denoted by Ψ , is :

$$\Psi = \begin{bmatrix} 1 & (T_1 - t_0) & \dots & (T_1 - t_0)^p \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & (T_n - t_0) & \dots & (T_n - t_0)^p \end{bmatrix} \quad (4.16)$$

and put

$$x = \begin{bmatrix} \mathcal{X}_1 \\ \mathcal{X}_2 \\ \cdot \\ \cdot \\ \cdot \\ \mathcal{X}_n \end{bmatrix} \quad \text{and} \quad \hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \cdot \\ \cdot \\ \cdot \\ \hat{\beta}_p \end{bmatrix}$$

Further, let $\mathbf{W}$ be the $n \times n$ diagonal matrix of weights:

$$\mathbf{W} = \text{diag} \{K_h(T_i - t_0)\} \quad (4.17)$$

Then the weighted least squares problem in Eq.4.14 can be written as,

$$\min_{\beta} (x - \Psi\beta)' \mathbf{W} (x - \Psi\beta) \quad (4.18)$$

with $\beta = (\beta_0, \dots, \beta_0)'$. The solution vector is provided by the weighted least square theory, and is given by

$$\hat{\beta} = (\Psi' \mathbf{W} \Psi)^{-1} \Psi' \mathbf{W} x \quad (4.19)$$

Here, as discussed before, the choice of bandwidth parameter h plays a crucial role. A too large bandwidth under-parameterize the regression function causing a large bias, while a too small bandwidth over-parameterize the unknown function and results in erroneous estimate. Ideally, the theoretical choice of bandwidth is easy to obtain, however, it is not directly usable since it depends on unknown quantities. In fact, finding a practical procedure for selecting the bandwidth parameter is one of the most important tasks [277].

Apart from bandwidth, the other unknown parameter for local polynomial fitting is the order of the polynomial function. Since the bandwidth primarily controls the modeling bias, the polynomial order is less crucial in kernel smoothing. However, for a given bandwidth h , a large value of p would expectedly reduce the modeling bias but would cause a large variance and a considerable computational cost. To reckon the optimal smoothing parameter, we propose a grid-search method, which would estimate the optimum bandwidth-polynomial-order combination that trade-off the bias and variance to get the desired smoothing. The following section describes the grid-search algorithm in detail.

4.2.3.1 Grid-Search Algorithm for Estimating the Optimal Kernel Parameters

While dealing with the bandwidth selection problem, a key issue is to have a good insight into the bias and variance of the estimators. Generally, the criteria for selecting the bandwidth is the trade-off between bias and variance, which nevertheless is a multi-objective optimization problem. For a locally weighted polynomial regression, the selection of optimal bandwidth and polynomial order can be defined as a derivative-free, unconstrained multi-objective optimization with a discrete domain of objective. It is discrete because the window size can take only an integer value, preferably odd. Though the polynomial could be of fractional order, it is not recommendable to use the entire real line. In view of the foregoing discussion, the estimation of optimal bandwidth and polynomial order can be defined as a grid-based, derivative-free, evolutionary multi-objective optimization problem. Generally speaking, an evolutionary multi-objective optimization problem would share more than one goal, but often conflicting. In our case, the minimization of variance leads to a larger bias and vice-versa. Without loss of generality, let us define a multi-objective optimization problem as follows;

$$\text{Minimize } F(\Phi) = (\Phi_1(\alpha), \Phi_2(\alpha), \dots, \Phi_m(\alpha)); \quad \text{Subject to: } \alpha \in \Omega \quad (4.20)$$

Where, α denotes a solution vector in the feasible solution space Ω , and Φ_i , ($i = 1, 2, \dots, m$) is the i^{th} objective to be minimized.

In our formulation, we defined the domain of optimization problem as a Cartesian grid made of bandwidth and polynomial order, as shown in Fig. 4.5. At the start, a sparse grid is defined with a bandwidth ranging from five to size of the data (N), with an equally spaced ten bandwidths. The range of polynomial order within the grid is set from two to five with a unit increment. We assume, based on the literature that the vehicle path can be represented using piece-wise polynomials of order that might vary, in a vague sense, between $[2, 5]$. At each grid-coordinate that intersects a bandwidth-polynomial order pair, the cost of the objective was defined using the Mean Squared Error (MSE), Absolute Mean Error (AME), and Error Variance (VAR). Initially, the cost is defined only using MSE and VAR. AME is used in the final stage of optimization to choose the optimal window size and polynomial order from the optimal grid. The proposed algorithm successively refined the grid by removing the grid-coordinates of higher costs. The grids of higher costs are identified by looking for the 75th percentile value of cost for the entire grid. The selection of 75th percentile cost was to ensure that the convergence and diversity are maintained during optimization. A too fast convergence may be possible with a lower percentile value of cost; however, it reduces diversity. At each stage of refinement, the grid spacing in the direction of bandwidth is reduced until the spacing between the grid is unity. The grid with unit spacing for bandwidth is referred to as Grid of Interest (GOI), from where the direction of grid refinement would be reversed, and the grid spacing is reduced to half in the direction of polynomial order, till the spacing is 0.25. This process would reinforce the selection pressure towards the optimal grid. Once we achieve the optimal grid, the optimal bandwidth-polynomial pair is identified for the minimum AME. The optimal bandwidth (h_{opt}) and polynomial order (p_{opt}), obtained after grid-search, is used in the minimization problem (Equation 4.18) to get the coefficients $\hat{\beta}$ of local polynomial function.

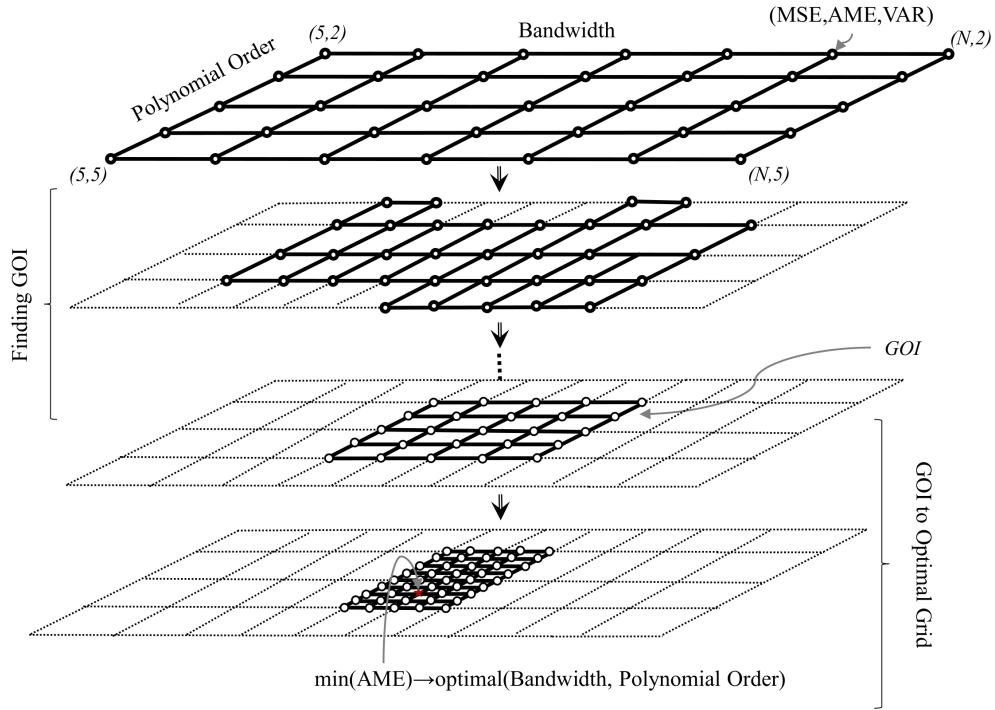


Figure 4.5: Graphical representation of the Grid-search algorithm

The three-stage vehicle path reconstruction technique described in this section can be applied to each vehicle data to achieve a smooth and continuously differentiable vehicle path that satisfies the conditions of internal and physical consistency. Next section presents the validation of the proposed reconstruction technique using the field data.

4.3 Performance Evaluation of the Proposed Reconstruction Technique

This section evaluates the reliability of the proposed reconstruction technique using an experimental and naturalistic vehicle path data. We also use the manually extracted NGSIM trajectory data reported by Coifman et al. [4], particularly to evaluate the performance of the missing data retrieval process.

4.3.1 Retrieval of Missing Data

As discussed earlier, the data loss during the sensor-based data collection is quite a common phenomenon, and the missing information needs to be retrieved efficiently for an effective microscopic traffic analysis. There are several effective interpolation techniques available to retrieve the missing information from time-series data. However, the main challenge in evaluating the performance of the retrieval process is that the actual data is unknown. To evaluate the efficacy of interpolation methods, we considered the manually extracted NGSIM (I-80) data reported by Coifman et al. [4]. We have synthesized the situation of data loss corresponding to occlusion by a fixed object

by removing certain observations at different sections along the roadway (Fig. 4.6a). The missing data were reconstructed, across 2283 vehicles using different interpolation methods such as linear, cubic, modified akima, pchip, and spline. Fig. 4.6b shows the comparison of different interpolation methods. It can be seen from the figure that the spline interpolation gives better results compared to the other interpolation techniques. It is also evident that as the size of occlusion increases, the prediction accuracy gets depreciated. To limit the error in the position within 5cm, it is advisable to select the data collection site so that the maximum possible occlusion is within 10m.

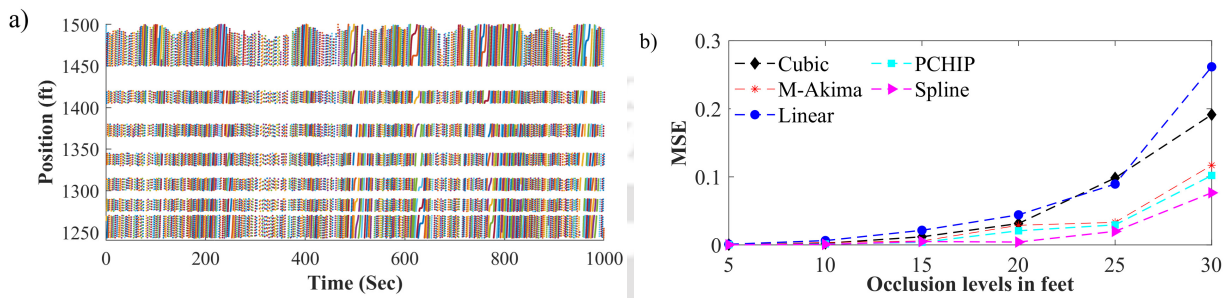


Figure 4.6: Performance evaluation of the missing data retrieval process; a) Manually extracted NGSIM data reported by Coifman et al. [4], with the synthesized occlusion conditions; b) Comparison of different interpolation techniques for the missing data retrieval

4.3.2 Reconstruction of Observed Vehicle Path

Potential of the proposed reconstruction method evaluated by performing a controlled experiment. A test vehicle equipped with a high-end GPS device (V-Box, Racelogic) (Fig. 3.5b) was driven at different speed levels on a testbed, where a ring road environment was set up (Fig. 3.5a). The speed of the instrumented vehicle was extracted from the V-Box file using RACELOGIC 2.13.5 (Build 211) 2003 software. At the same time, the test vehicle was video-graphed using a drone that has good gimbal support (DJI-Inspire 1). Further, the collected video footage was stabilized to a tripod-mode condition by transforming all video frames to the orientation of a reference frame using the Stabilizer tool. From the stabilized video footage, the vehicle path and the other derived quantities such as speed and acceleration were extracted using the trajectory extraction tool, SAVETRAX.

The observed vehicle path and the reconstructed path are shown in Fig. 4.7a. It is evident from the figure that the observed path carries unwanted noise as undulations (represented using the black line). The proposed reconstruction technique has significantly reduced the noise from observed data, moreover resulted in a smooth vehicle path, as shown in Fig. 4.7a (represented using red line). It is apparent that the smoothing technique has removed all the unwanted undulations in the observed vehicle path without significantly altering the observed vehicle path. Furthermore, we compared the speed obtained from V-Box (benchmark speed) and speed derived from the reconstructed path through numerical differentiation. Fig. 4.7b shows the comparison of speeds, and they are acceptably matching. The mean difference of the reconstructed and the benchmark speed was observed as 0.495 m/s with a standard deviation of 0.401 m/s . The comparison clearly shows that the proposed reconstruction framework is trustworthy, and can be directly used for path

smoothing and for an accurate derivation of kinematic quantities such as speed, acceleration, and jerk.

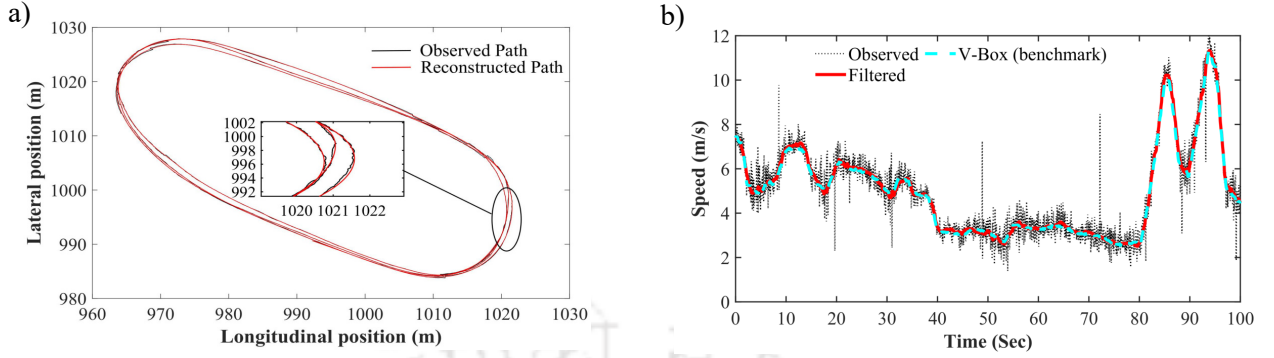


Figure 4.7: Performance evaluation of the proposed reconstruction technique using the data from a controlled experiment; a) Comparison of the reconstructed path and the observed path; b) Comparison of reconstructed speed with the benchmark speed

To elaborate on the performance at each stage of the proposed reconstruction method, results from each stage have been analyzed. Coming to the proposed RELP algorithm, the main emphasis was to get a kinematically consistent vehicle path. Fig. 4.8a shows the distribution of the acceleration after applying the RELP (the observed distribution is shown in Fig. 2.3a). It can be seen from the figure that the acceleration is not distributed within the realistic range, indicating the presence of noise. The Kernel smoothing is applied to the path to remove this noise using the estimated smoothing parameters. The acceleration distribution after Kernel smoothing is presented in Fig. 4.8b, which is well within the realistic range for acceleration. Figure Fig. 4.8c shows a comparison of two stages mentioned above and the benefit.

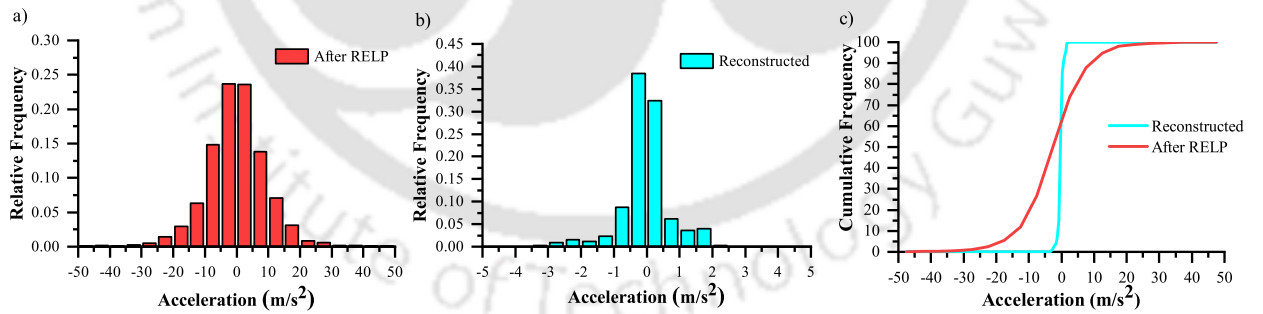


Figure 4.8: Physical consistency analysis of the estimated acceleration; a) Distribution of acceleration after RELP (ranges between $-50m/s^2$ to $50 m/s^2$); b) Distribution of acceleration after Kernel smoothing (ranges between $-3.5 m/s^2$ to $2.5 m/s^2$); c) Cumulative Frequency (in %) distribution after RELP and Kernel smoothing

It can be seen from Fig. 4.9a that RELP has efficiently denoised the speed data, though the error amplitude was non-uniform across the profile. Fig. 4.9b and c show the increasing benefit of the proposed method in speed and acceleration profile at consecutive stages of reconstruction.

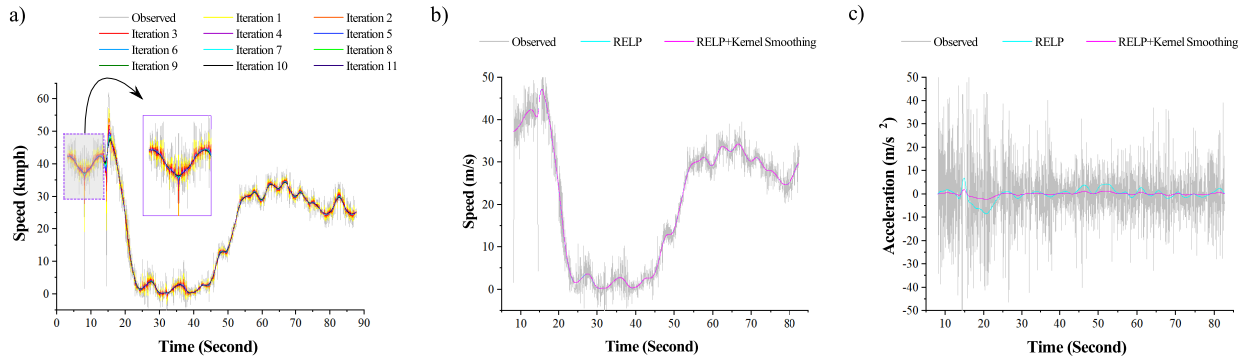


Figure 4.9: Increasing benefit of the proposed reconstruction technique, a) Performance of RELP in denoising the gradient of vehicle path b) Improvement in speed after each stage of reconstruction, c) Improvement in acceleration after each stage of reconstruction

Further, the performance of the proposed method was evaluated using the NGSIM trajectories, and compared with the reconstructed NGSIM trajectory reported by Montanino et al. [5]. For the comparison, we have selected the vehicle with ID 1929, which is commonly used in many of the trajectory reconstruction related studies.

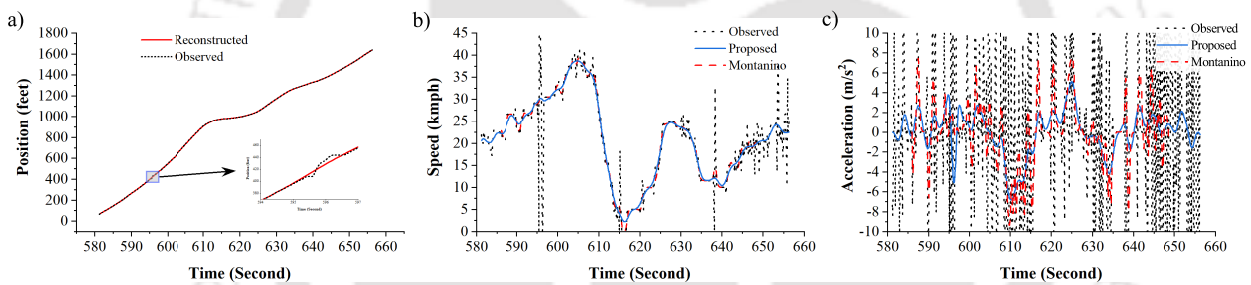


Figure 4.10: The performance of the proposed method on the NGSIM I-80 dataset presented on a sample trajectory (Vehicle ID 1929) and a comparison with the method proposed by Montanino et al. [5] (a) Trajectory, (b) Speed profile, and (c) acceleration profile.

It is evident from the Fig. 4.10 that the proposed method could reconstruct the vehicle trajectory, realistically. The studies using naturalistic driving data [279–281] have reported that the safety-critical events, such as the crashes or near crashes, could occur when the longitudinal acceleration is higher than the threshold of $-0.6g$ ($\approx 5.88m/s^2$). The acceleration profile corresponding to the raw data indicates a crash event, but the crash has not occurred in reality. It is to be noted that the acceleration profile obtained after applying the proposed method maintains the acceleration profile within the no-crash range.

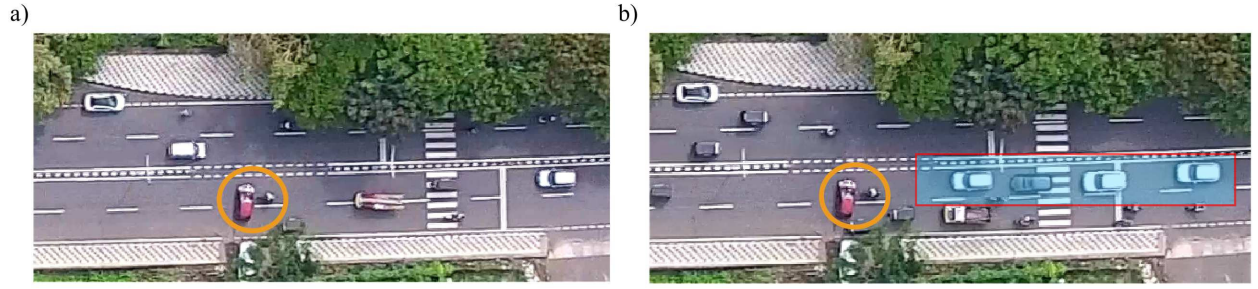


Figure 4.11: A 90 degree near-crash, a) Vehicles involved in the near-crash (circled in the figure), b) Piled-up vehicles due to the bottleneck created by the near-crash (highlighted in the figure)

We have also analyzed the trajectory corresponding to a vehicle that has undergone a near-crash situation. It was a 90-degree near-collision by a PTW (Fig. 4.11a) to a car that entered from a minor road. The incident also resulted in the piling up of vehicles behind (Fig. 4.11b). Fig. 4.12a shows the observed trajectory of PTW (PTW 1) and the piled up cars (shown only the first four of them), and their speed and acceleration profiles (Fig. 4.12 b and c, respectively). The reconstructed trajectories, speed, and acceleration profiles are shown in Fig. 4.12 c, d, and e, respectively. It is to be noted that PTW 1, the vehicle that has undergone near-collision, shows a deceleration rate of -5.07m/s^2 , which indicates the occurrence of an incident. However, the deceleration rate of the piled-up vehicles was well within the realistic range, which is expected as they have enough time for a smooth deceleration (Fig. 4.11). Hence, it is clear that the proposed method is not causing any over-smoothing, thus not misinterpreting the events, which is essential for safety-related traffic analyses.

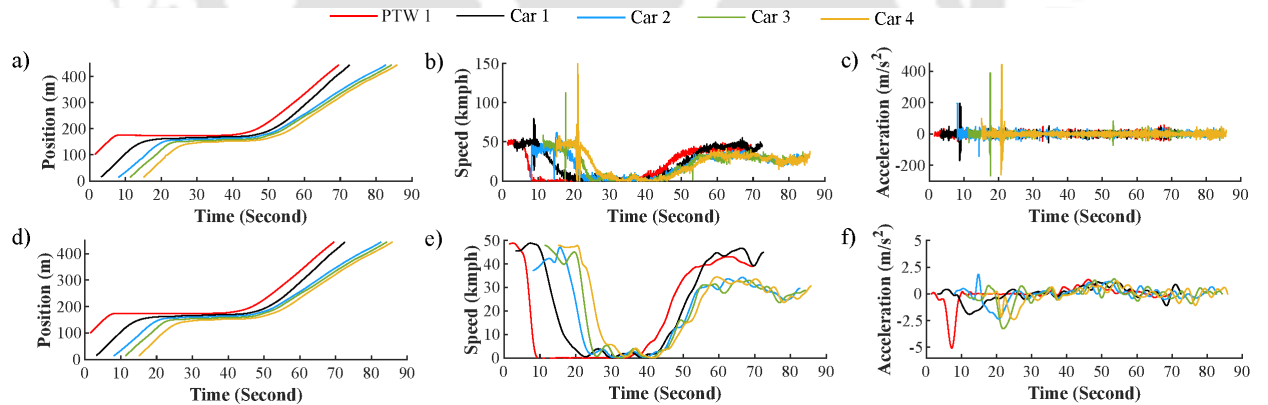


Figure 4.12: Position, speed, and acceleration profiles of the vehicle involved in near-crash (PTW 1) and the piled-up vehicles (Car 1, Car 2, Car 3, and Car 4), (a,b,c) Observed position, speed, and acceleration profiles, (d,e,f) Reconstructed position, speed, and acceleration profiles.

As discussed, the difficulty associated with the implementation of the advanced smoothing techniques compels the researchers to use off-the-shelf smoothing techniques. In order to avert such a situation, and for direct use of the proposed smoothing framework, we developed an application with a user-friendly interface. The application was developed and tested on MATLAB®2019b, in Windows 10® platform. The intended user can directly install the application into the MATLAB interface following the general procedure of app installation in MATLAB [282]. The observed

path data can be directly loaded into the interface, with predefined column headings, for smoothing. After smoothing, the smoothed path and the derived quantities such as speed and acceleration can be exported from the application. Fig. 4.13 shows the screen-shot of developed user-interface of the smoothing application.

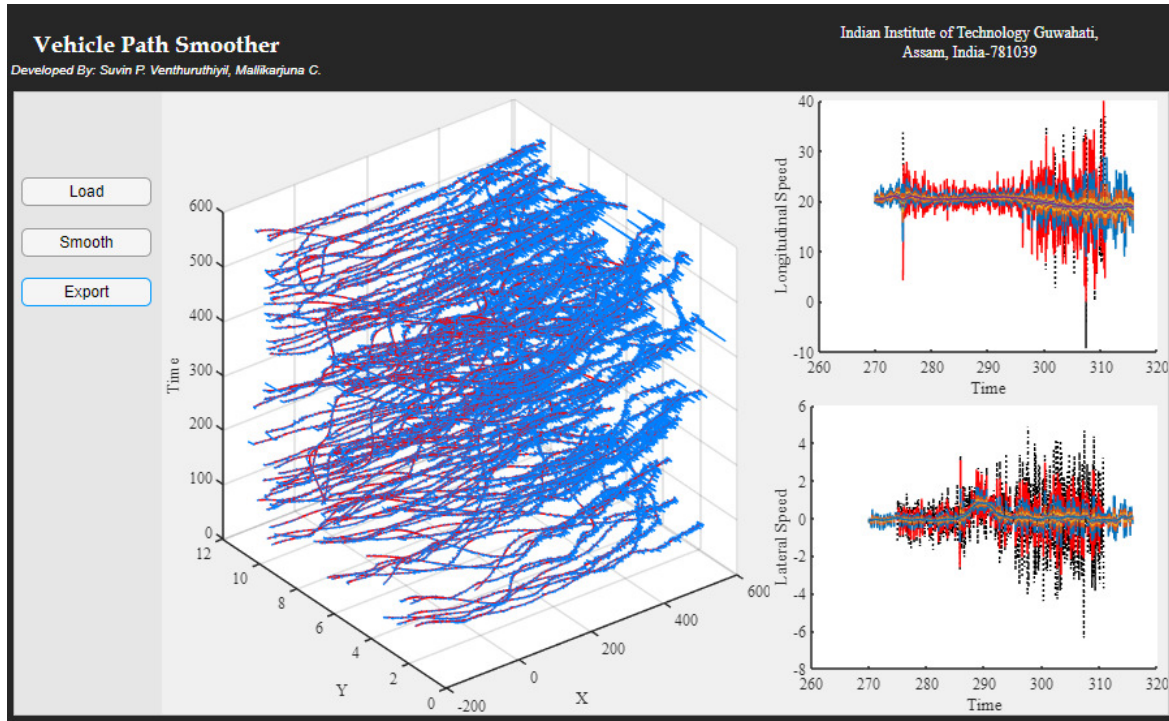


Figure 4.13: User interface of MATLAB application developed for vehicle path smoothing

4.3.3 Residual Analysis

In order to characterize the error generated during the drone-based traffic data collection, the residual between the observed path and the reconstructed path were analyzed. The residual was adopted as a proxy to the measurement error based on the assumption that the reconstructed trajectories are a better estimate of the actual than the observed data. Fig. 4.14 shows the distribution of error observed in the position, speed, and the acceleration in both the longitudinal and lateral directions. It is evident from the figure that the residuals are distributed as non-Gaussian, but has a zero mean. We verified that the residuals are distributed with zero mean through the t-test with a 99% confidence. The test results shown in Table 4.1 signifies that the residuals are distributed with a mean of zero for the position, speed, and acceleration in both the longitudinal and lateral directions. This indicates that the vehicle path data is free from systematic error, by virtue of the tracking accuracy achieved in the recent past as well as the advantage of camera stabilization.

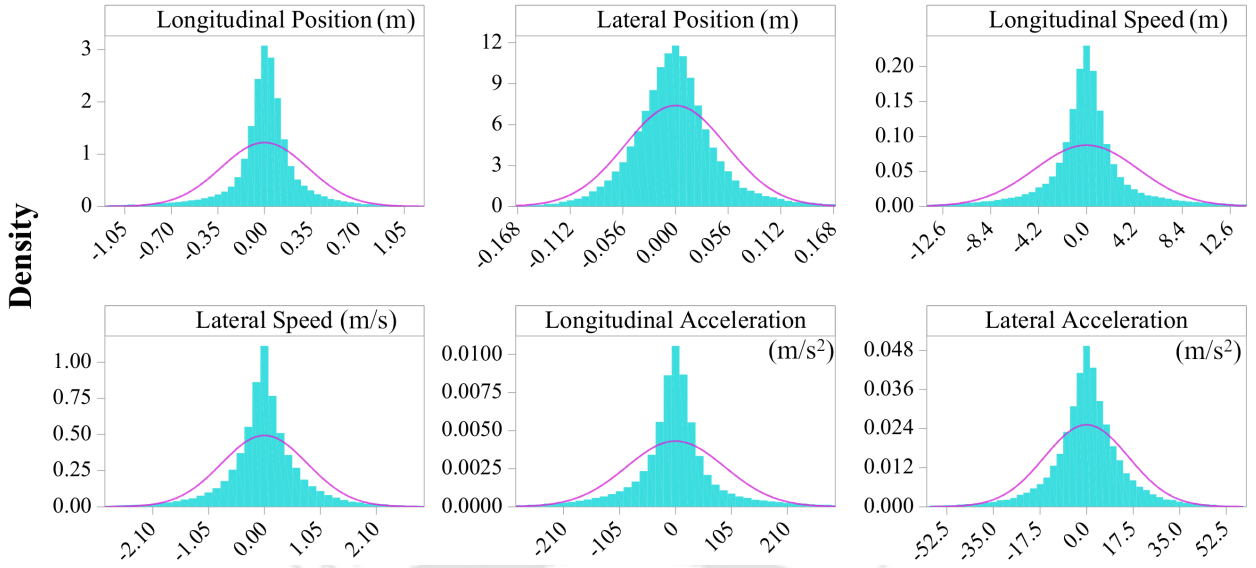


Figure 4.14: Distribution of residuals and a fitted Gaussian function

Though visually it is evident from Fig. 4.14 that the residuals are not Normally distributed, we have performed the Anderson-Darling Normality test to verify this statistically. The test statistics, along with the descriptive statistics of the residuals, are given in Table 4.2. The Normality test results at 95% confidence level show the non-Normality of residuals for the position, speed, and acceleration. The 95% tolerance interval of residuals in the discrete observation of longitudinal position (at every 0.0417 seconds) was found as $(-0.755 \text{ m}, 0.611 \text{ m})$, and for lateral position $(-0.091 \text{ m}, 0.114 \text{ m})$, which imply an error magnitude of significant effect. Further, an unacceptable amplification and propagation of the error can be seen in the derived quantities such as speed and acceleration. The 95% tolerance interval for longitudinal speed was $(-10.149 \text{ m/s}, 10.493 \text{ m/s})$ and $(-1.719 \text{ m/s}, 1.703 \text{ m/s})$ for lateral speed. The acceleration has shown an extremely outlying values of $(-201.225 \text{ m/s}^2, 220.214 \text{ m/s}^2)$ and $(-34.809 \text{ m/s}^2, 31.666 \text{ m/s}^2)$ for longitudinal and lateral acceleration, respectively.

One of our postulations about the characteristics of the error in the drone-based trajectory data was the heavy-tailed distribution. Kurtosis is generally considered as a measure of the "tailedness" of the probability distribution function of a real-valued random variable. The data sets with high Kurtosis tend to have heavy tails or outliers. For a Normal distributed random variable, the Kurtosis would be 3. The Kurtosis value obtained for the residuals, shown in Table 4.2, clearly indicates the heavy tailedness with a magnitude which is much higher than three.

Another claim from this study was about the difference in the level of smoothing required for different vehicles' paths due to the dissimilar noise contamination. We validated this claim by analyzing the time-series of mean position error per frame, which is presented in Fig. 4.15a. It is evident that the average error in the position per frame is not uniform across the length of the video, but highly random. Further, we verified that the residuals are not auto-correlated throughout the observation using the partial auto-correlation analysis of the mean error (Fig. 4.15). This clearly postulates that the vehicles entering the monitoring area at different observation times are

Table 4.1: Results of t-test to verify the zero mean of residuals

Null hypothesis	$H_0 : \mu = 0$	
Alternative hypothesis	$H_1 : \mu \neq 0$	
Sample	t-Value	P-Value
Longitudinal Position	0.00	>0.99
Lateral Position	0.00	>0.99
Longitudinal Speed	-0.00	>0.99
Lateral Speed	0.00	>0.99
Longitudinal Acceleration	-0.23	0.818
Lateral Acceleration	0.03	0.979

Table 4.2: Residual statistics and the Normality test results

Residual	Anderson-Darling Normality Test		Mean	Standard Deviation	Variance	Skewness	Kurtosis
	p-Value	A					
Longitudinal Position (m)	<0.005	5790.96	0.000	0.327	0.107	-1.446	18.262
Lateral Position (m)	<0.005	2105.82	0.000	0.054	0.003	0.133	88.082
Longitudinal Speed (m/s)	<0.005	4890.43	0.000	4.562	20.814	-0.003	8.718
Lateral Speed (m/s)	<0.005	3095.51	0.000	0.807	0.652	1.199	56.052
Longitudinal Acceleration (m/s)	<0.005	4422.54	-0.06	92.480	8551.9	0.530	7.749
Lateral Acceleration (m/s ²)	<0.005	2559.98	0.001	15.836	250.765	-0.414	22.895

contaminated with dissimilar noise levels; therefore, they need different levels of smoothing.

Further, we have analyzed the difference in the level of noise contamination in the longitudinal and lateral direction and the need for separate smoothing parameter estimation for these components. We validated the claim by testing error variance in the longitudinal and lateral directions to see whether they are statistically different. The test results are shown in Table 4.3. It is evident from the table that, at 5% significance, the variance of error in the longitudinal and lateral directions is significantly different. Thus the current practice of invariant smoothing of longitudinal and lateral dynamics would result in inappropriate smoothing of the data.

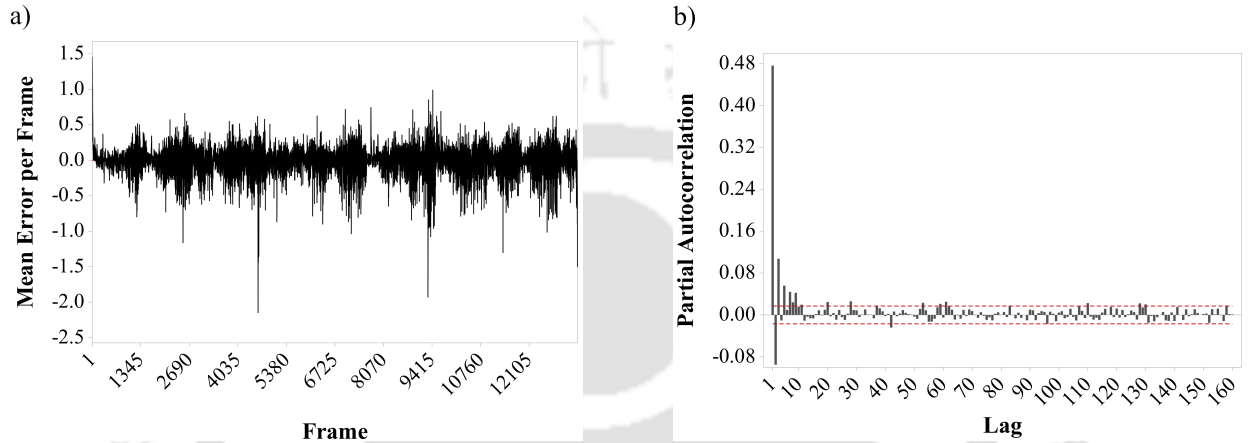


Figure 4.15: a) Time-series of mean error per frame; b) Partial auto-correlation analysis

The residual analyses clearly indicate the need of a novel reconstruction technique to deal with the noises generated during the contemporary traffic data collections. The present study fulfills this necessity by proposing the adaptive, and data driven smoothing technique. We further evaluated the reliability of the proposed reconstruction technique by verifying the internal and physical consistency of the reconstructed path.

4.3.4 Consistency Analyses

The reliability of the proposed reconstruction technique was evaluated through internal and physical consistency analyses. For this purpose, we considered two datasets; one was collected using a quadcopter drone (DJI-Phantom 4), from a rural highway covering 730m, for 15 minutes. The other data were collected using a fixed camera setup placed on a vantage point near an urban midblock section covering 100 meters and recorded for 30 minutes. The path data corresponding to all the vehicles were extracted from the videos using SAVETRAX, and applied the proposed smoothing technique. We have also considered some of the off-the-shelf smoothing techniques for comparative analysis. The internal consistency of the reconstructed path using different techniques was estimated using Equations 4.1a&b, respectively, for position and speed. Fig. 4.16a&b show the comparison of the internal consistency of position, and speed, respectively, using different reconstruction methods, for the drone data. Similarly, Fig. 4.16c&d show the internal consistency analysis results corresponding to the data from a fixed camera setup. It is evident from the

Table 4.3: Test for equal variance of residuals in longitudinal and lateral directions

Null hypothesis	H_0 : All variances are equal		
Alternative hypothesis	H_1 : At least one variance is different		
Significance level:	$\alpha = 0.05$		
<u>95% Bonferroni Confidence Intervals for Standard Deviations</u>			
Sample	N	Standard Deviation	Confidence Interval
Longitudinal Position	107863	0.326532	(0.321555, 0.331592)
Lateral Position	107863	0.053907	(0.052189, 0.055682)
<i>Individual confidence level = 97.5%</i>			
<u>Tests for Equal Variance</u>			
Method	Test Statistics		P-Value
Multiple comparisons	17735.4		0
Levene	33126.34		0

figure that the internal consistency is well-maintained for the trajectories reconstructed using the proposed method, compared to any other off-the-shelf methods. The maximum error in the longitudinal position was close to zero in the case of the proposed method, but the Moving Average and Gaussian filter have produced an error close to half a meter for the observed instance. This would be further amplified while estimating speed and acceleration through numerical differentiation. It is to be noted that, as presented in Section 2.3, the Moving Average technique is widely used for trajectory reconstruction. Furthermore, the Savitzky-Golay filter and LOESS methods have produced relatively less error that was less than 15cm. Similarly, the proposed method has outperformed all other methods (ref. Fig. 4.16b&d) in terms of the error in the longitudinal and lateral speed. The results indicate that the path data reconstructed using the proposed method closely follows the fundamental kinematic equations of motion.

In the physical consistency analysis, we have compared the acceleration profiles obtained after applying different smoothing techniques. The results of the physical consistency analyses are shown in Fig. 4.17. The consistency of longitudinal acceleration corresponding to different smoothing techniques are shown in Fig. 4.17a-f and m-r, respectively for drone data and fixed camera setup. Corresponding to the lateral acceleration, the distribution is shown in Fig. 4.17g-l and s-x, respectively for drone data and fixed camera setup. It is apparent from the figures that the proposed reconstruction technique maintains the vehicle accelerations within the realistic range of comfortable driving corresponding to uninterrupted traffic conditions observed on a rural highway. However, the acceleration distribution obtained after applying other methods (Fig. 4.17c-f) clearly indicate an unrealistic acceleration values. Similarly, for the lateral acceleration, the proposed method is giving realistic ranges for the study site. Lateral acceleration obtained after applying the other methods corresponds to aggressive driving, but such a behavior was hardly observed during

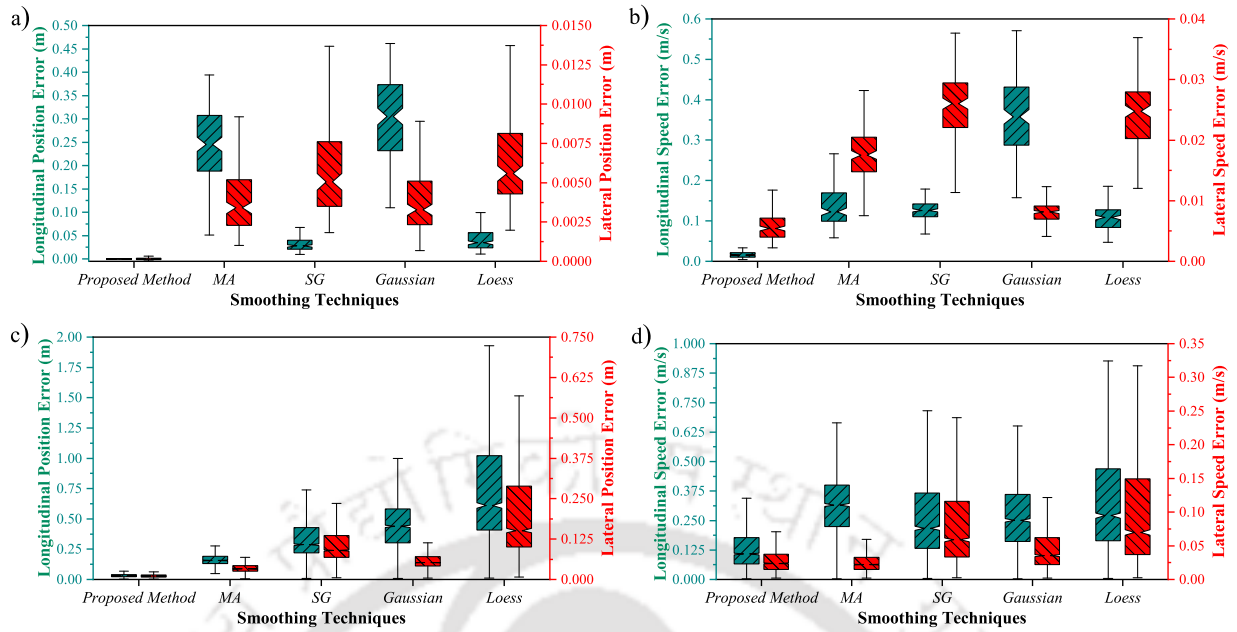


Figure 4.16: Internal consistency analysis of the proposed method and comparison with some of the off-the-shelf smoothing techniques

the data collection. The acceleration distribution observed from the fixed camera setup also gives satisfactory results for the proposed smoothing technique (Fig. 4.17h&t). The range of acceleration is relatively higher for data from fixed camera setup compared to drone, and is obvious since the data is from an urban environment.

From the internal and physical consistency analyses, it is evident that the proposed method outperforms most of the existing smoothing techniques. Moreover, the reconstructed path from the proposed method ensures the realistic driving behavior and strictly follows the physics of vehicle movement.

4.4 Summary

The present study proposes a trajectory reconstruction framework that is capable of handling a wide range of noises produced while collecting the trajectory data using the recent advances in drone technology and computer vision. The first and foremost merit of the proposed reconstruction technique is its adaptiveness contingent to the underlying noise characteristics. In this study, we have noticed that the noises in the observed vehicle path follow a heavy-tailed distribution, which often needs a special treatment. Consequently, the direct use of the existing smoothing techniques may not ensure sufficient smoothing of the observed data. We also found that the noise in the vehicle paths corresponding to different vehicles and directions shows unique characteristics, thus needing a separate smoothing for each vehicle and the longitudinal and lateral directions. However, the existing practices of vehicle trajectory reconstruction have used invariant smoothing parameters across different vehicle trajectories, and the lateral component of the path. Highlighting the need

for a separate smoothing, the present study proposed the trajectory reconstruction technique that is capable of handling the above challenges.

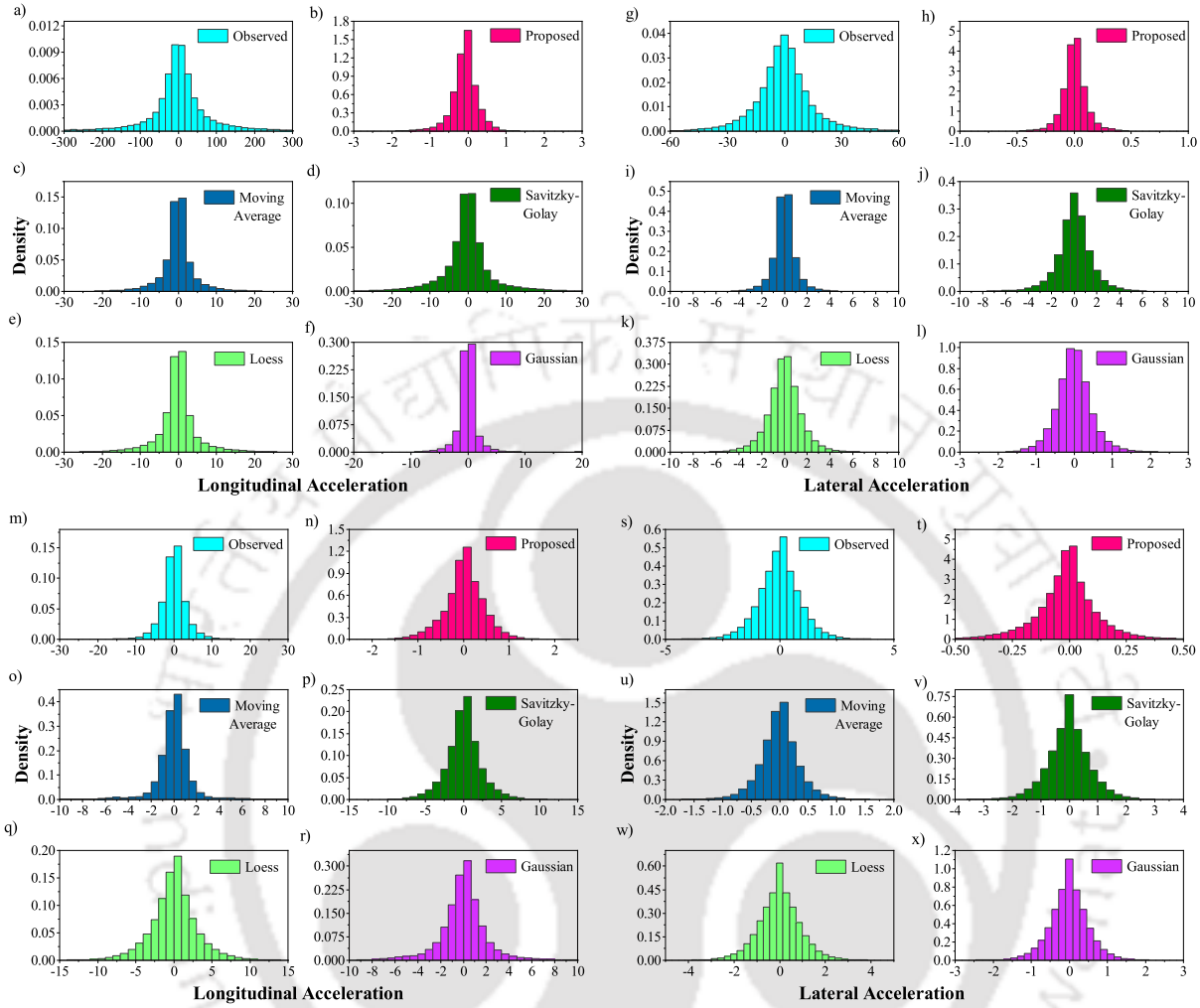


Figure 4.17: Physical consistency and Frequency spectrum analysis of acceleration

The proposed method comprises three stages that collectively perform the attenuation of the noises associated with the contemporary trajectory data collection. During the sensor-based data collection, it is expected to have a possible data-loss due to several factors, such as sensor malfunctioning or occlusion. In the first stage of our trajectory reconstruction method, all the missing information was retrieved using the spline interpolation technique, which has shown an accuracy up to $5cm$ when occluded for about $10m$. The selection of the interpolation technique was made after an extensive comparison of different interpolation techniques using a synthesized data-loss condition due to occlusion. In the second stage, we proposed a Recursively Ensembled Lowpass (RELP) filter, which is capable of handling the heavy-tailed noise distribution by recursively applying low-pass filtering to the path data, till the stopping criterion is achieved. RELP is built on the well-known Savitzky-Golay filter, which has the property that it is equivalent to a discrete convolution with a fixed impulse response. RELP was applied to the gradient of the path, to achieve a realistic and reasonable speed profile across the vehicle path. The main advantage of this stage,

apart from dealing with the heavy-tailed noise, is the ability to inherently ensure the internal consistency requirements of the path followed by the vehicles. However, this stage does not ensure complete denoising of the observed path data. Even after RELP, there could be certain white Gaussian noise remaining in the observed data. To cope up with this, we considered an adaptive tri-cubic kernel smoothing, where the smoothing parameters are estimated using a grid-search algorithm, which is also a proposal of the present study. In the grid-search algorithm, the smoothing parameters were tuned in such a way that they would trade-off the bias and variance of the estimate. After this stage, the vehicle path would be smooth and continuously differentiable and ensure the physical consistency of the vehicle path.

To validate the proposed method, we have performed a controlled experiment, where an instrumented vehicle was monitored using a drone. The error associated with the path data extracted from the drone video has shown the heavy tailedness, which is primarily due to the unwanted camera movements. The speed estimated after applying the proposed reconstruction method was compared with the speed obtained from V-Box device. It was found that both the speeds are matching with a mean difference of 0.495 m/s , with a standard deviation of 0.401 m/s . The analysis of the residuals has shown that the noise in the data has a zero-mean distribution, which can be attributed to the tracking accuracy achieved in the recent past. Moreover, the kurtosis of the noise distribution is significantly higher than three, indicating a heavy tailedness. The mean error in the position of vehicles at each frame was found to be statistically different and is not partially auto-correlated, which indicates that the vehicles entering the monitoring area at different times of observation contaminated with dissimilar levels of noise and need a different smoothing. The error variance of longitudinal and lateral positions of a vehicle was found statistically different, which indicates that the underlying errors are dissimilar.

Further, we have considered the field data collected using a drone as well as the fixed camera setup to ensure the reliability of the proposed method. The internal and physical consistency analyses have clearly shown that the proposed method is capable of achieving a better internal and physical consistency compared to any of the off-the-shelf smoothing techniques. Moreover, the application of the proposed method to the trajectory corresponding to a vehicle that has undergone a near-crash situation revealed that the proposed method is preserving the actual vehicle dynamics, which indicates the adequacy of the proposed method in safety-related studies. Accuracy levels required for the path data used in training and validating the path planning models for autonomous driving are stringent, and the proposed approach could cater such needs. Towards this direction, there are three major contributions; *i*) the proposal of a novel filtering technique, RELP, which would ascertain the underlying vehicle dynamics by inherently ensuring the internal consistency requirements of the reconstructed paths, *ii*) a method to identify an optimal set of smoothing parameters to reconstruct each vehicle path towards an optimal smoothing, *iii*) highlighted the need for separate smoothing of lateral and longitudinal components with an independent parameter estimation. For direct use of the proposed smoothing framework for the prospective users, we developed a MATLAB application and is publicly available.

Chapter 5

Anticipated Collision Time (ACT)

This chapter discusses a 2D surrogate safety indicator called Anticipated Collision Time (ACT) developed as part of this research. The detailed procedure of automatically detecting the conflict situations, extracting ACT profiles, and identifying the crash-types from the input trajectory data. This chapter also presents another safety indicator derived from the ACT profile called Time of Evasive Action (TEA). Other derived indicators, such as Time Exposed ACT (TE-ACT) and Time Integrated ACT (TI-ACT), which capture crash exposure and severity, are also described. Section 5.2 describes the derivation of ACT and a detailed methodology to estimate ACT. This section also discusses the other SSMs derived from an ACT profile, such as Time Extended ACT (TE-ACT), Time Integrated (TI-ACT), and TEA. Section 5.3 summarizes and concludes the chapter.

5.1 Introduction

Approximately 1.35 million people die annually from road traffic crashes, which is ranked as the eighth leading cause of global mortality by World Health Organization [2]. According to the 'Vision-Zero' initiative [283], "it can never be ethically acceptable that people are killed or seriously injured when moving within the road transport system". Taking this into account, many researchers and organizations are actively working to assess and mitigate the road fatality risk [24, 26, 161, 170, 172, 284]. All these attempts are envisioned to accomplish a sustainably safe road transport system with zero fatalities or no severe injuries [285–287]. The contributions of these researches are evident from the declining trend of the annual road crash rates across countries. Nevertheless, road fatalities are still a serious issue worldwide, particularly at complex traffic facilities and mixed traffic scenarios with a considerable proportion of Vulnerable Road Users (VRUs).

To date, however, there is no overarching and unifying framework available for a comprehensive safety analysis across various transportation facilities [159, 161]. The key challenge associated with developing such a measure is the 2D nature of the vehicle dynamics and the collision scenarios. Moreover, the observation and generalization of all the possible conflict scenarios from a trajectory database is a real challenge. To bridge the gap in the prevailing SSM-based safety as-

assessment methods, the present study proposes a novel surrogate safety indicator called Anticipated Collision Time (ACT). ACT can be applied to all scenarios irrespective of collision types, road geometry, and vehicle state configurations. For the vehicles that are on a collision course, the ACT estimation procedure considers the shortest distance between the vehicles and the closing-in rate in the shortest distance direction. In addition to providing a safety measure, ACT also offers information about the collision types such as head-on, rear-end, side-swipe, angled, and Run-Off road. Unlike the existing SSMs, the ACT estimation simultaneously considers all the vehicles interacting with the subject vehicle. Apart from ACT, this study proposes another SSM, derived from the ACT profile, and termed as Time of Evasive Action (TEA). This is the time at which the vehicle initiates an evasive action once found itself in an unsafe situation. To date, there is no SSM that can capture the drivers' response to a collision situation, which is vital while designing an effective collision warning system. TEA is such a measure that can help improve the design of existing collision warning systems and make them more customized to the user behavior.

5.2 The Anticipated Collision Time

In this study, we propose a new method for temporal conflict-proximity estimation, and the corresponding safety indicator is termed as Anticipated Collision Time (ACT). ACT is similar to TTC by definition, but functionally different. The shortest distance between the vehicles and the closing-in rate in the 'shortest distance direction' are the primary inputs to estimate ACT. ACT estimation considers the vehicles' closing-in rate to remain constant during the time left to collision. The closing-in rate amalgamates all the state variables that affect the likelihood of a potential conflict, thereby generalizing the ACT for extensive safety analysis.

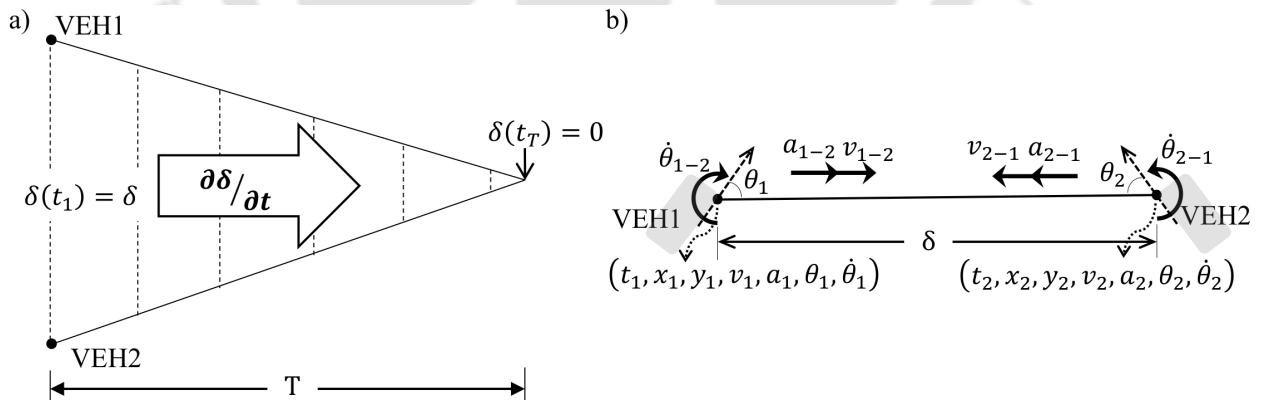


Figure 5.1: Simplified illustration of the concept of Anticipated Time to Collision (ACT); a) Closing-in of two vehicles till collision; b) Factors influencing the closing-in rate of two vehicles

The closing-in rate captures the rate at which vehicles approach each other. Fig. 5.1a illustrates a conflict scenario involving two vehicles. Let the shortest distance between the approaching vehicles at a time instance t_1 be δ , and the rate at which the vehicles approach each other (closing-in rate) is $\frac{\partial \delta}{\partial t}$. The vehicles would collide after T seconds if they retain the closing-in rate. In other words, after T seconds, the shortest distance between the vehicles would be zero. From this

interpretation, the closing-in rate of the vehicles can be estimated as:

$$\frac{\partial \delta}{\partial t} = \frac{\delta(t_1) - \delta(t_T)}{t_T - t_1} = \frac{\delta}{T} \quad (5.1)$$

Here, T is the Anticipated Collision Time (ACT) indicating the time at which a potential collision would occur. Rearranging Equation 5.1, we can write the equation for ACT as:

$$ACT = \frac{\delta}{\left(\frac{\partial \delta}{\partial t}\right)} \quad (5.2)$$

It is evident from Eq.5.2 that $\frac{\partial \delta}{\partial t}$ should be strictly greater than zero for a conflict to occur. A negative value of $\frac{\partial \delta}{\partial t}$ indicates that the vehicles are shying-away. Therefore, more precisely, the equation for ACT can be written as;

$$ACT = \begin{cases} \frac{\delta}{\left(\frac{\partial \delta}{\partial t}\right)}, & \text{if } \frac{\partial \delta}{\partial t} > 0 \\ \infty, & \text{otherwise} \end{cases} \quad (5.3)$$

The factors determining the closing-in rate of two vehicles are their speed, acceleration, heading angle, and yaw rate. Therefore, the closing-in rate of a pair of vehicles can be determined by taking the resultant of speed, acceleration, and steering in the direction of the shortest distance between the vehicles. Fig. 5.1b schematically represents the components that decide the closing-in rate of a pair of vehicles. The closing-in rate between these vehicles would be the summation of the resultant speed, acceleration, and yaw rate. Mathematically, it can be described as:

$$\frac{\partial \delta}{\partial t} = \mathcal{Rel}(\vec{v}_{1-2}, \vec{v}_{2-1}) + \mathcal{Rel}(\vec{a}_{1-2}, \vec{a}_{2-1}) \times t + \mathcal{Rel}(\dot{\theta}_1, \dot{\theta}_2) \times \delta \quad (5.4)$$

Where, $\vec{v}_{1-2}$, $\vec{a}_{1-2}$ are the components of speed and acceleration of vehicle-1 towards vehicle-2 in the direction of the shortest distance. $\vec{v}_{2-1}$, $\vec{a}_{2-1}$ are the components of speed and acceleration of vehicle-2 towards vehicle-1 in the direction of the shortest distance. $\dot{\theta}_1, \dot{\theta}_2$ are the steering rates of vehicle-1 and vehicle-2, respectively. Here, $\mathcal{Rel}$ is an operator that takes the vector sum of the quantities.

Broadly, the procedure for estimating ACT can be described as follows;

1. Get the state parameters of the vehicle pairs for which ACT is to be estimated.
2. Orient the vehicles based on their heading angles.
3. Estimate the shortest distance δ , between the vehicles.
4. Get the components of the vehicles' state vector in the direction of shortest distance (δ) and

estimate the rate at which vehicles are closing-in ($\frac{\partial \delta}{\partial t}$).

5. Estimate *ACT* inputting δ and ($\frac{\partial \delta}{\partial t}$) in Eq.5.3

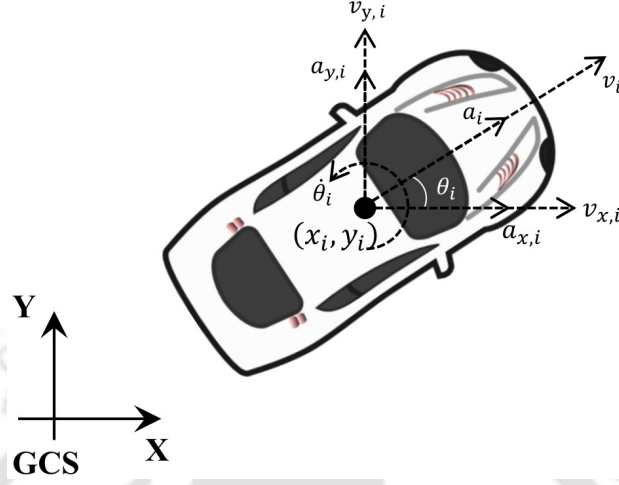


Figure 5.2: The state of a vehicle i at a time instance t in GCS

5.2.1 Procedure for Orienting Vehicles, Distinguishing the Vehicles on a Collision Course, and Identifying the Collision Type

The input for estimating the ACT is the vehicle state data $\mathcal{X} = [t, x, y, v_x, v_y, a_x, a_y, \theta, \dot{\theta}]$, where t is the time, x is the longitudinal position, y is the lateral position, v_x and v_y are the longitudinal and lateral speeds, a_x and a_y are the longitudinal and lateral accelerations, θ is the heading angle, and $\dot{\theta}$ is the yaw rate. It should be noted that the vehicle's state is measured with reference to a Global Coordinate System (GCS). Fig. 5.2 represents the state of a vehicle i at a particular time instance t .

If the heading angle information is not available, which is the case with the existing trajectory data, it can be estimated using the available state information. The heading angle $\theta(t)$ of a vehicle at a time instance t can be estimated using the following equation,

$$\theta(t) = \tan^{-1} \left(\frac{v_y(t)}{v_x(t)} \right) \quad (5.5)$$

5.2.1.1 Orienting the Vehicle based on the Heading Angle

This section presents a procedure to orient the vehicle in such a way that it makes an angle (θ) with reference to GCS. The following rotation matrix (R_z) (Eq.5.6) can be used to achieve a yaw

of (θ) with GCS:

$$\mathbf{R}_z = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (5.6)$$

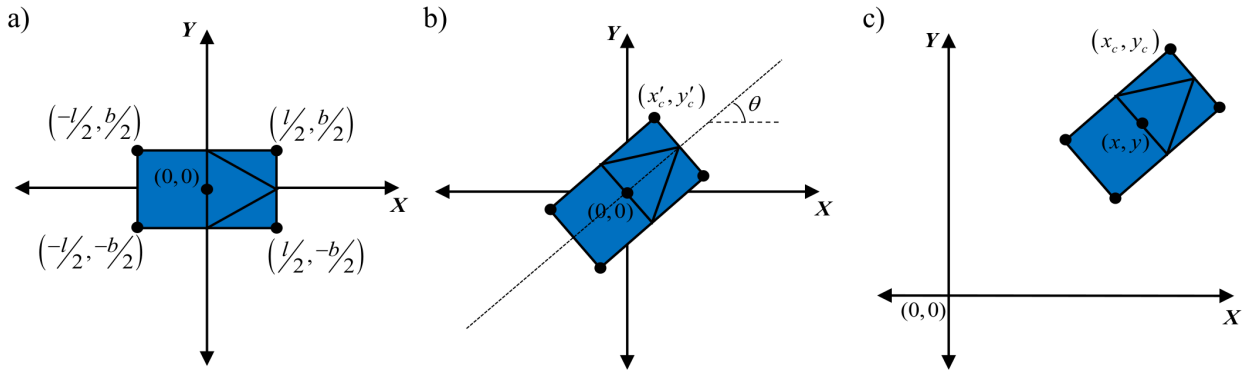


Figure 5.3: Procedure for orienting the vehicles in the 2D space according to the heading angle; a) Base orientation; b) Rotation; c) Translation

Initially, it was assumed that the vehicle is oriented, as shown in Fig. 5.3a, with its centroid as the origin and heading angle aligned with the X-direction of GCS. The corner coordinates of the vehicle can easily be estimated based on the known vehicle dimension (length l , and width b). The vehicle was rotated from the basic orientation to actual heading angle using the rotation matrix $\mathbf{R}_z$ (Fig. 5.3b). After this step, the corner points of the vehicle are super-positioned to match with its actual position (based on any of the known point, say the centroid (x, y)), as shown in Fig. 5.3c. For example, the actual corner points (x_c, y_c) (Fig. 5.3b) of the vehicle can be estimated from (x'_c, y'_c) (Fig. 5.3c) as,

$$x_c = x'_c + x$$

$$y_c = y'_c + y$$

5.2.1.2 Estimating the Shortest Distance Between the Vehicles

Once the vehicles are oriented with respect to their heading angle, it is essential to know the shortest distance between the vehicles. Fig. 5.4b presents the logic used to estimate the shortest distance. First, the nearest corners of the vehicles on a collision course were identified by checking the Euclidean second norm of all the corner point pairs, using the following equation:

$$\begin{bmatrix} x_{c1} & y_{c1} \\ x_{c2} & y_{c2} \end{bmatrix} = \min (\|\mathbf{P}_1 - \mathbf{P}_2\|_2) \quad (5.7)$$

where, (x_{c_i}, y_{c_i}) , $i = 1, 2$ are the closest points between the subject and the interacting vehicles, $\mathbf{P}$ is the vector of boundary points of vehicle i . It is to be noted that, to estimate the shortest distance, the collision points must be known. The collision point can be estimated from the known closest corner points using the following procedure.

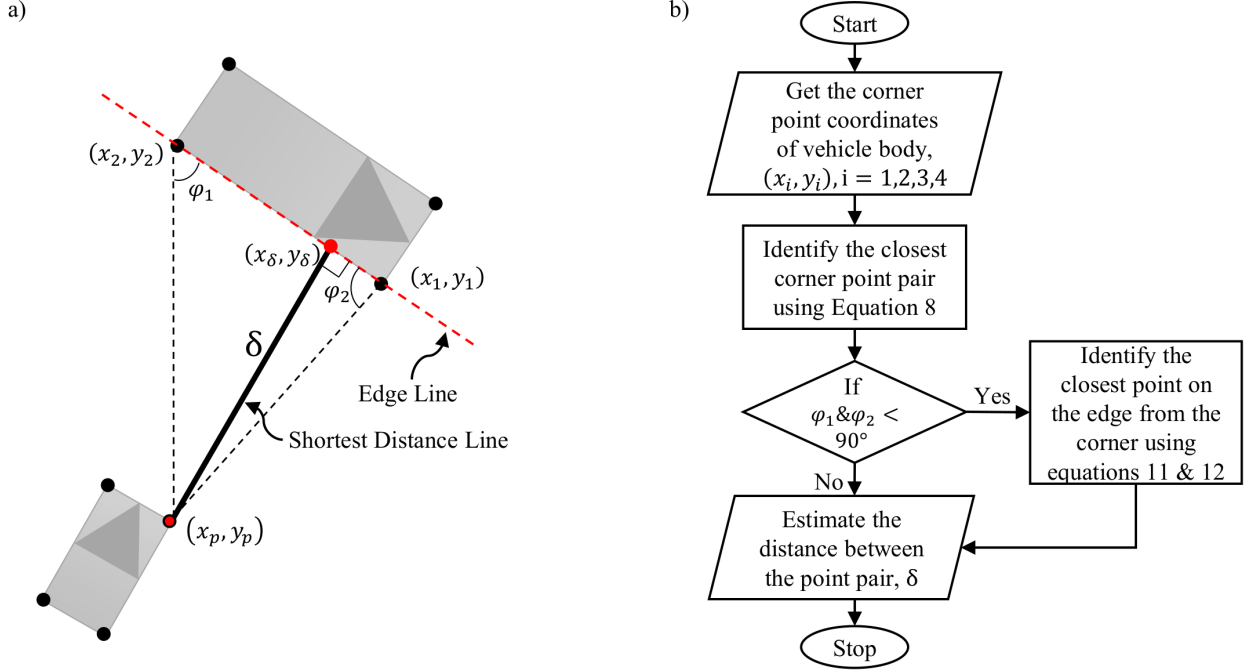


Figure 5.4: Estimation of shortest distance between vehicles; a) a general orientation of vehicles; b) Algorithm flowchart of closest point estimation

For this purpose, first let's assume a general case of the conflicting vehicles' orientation as shown in Fig. 5.4a. Let (x_p, y_p) be the nearest corner points of the subject vehicle from the conflicting vehicle. Let (x_1, y_1) and (x_2, y_2) be the corner points constituting the nearest edge of the vehicle to (x_p, y_p) . In order to estimate the shortest distance, it is essential to know (x_δ, y_δ) . Then the shortest the distance δ is,

$$\delta = \sqrt{(x_p - x_\delta)^2 + (y_p - y_\delta)^2} \quad (5.8)$$

As (x_δ, y_δ) is the right angle intersection of the vehicles' edge line and the shortest distance line, it can be estimated by solving the equations of both the lines simultaneously. Considering the known end points, the equation of edge line (in the parametric form, $y = \alpha \times x + \beta$. where, α is the slope, and β is the intercept) is,

$$y = \alpha \times x + \beta_1 \quad (5.9)$$

Where, $\alpha = \left(\frac{y_2 - y_1}{x_2 - x_1} \right)$ and $\beta_1 = y_1 - \alpha_1 \times x_1$

Similarly, the equation of the shortest distance line would be,

$$y = \frac{-1}{\alpha} \times x + \beta_2 \quad (5.10)$$

Where, $\beta_2 = y_p + \frac{x_p}{\alpha}$.

It is to be noted that the slope of the shortest distance line would be the negated inverse of the slope of edge line since they are intersecting at right angles. Solving Eq.5.9 and Eq.5.10 for intersection point yields (x_δ, y_δ) :

$$x_\delta = \frac{\beta_2 - \beta_1}{\alpha + \frac{1}{\alpha}} \quad (5.11)$$

$$y_\delta = \frac{\alpha \times (\beta_2 - \beta_1)}{\alpha + \frac{1}{\alpha}} \quad (5.12)$$

Then the shortest distance between the vehicles can be estimated by substituting Eq.5.11 and Eq.5.12 into using Eq.5.8.

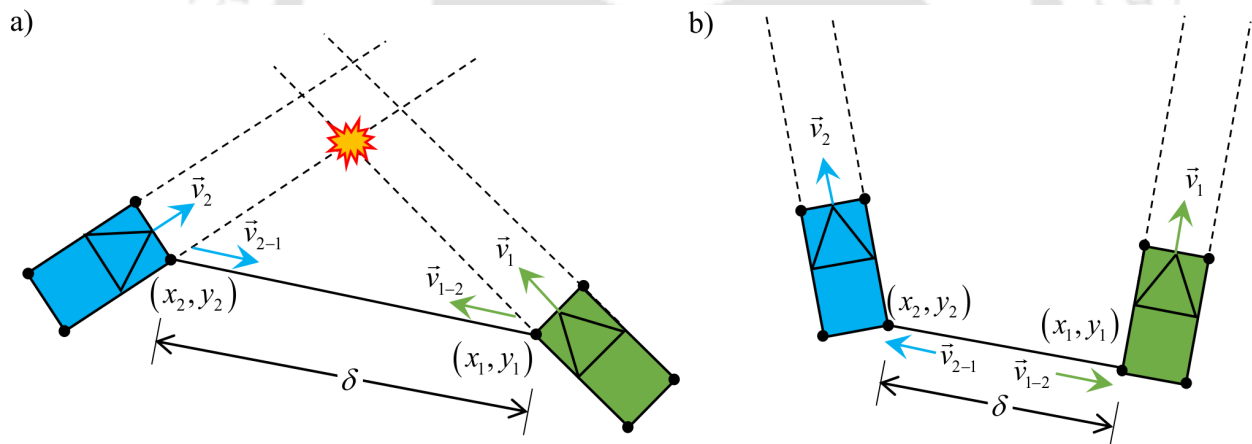


Figure 5.5: State configuration of vehicles involved in angled collision scenario; a) Vehicles are closing-in; b) Vehicles are shying away

5.2.1.3 Checking whether the Vehicles are on a Collision Course

In the basic TTC formulation, the collision course was identified based on the speeds of vehicles in car-following. Later, Ozbay et. al., [288] have expanded the condition for a collision course and tabulated the possible scenarios between the car-following vehicles to involve in a collision course. Fundamentally, these criteria indicate whether the vehicles are either closing-in or shying-away. For vehicles to be on a collision course, they have to close-in on each other. The vehicle pair shown in Fig. 5.5a corresponds to a closing-in movement, where there is a potential of future collision. In contrast, the vehicle pair in Fig. 5.5b shies-away as they move forward, thus no chance of collision.

Therefore, the necessary condition for vehicles to be on a collision course is a positive closing-in rate of vehicles. The sufficient condition is that the vehicles' movement must intersect.

In the case of rear-end collisions, the intersecting vehicles are generally identified by checking the lateral overlapping of the vehicle bodies [284], assuming that the vehicles are aligned in such a way that their centerlines are parallel (as shown in Fig. 5.6a). The criterion for the overlapping of vehicles (g_y) is identified using the following equation:

$$g_y = |y_l - y_f| - \frac{w_l}{2} - \frac{w_f}{2} \quad (5.13)$$

where (x_l, y_l) and (x_f, y_f) are the front center coordinates of the leader and follower vehicles, respectively. w_l and w_f are the width of the leader and follower vehicles, respectively. If $g_y < 0$, vehicle bodies are overlapping laterally, and if they have a positive closing-in rate, they are on a collision course. However, in the real field, the rear-end collision can happen even though there is no overlap between the vehicles in the lateral direction (see Fig. 5.6b). Such a scenario could be observed during a lane change operation, overtaking, merging, and so on. Similarly, there can be scenarios such as no rear-end collision, even though there is a lateral overlap between the vehicles (see Fig. 5.6c). Here, the distinguishing factor is the vehicles' orientation. Therefore, the orientation of vehicles needs to be appropriately considered for collision risk estimation.

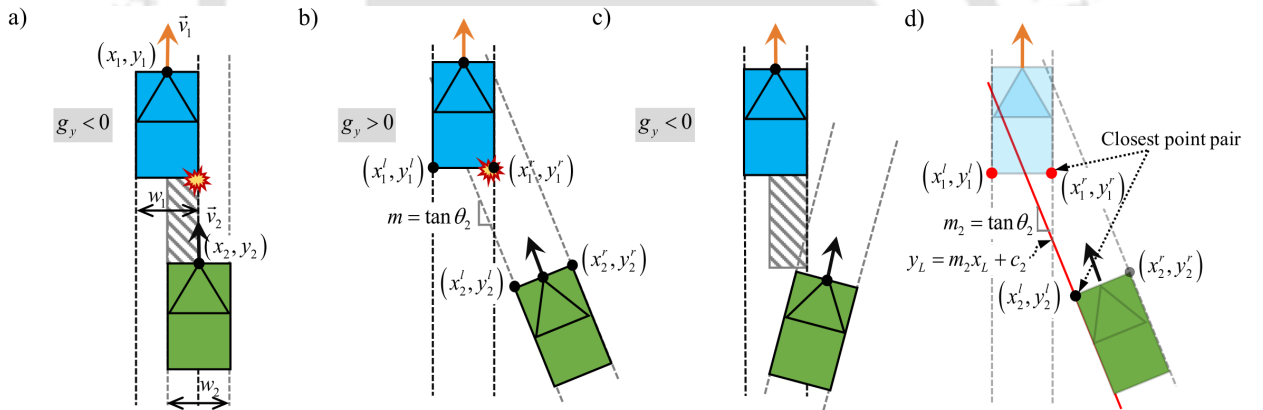


Figure 5.6: State configuration of vehicles in rear-end collision scenario; a) Conflict scenario of overlapping vehicles with parallel center-line; b) Conflict scenario for non-overlapping vehicles with angled center-line; c) Non-conflict condition for overlapping vehicles with angled center-line; d) Criteria for rear-end conflict

To generalize the sufficient condition for vehicles to be on a collision course, we introduced the following procedure. A collision happens only when the line drawn from the approaching vehicle's closest corner crosses through either of the other vehicles' corner points (Fig. 5.6d). The condition for a line to pass between a pair of points is stated as follows: A line L passes between two points (x_1^l, y_1^l) and (x_2^r, y_2^r) iff the y-offset of the points from the line have opposite sign. Mathematically, it can be proved as follows. Let the equation of the line L be:

$$y_L = m_2 x_L + c_2$$

If the line passes through any of the corner points, then:

$$y_1^l - m_2 x_1^l + c_2 = 0, \quad \text{or} \quad (5.14a)$$

$$y_2^r - m_2 x_2^r + c_2 = 0 \quad (5.14b)$$

If the line is not passing through a point, the two possibilities are either the point is above the line or below the line. Manifestly, when the pair of points lie on both sides of a line, the y-offset of the points from the line would have opposite sign. Therefore, the condition for the line to pass between the points can be formulated as:

$$\text{sign}(y_1^l - m_2 x_1^l + c_2) \times \text{sign}(y_2^r - m_2 x_2^r + c_2) \leq 0 \quad (5.15)$$

Therefore, the vehicle pair on a collision course can be identified considering the criteria given in Equation 5.15 and a positive closing-in rate.

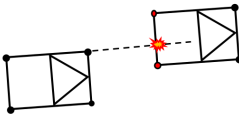
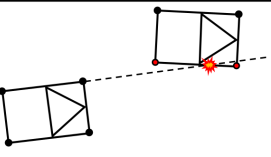
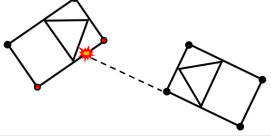
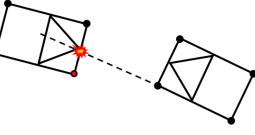
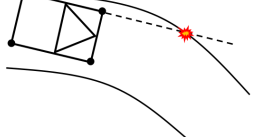
5.2.1.4 Determination of the Collision Type

The ability to differentiate various crash types is one of the unique attributes of ACT. Such an additional information on crash types enables the researchers to perform a comprehensive behavior modeling of different crash types. Besides, gaining more knowledge about different collision types allows for developing a more effective collision warning systems, incorporating all forms of collision possibilities into the existing rear-end collision warning systems. We found that the position of the impact point on the vehicle-body and the nature of vehicles' relative motion are the traits that determine the collision type. The condition used for the intersecting movement identification, presented in Section 5.2.1.3, could be used to identify the conflict point. Table 5.1 lists various collision types and the corresponding combinations of the impact points and relative movement characteristics. This decision table helps a straightforward classification of collision types. A line from the nearest corner, with a slope that equals its heading angle, is projected onto the interacting vehicles' body and identifies the edge through which it passes (using Eq. 5.15). Upon identifying the impact point, the information on vehicles' relative motion was used to get the crash type from the decision table. For example, in case of a rear-end collision, it is essential that the extended line from the follower vehicle to pass through the leader vehicle's rear-end corner points. It is to be noted that the vehicles should have the same direction of motion for rear-end and side-swipe collisions. Whereas, for angled and head-on collisions, the vehicles should approach each other. A run-off-road collision is a single-vehicle crash where the vehicle crosses the road boundary.

5.2.2 Other Crash Surrogates Derived from ACT-Profile

The ACT profile of a vehicle can be used to derive several other safety indicators. This section presents three other surrogate safety measures derived from ACT, namely the Time of Evasive

Table 5.1: Criteria for classification of conflict types

Collision Type	Collision Configuration	Relative Motion*
Rear-End		
Side-Swipe		
Angled		
Head-On		
Run-Off Road		

* Size of the arrow represent the magnitude of the vector

Action (TEA), Time Exposed ACT (TE-ACT), and Time-Integrated ACT (TI-ACT).

5.2.2.1 Time of Evasive Action (TEA)

Time of Evasive Action (TEA) is defined the time at which the vehicle starts to respond, in terms of an evasive action, to a perceived potential conflict situation ahead. TEA is one of the new SSMs of its kind, which measures the drivers response pattern during an unsafe situation. Some of the main applications of TEA are, but not limited to, (i) user specific customization of autonomous vehicle control algorithm; (ii) designing collision warning system with sufficient buffer time to respond; (iii) crash specific and vehicle type specific analyses of response pattern to enhance the safety at various road facilities. Fig. 5.7 shows the TEA estimation procedure using the ACT and acceleration profiles. First, the time instance at which the vehicle enters into an unsafe situation (t_{unsafe}) is identified from the ACT profile considering an ACT threshold (ACT^*). If the crash frequency data are available, ACT^* can be estimated by mapping the estimated crash frequency to the actual. In the absence of crash data, it is convenient to assume ACT^* as 3sec as suggested by several past studies [289, 290]. The collision would happen ACT^* seconds later from the time instance at which the vehicle enters into an unsafe situation (t_{unsafe}). Once the t_{unsafe} is estimated, a look-back procedure can be performed from the instance t_{unsafe} to a prior time instance where

the vehicle started to decelerate (Fig. 5.7b). This is nothing but the time at which the vehicle started taking an evasive action, in terms of deceleration ($t_{deceleration}$). Then, TEA can be estimated as the summation of ACT^* and the average absolute difference of t_{unsafe} and $t_{deceleration}$. i.e.,

$$TEA_i = \frac{1}{n} \sum_{k=1}^n |t_{unsafe}^k - t_{deceleration}^k| + ACT^* \quad (5.16)$$

It is noteworthy that the TEA measure gives an idea about, on an average, how early a vehicle/driver is responding to a potential collision situation. Estimating TEA for different drivers, vehicle classes, and facilities would enable an effective design of warning mechanism to alert the users ahead of an unsafe situation, with a sufficient buffer time to respond.

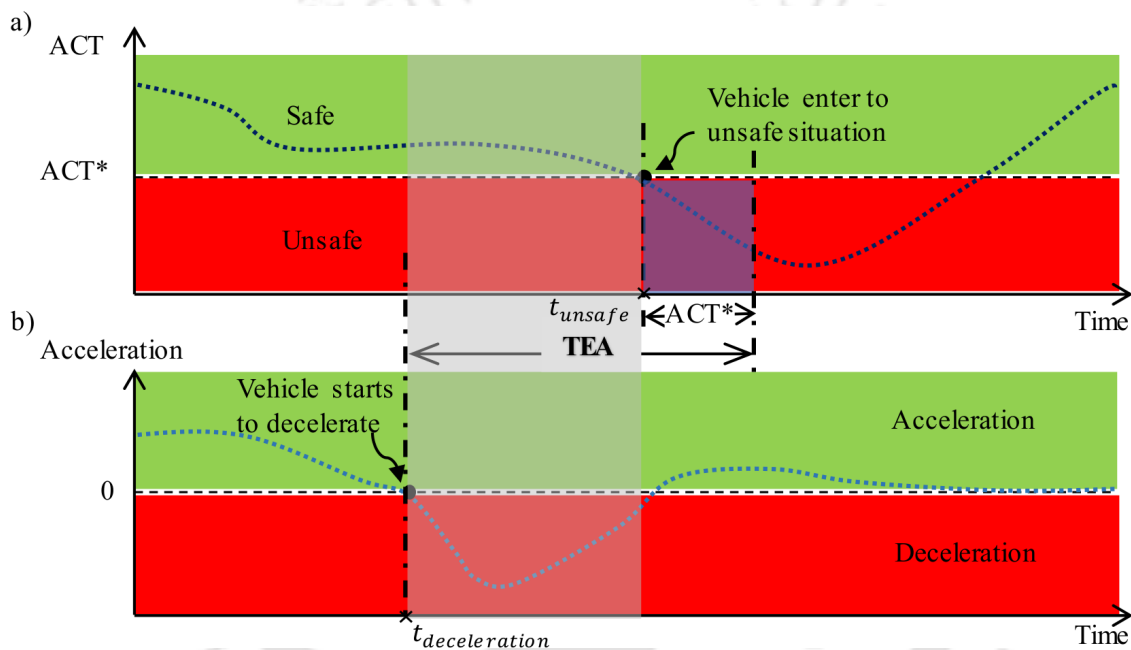


Figure 5.7: Estimation of TEA from ACT and acceleration profiles; a) ACT profile with safe and unsafe regions separated by ACT^* ; b) Acceleration profile showing acceleration and deceleration regions

5.2.2.2 Time-Exposed and Time-Integrated ACT

Using ACT profile, two other safety indicators termed as Time-Exposed ACT (TE-ACT) and Time-Integrated ACT (TI-ACT) were also derived. TE-ACT measures the duration for which the vehicle is exposed to an unsafe scenario. Whereas, the TI-ACT measures the severity of the potential collision. TE-ACT and TI-ACT are similar to Time-Extended TTC (TET) and Time-Integrated TTC (TIT), but TET and TIT consider only rear-end collision risk, whereas TE-ACT and TI-ACT consider all collision scenarios encountered while driving. Fig. 5.8 shows an ACT profile divided into two parts: safe and unsafe region, demarcated by ACT^* . The ACT profile segments falling in the unsafe zone (highlighted with yellow color in Fig. 5.8) determines the risk exposure and severity of a potential crash. Once all the ACT segments falling in the unsafe region are abstracted and combined, the resulting total time gives the duration of crash exposure (TE-ACT), and the

resulting total area of the yellow regions represents the crash severity (TI-ACT). Mathematically they can be expressed as follows:

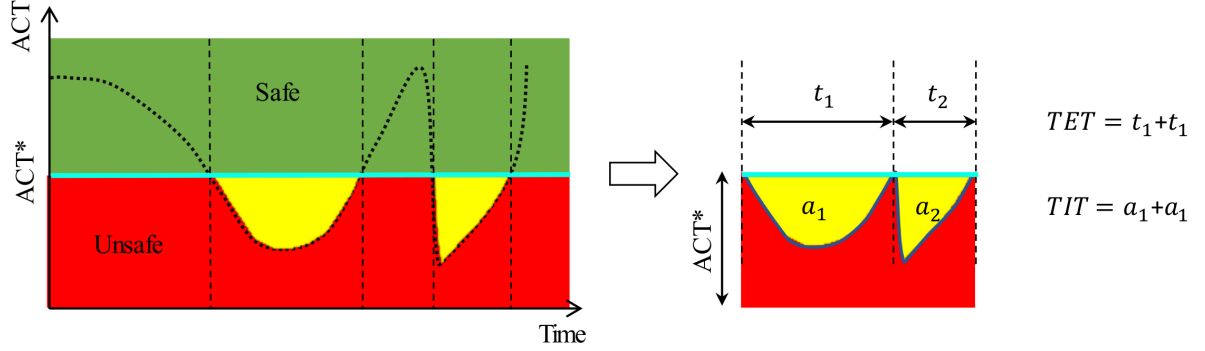


Figure 5.8: Estimation of TE-ACT and TI-ACT from ACT profile

$$TEACT_i = \sum_{t=0}^T \tau_i(t) \times \Delta t \quad (5.17a)$$

$$\tau_i(t) = \begin{cases} 1, & \text{if } \forall 0 \leq ACT_i(t) \leq ACT^* \\ 0, & \text{Otherwise} \end{cases} \quad (5.17b)$$

$$TIACT_i = \int_0^T (ACT^* - ACT_i(t)) dt, \quad \forall 0 \leq ACT_i(t) \leq ACT^* \quad (5.18)$$

5.2.3 Comparison of ACT with TTC

Many current surrogate safety analysis techniques and the majority of collision warning systems are based on the TTC measure. To emphasize the benefits of ACT in safety assessment studies, we performed a comparative assessment of ACT with TTC. Fig. 5.9 shows the distribution of duration of crash exposure and and severity derived from ACT and TTC. It is evident from the figure that the quantification of crash exposure and crash severity using the conventional TTC measure underestimates the risk levels. Table 6.7 shows the descriptive statistics of the exposure and severity obtained from ACT and TTC. The average duration of crash exposure and crash severity values obtained from ACT and TTC are significantly different. It can be seen from the table that the average duration of crash exposure estimated from TTC (TET) and ACT (TE-ACT) differ by 7.17 sec, indicating that TTC significantly underestimates the crash exposure experienced by PTWs. Similarly, the severity of the potential crashes estimated using ACT (TI-ACT) is nearly twice that obtained from TTC (TIT), which also highlights the limitation of TTC in capturing the overall crash risk. In order to statistically verify the difference of crash exposure and severity estimated from ACT and TTC, a two-sample *t*-test was performed. Table 5.2 shows the test results at 95% confidence level. Test results indicate that the crash exposure and severity estimated from

TTC and ACT are statistically different. The primary cause of the underestimation of the risk is owing to the inherent limitation of TTC that it can only capture the rear-end crash risk. However, in the real field, rear-end crash is among the several other crash types. Therefore, while evaluating the safety levels at different traffic facilities, it is imperative to consider ACT instead of TTC since the conflict types are not only the rear-end collisions.

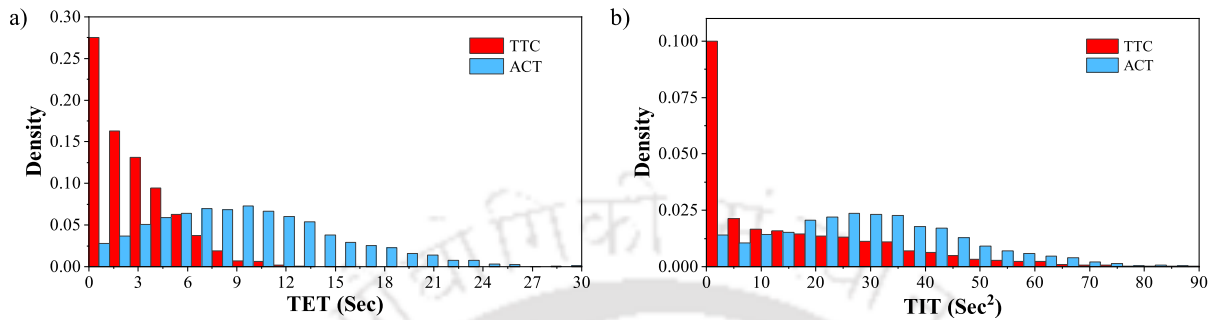


Figure 5.9: Comparison of distributions of crash exposure and severity estimated from ACT and TTC; a) Distribution of duration crash exposure derived from ACT and TTC; b) Distribution of crash severity derived from ACT and TTC

5.2.4 Comparison of Actual and Estimated Crash Frequencies

It is to be noted that the ACT captures all the interactions that happen in a traffic stream. However, all interactions may not end up in a crash. One obvious fact is that if the time remaining between two vehicles is short, they are more likely to collide. Therefore, it can be considered that the vehicles with shorter time-proximity are prone to a crash. The ACT* helps to categorize safe and unsafe interactions based on time proximity. There are several arguments on selecting an appropriate ACT* for demarcating safe and unsafe scenarios in the literature. The authors believe that it is vital to map the actual crash events to the estimated ones to identify the ACT*. However, researchers assume particular ACT* (1,1.5,2,3,4 sec are some of the commonly chosen ACT* values) in the absence of actual crash data.

The comparability of the collision estimates from SSMs with the actual crash occurrences was always been a debate in the proactive safety studies. The threshold value has a significant role in

Table 5.2: Statistical comparison of duration of crash exposure and crash severity estimated using TTC and ACT using two sample t-test

Parameter	Condition	t Statistic	DF	Prob> t
Exposure	Equal Variance Assumed	-33.94	6250	6.22E-232
Severity		-23.53	6250	2.32E-117
Exposure	Equal Variance NOT Assumed (Welch Correction)	-33.94	3392.30	1.12E-217
Severity		-23.53	4836.94	5.46E-116

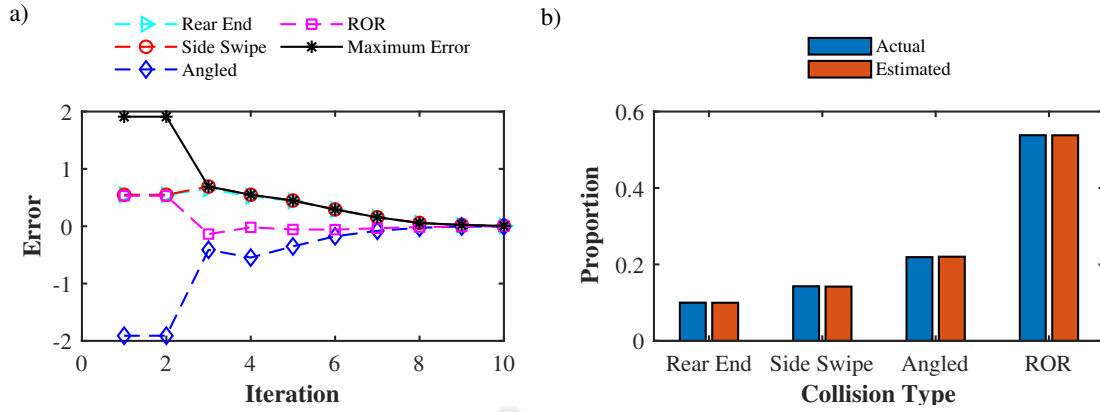


Figure 5.10: Estimation of the ACT-threshold using the simultaneous optimization; a) Error minimization; b) Comparison of the actual and estimated conflicts using the optimal ACT-thresholds.

the projection of the estimated collisions to the actual crash scenario of the study sites. Theoretically, the estimated crash frequency should be mapped to the actual for estimating an appropriate threshold value that can replicate the field scenario. We highlight that ACT can replicate the actual field scenario after optimally selecting the threshold values for different crash types. In order to prove this, we used data from a location where both crash history and trajectory data are available (Shillong bypass, NH-48 connecting Guwahati and Shillong). The threshold values were estimated using a simultaneous optimization method. Fig. 5.10a shows the error minimization at each iteration, and Fig. 5.10b shows the collision frequency estimated using the optimal threshold values. The optimal threshold values achieved after the optimization for each conflict type were: $ACT_{RearEnd}^* = 5.58$, $ACT_{SideSwipe}^* = 4.31$, $ACT_{Angled}^* = 0.47$, and $ACT_{ROR}^* = 1.78$. Notably, rear-end crashes account for only 10% of all crashes in the study area; thus, using TTC to assess crash risk would significantly underestimate the risk.

5.3 Summary and Conclusion

This study developed a novel surrogate safety indicator, the Anticipated Time to Collision (ACT), which is independent of the collision type, road geometry, and the traffic scenario. With this contribution, the present study opens up a new avenue for proactive safety analyses, where the overall crash risk could be estimated without relying on historical crash data. The ACT comparison with TTC reveals that the latter underestimates the crash risk as it may not be applicable to all the crash types.

The key benefits of the proposed method include:

1. It works for all collision types, regardless of the angle of approach
2. This method can be applied to estimate the crash likelihood at various transportation infrastructures such as junctions, curves, mid-blocks, and also at transport network-level
3. A common severity hierarchy can be achieved, since the proposed approach works equally

across all conflict situations and transport facilities

4. In addition to the safety measure, this approach offers conflict-type detail, such as the head-on, rear-end, side-swap, angled, and Run-Off-Road (ROR) crashes
5. Several other indicators can be derived to quantify the crash exposure, crash severity, and the time at which vehicle respond to an unsafe situation
6. The conflict that involves multiple vehicles can be evaluated

The Time of Evasive Action (TEA) is one of the important contributions of this research work. This indicator enables a detailed understanding of the response patterns of vehicles towards an unsafe situation. A comprehensive knowledge of vehicle response patterns toward an unsafe situation is essential to design reliable control strategies for autonomous vehicles and develop better collision warning systems. To capture the crash exposure and severity, two other measures were also derived from ACT profile, which are TE-ACT and TI-ACT, respectively. All the measures together give a clearer picture of crash occurrences in terms of their frequency, exposure time, severity, and evasive actions.

This study also highlights the importance of choosing an appropriate ACT threshold value according to the study purpose. Different threshold values for different vehicle classes needs to be considered to estimate the crash likelihood corresponding to a predefined crash exposure or severity. Whereas, a single threshold value for all vehicle types is useful while comparing the crash severity or exposure.

Chapter 6

Application of Proposed Trajectory-Based Surrogate Safety Analysis at Various Traffic Scenarios

This chapter presents some of the applications of the proposed trajectory-based proactive safety analysis. First, the safety and comfort during an LC maneuver were assessed. In this regard, we proposed a method to automatically identify the LC vehicles and the time window in which LC occurs. Further, the LC maneuvers were classified into interrupted (ILC) and uninterrupted (ULC) LCs, and estimated risk and discomfort, separately for each LC type. Second, powered-Two-Wheelers (PTWs) safety in an urban environment was evaluated and compared to that of other vehicle types. Third, the safety of vehicles traversing a 3-Dimensional road geometry was evaluated and studied the impact of the road geometric features on the vehicle's safety. Last, we evaluated the safety at various components of a transportation network and identified the high-risk sites in terms of crash frequency, exposure, and severity.

6.1 Microscopic Outlook of the Lane-Change Dynamics

A comprehensive understanding of lane changing dynamics is important for developing realistic simulation tools, autonomous LC systems with improved safety and comfort, or proactive LC safety assessment tools. The LC maneuver is generally considered as a continuous process from the original to the target lane, without any intermediate interruptions. Such an assumption helps in a deterministic and straightforward LC modeling and trajectory planning. However, recent studies have shown that LC dynamics are not necessarily continuous but disrupted intermittently [143, 189, 291]. The disruptions could be attributed to the intermittently perceived risks or discomfort and the associated evasive actions by the driver. The risks could arise due to the closing-in of lag vehicle in the target lane faster than the anticipated rate, emergency braking of the leader vehicle, or personal preference of LC drivers. Excessive lateral acceleration, steering, or frequent braking,

might result in driver discomfort. These interruptions and the evasive actions cause a longer LC duration, compared to an uninterrupted LC operation.

Traffic simulation models consider the lane-changing as either instantaneous or a process performed during a fixed period, disregarding the intermediate actions. Similarly, in autonomous LC applications, the trajectory planning depends predominantly on an S-shaped parametric curve (Fig. 2.4d) or a sinusoidal lateral acceleration profile (Fig. 2.4f). A preliminary investigation into the real-field LC scenarios has revealed that the above assumptions are highly restrictive. Fig. 2.4 shows two LC trajectories, which are different in terms of their profile and duration. The existing LC studies primarily assume all LCs to be similar to Fig. 2.4a. But, for an effective modeling of LC operation, it is necessary to differentiate and gain a detailed understanding of the two types of LC operations shown in Fig. 2.4. The benefits of such an understanding are, but not limited to, *i*) train the CAVs to understand interrupted LC operations of human drivers in a mixed traffic environment; *ii*) development of a more dynamic LC trajectories in the case of autonomous LC; *iii*) realistic simulation of LC operations.

The present study performs a detailed investigation of LC operations and proposed to classify them into interrupted LC (ILC) and uninterrupted LC (ULC), besides highlighting their characteristic differences. The scope of this study is limited to a single lane change operation, and the dynamics of multiple LCs are not addressed. Here, we make three major contributions. First, a methodology to identify the time window (referred to as LC window here on) within which the LC maneuver takes place. The lateral speed profile of the LC vehicle was segmented to identify the LC window. Second, a method to classify the LC trajectories into ILC and ULC, based on the interruption levels during an LC. The interruption level was quantified based on the cumulative distance traversed by an LC vehicle during the deceleration and acceleration phases, respectively. Third, highlighting the characteristic difference of ILC and ULC. The risk and comfort during ILC and ULC were compared to quantify the characteristic differences. This study has shown that the ILC takes more time compared to ULC, to complete the LC operation. Therefore, it is necessary to develop separate LC duration models for ILC and ULC.

6.1.1 Data Preparation and Filtering

To investigate the characteristics of ILC and ULC, we used the Next Generation Simulation (NG-SIM) trajectories. The trajectory data collected from both I-80 and US-101 were considered for the study. There was a total of 45 minute data from each location, covering 640 meters in the case of US-101, and 500 meters for I-80. All of the trajectories were reconstructed to remove the embedded noise using the method proposed in this research work [13, 64]. The kinematic variables, such as speed and acceleration, in lateral and longitudinal directions, were estimated from the reconstructed path data.

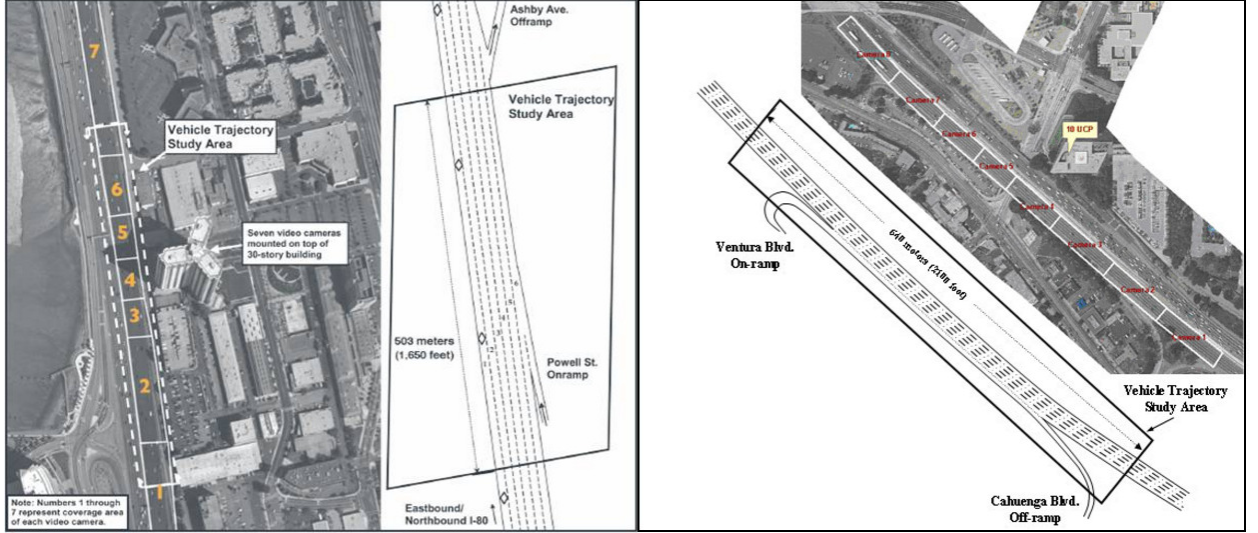


Figure 6.1: Locations of the trajectory data collection

6.1.2 Identification of the Time Window Corresponding to Lane Change Operation

Unlike the existing methods, this study proposes an automated method for defining the LC window as follows:

Let $\mathcal{X}_j = \mathbf{p}_1^j, \mathbf{p}_2^j, \dots, \mathbf{p}_{N_j}^j$ be the trajectory of j^{th} vehicle, where N_j is the length of the trajectory. In this work we considered each point $\mathbf{p}_i^j = [t_i^j, x_i^j, y_i^j, v_{y,i}^j, a_{y,i}^j]$ as a vector in five dimensional space, where t_i^j corresponds to time, x_i^j , y_i^j , $v_{y,i}^j$, and $a_{y,i}^j$ correspond to the longitudinal position, lateral position, lateral speed, and lateral acceleration of the vehicle, respectively. $\mathcal{X}_j$ is the input for both LC widow detection as well as LC classification.

In the first stage, all the trajectories involving LC operation were identified using the lane occupancy count, and the maximum lateral displacement. Here, the lane occupancy count is the number of lanes occupied by the vehicles while traversing the study road stretch. The trajectories with lateral displacements more than 1.5 times the vehicle width were identified, and then verified for the lane occupancy count. It is noteworthy that occupying multiple lanes can be corresponding to either LC or overtaking operations. In this study, we consider only the cases corresponding to single lane change. Therefore, the trajectories corresponding to overtaking operations and multiple lane changes are separated and excluded from the further analysis. Equation 6.1 shows the conditions for grouping the lateral maneuvers:

$$\text{Maneuver Type} = \begin{cases} 1 \text{ (Single LC)}, & \text{if } \Delta y_j > 1.5 \times w_j \text{ \& } \Phi(l^j) = 2 \text{ \& } l^j(1) \neq l^j(N_j) \\ 2 \text{ (Multiple LC)}, & \text{if } \Delta y_j > 1.5 \times w_j \text{ \& } \Phi(l^j) > 2 \text{ \& } l^j(1) \neq l^j(N_j) \\ 3 \text{ (Overtaking)}, & \text{if } \Delta y_j > 1.5 \times w_j \text{ \& } \Phi(l^j) > 1 \text{ \& } l^j(1) = l^j(N_j) \end{cases} \quad (6.1)$$

where, ‘Maneuver Type’ is a categorical variable; Δy_j is the maximum lateral displacement for vehicle j ; w_j is width of vehicle j ; l^j is the vector of lane occupancy indicator; and Φ is a function that returns the size of the unique values in l^j . Further, the LC window corresponding to each LC

trajectory is defined as follows:

Definition 6.1.1. LC Window: LC window is the time interval where the following conditions are satisfied. 1) the lateral speed of the vehicle increases from zero state, reaches the maximum, then decreases back to zero. 2) within this time frame, the vehicle must be occupying exactly two lanes. 3) Maximum lateral displacement during this period should be more than 1.5 times of the vehicle width.

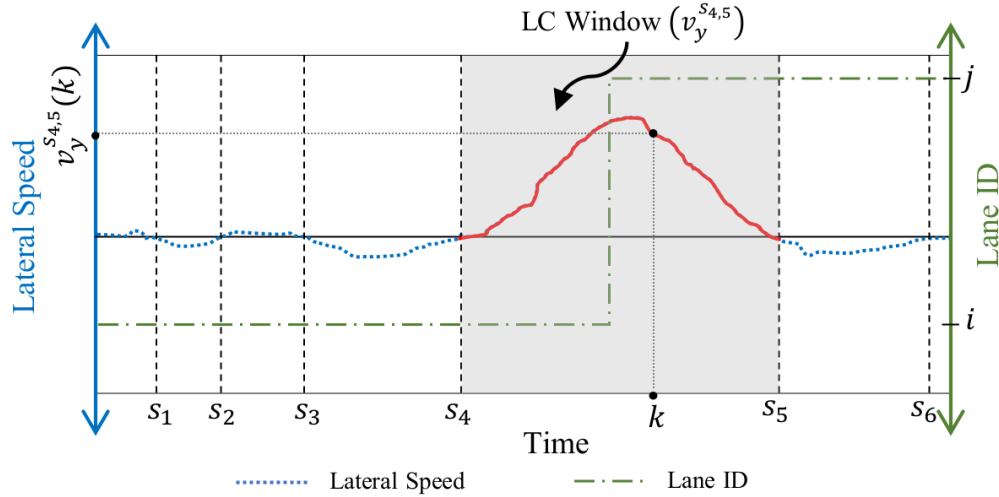


Figure 6.2: Method of lateral speed segmentation for LC window identification

Each LC vehicle trajectory is segmented based on the lateral speed profile, and analyzed for the conditions corresponding to LC window. Fig. 6.2 shows a schematic representation of the LC window identification procedure. To be specific, for each trajectory $\mathcal{X}_j$, the lateral speed profile v_y is segmented as $v_y^s = v_y(s_l : s_m)$, where s_l and s_m are the start and end times of the segment. The criteria for segmentation are as follows:

$$v_y^{s_l, m} = v_y(s_l : s_m) \subset v_y \mid s_l < s_m \quad \forall l, m = 1, 2, \dots, s, \text{ and} \quad (6.2)$$

$$|v_y^{s_l, m}(k)| > 0 \quad \forall s_l < k < s_m$$

From the identified segments, the LC window is selected if the following criterion (Equation 6.3) is satisfied.

$$\text{LC Segment}_j = \{\Phi(l^j((s_l : s_m))) = 2 \ \& \ \Delta y_j > 1.5 \times w_j\} \quad (6.3)$$

Applying Eqns. 6.2 & 6.3, the LC windows corresponding to all the LC operations were identified. Fig. 6.3a shows an example of the identified LC window, and Fig. 6.3b shows some of the identified (normalized) vehicles paths.

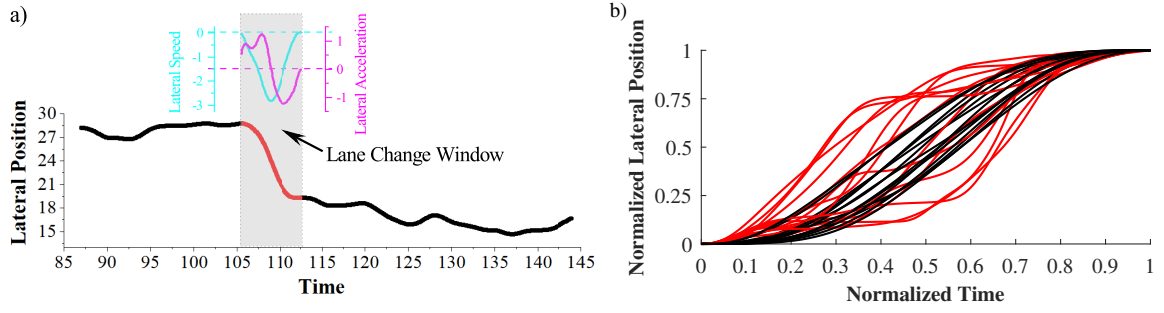


Figure 6.3: a) An example of LC window detected using the proposed method; b) Normalized lateral position profiles of a few vehicles within the LC window (ULC (black), ILC (red))

6.1.3 Classification of LC Trajectories into ILC and ULC

LC trajectories presented in Fig. 6.3b clearly shows two patterns: a smooth and S-shaped curve and the other is with a stair-like profile. It is to be noted that a continuous, s-shaped LC profile corresponds to ULC (highlighted using black color in Fig. 6.3b). In contrast, the stair-like profiles correspond to ILC (highlighted using red color in Fig. 6.3b). The stair-like profile of ILC could be due to intermittent interruptions. Most of the LC path planning studies have used S-shaped parametric curves [199, 292, 293], such as B-spline, polynomial, circular, spiral, trapezoidal, clothoid, Bezier's, and hyperbolic tangent [294]. However, these models cannot capture the interrupted LC operations [143]. In order to capture the intermediate interruptions, studies have considered dynamic methods that optimize safety and comfort [143, 166]. Thus, it is essential to separate ILC and ULC and develop separate models.

An LC operation can be divided into two phases. First, a laterally accelerating phase till it reaches the maximum lateral speed (v_y^{max}) at time τ^* . In the second phase, the vehicle laterally decelerates till its lateral speed is zero. In ideal conditions, the lateral speed profile corresponding to LC would be perfectly concave (Fig. 2.4e). Considering this feature, The present study proposes a method to distinguish the ILC and ULC.

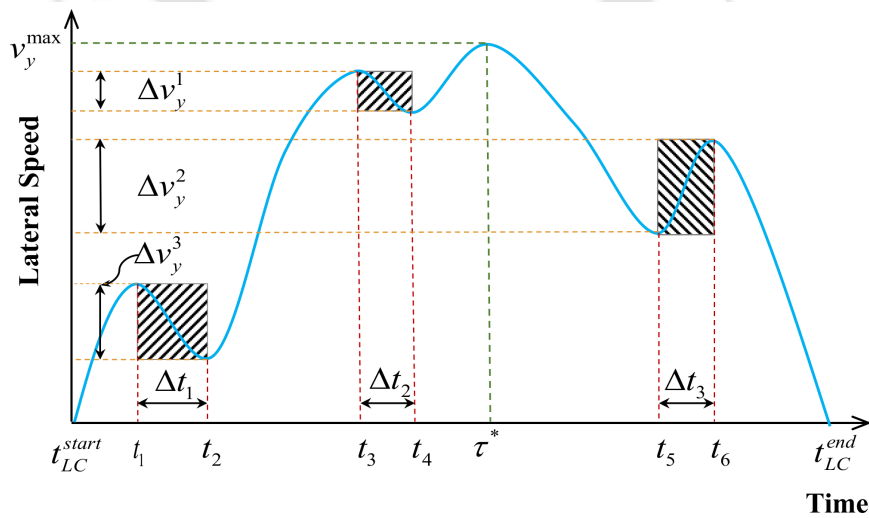


Figure 6.4: Schematic representation LC interruptions reflected in the lateral speed profile

Let t_{LC}^{start} be the starting time of the LC operation, and t_{LC}^{end} corresponding to the end. The acceleration phase Δt_{LC}^{acc} can be defined as,

$$\Delta t_{LC}^{acc} = t_i \mid t_{LC}^{start} \leq t_i \leq \tau^*, \quad i = 1, 2, \dots, \quad \text{and} \quad |\Delta t_{LC}^{acc}| = \tau^* - t_{LC}^{start} \quad (6.4)$$

Similarly, the deceleration phase, Δt_{LC}^{dec} can be defined as,

$$\Delta t_{LC}^{dec} = t_i \mid \tau^* \leq t_i \leq t_{LC}^{end}, \quad i = 1, 2, \dots, \quad \text{and} \quad |\Delta t_{LC}^{dec}| = t_{LC}^{end} - \tau^* \quad (6.5)$$

Within Δt_{LC}^{acc} , ideally, it is expected that the vehicle continuously accelerates laterally till the lateral speed is maximum by time, τ^* . Yet, when there is an interruption to the LC maneuver, the vehicle would decelerate to avoid any potential crash risk with any of the surrounding vehicles or to achieve the desired comfort. Fig. 6.4 shows that, from time t_1 to t_2 and t_3 to t_4 , the LC vehicle is interrupted by the surrounding traffic. Similarly, within Δt_{LC}^{dec} , the LC vehicle is expected to decelerate continuously from τ^* to a near-zero lateral speed. However, Fig. 6.4 indicates that from time t_5 to t_6 , the LC vehicle accelerated probably due to perceived danger or discomfort.

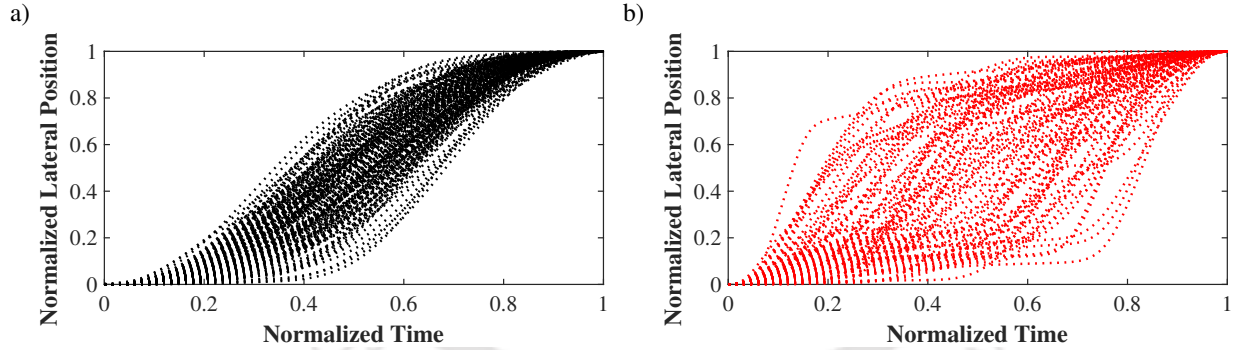


Figure 6.5: Lane change trajectories corresponding to; a) ULC; b) ILC

It is hypothesized that, in the case of an interrupted LC, the vehicle covers some extra distance in order to minimize the crash likelihood or to gain comfort. The proposed method uses the extra cumulative distance traversed by LC vehicles during the lateral acceleration and deceleration phases to classify LC into ILC and ULC. Moreover, this quantity could be used as an indicator for the level of interruption. Shaded rectangles in Fig. 6.4 represent this measure graphically. The level of LC interruption is nothing but the total area of the rectangles. Mathematically the level of interruption (Υ_{LC}) can be defined as follows,

$$\Upsilon_{LC} = \underbrace{\sum_{i=1}^n \Delta v_i \times \Delta t_i}_{\text{Acceleration phase}} + \underbrace{\sum_{j=1}^m \Delta v_j \times \Delta t_j}_{\text{Deceleration phase}} \quad (6.6)$$

For a perfectly uninterrupted LC operation, $\Upsilon_{LC} = 0$. But, in reality, there could be certain deviations. We found that a threshold of $\Upsilon_{LC}^{threshold} = 0.05$, which was selected after a detailed investigation of the ULC profiles, could be more appropriate for distinguishing ILC and ULC. After grouping the LC trajectories into ILC and ULC, an automatic verification is performed by comparing the LC duration, and the extra distance traveled, to rectify the incorrect classifications. If an LC is classified as uninterrupted, but the duration is higher than 25 seconds and the area is close to $\Upsilon_{LC}^{threshold}$, then it is reclassified as ILC. An ILC with the area close to $\Upsilon_{LC}^{threshold}$ and the duration less than 6 second is reclassified as ULC. Such a fine-tuning would help to minimize the effect of the pre-set threshold value on the classification process. Fig. 6.5a shows the ULC trajectories extracted using the proposed measure, and Fig. 6.5b shows the trajectories corresponding to ILC.

6.1.4 Characteristics of Interrupted and Uninterrupted Lane-Changes

This section presents the characteristics of ILC and ULC such as the crash likelihood and the discomfort of the driver. Further, the LC duration was compared and separate duration models were developed for ILC and ULC.

6.1.4.1 Risk and Discomfort Associated to ILC and ULC

This study conjunctures that the crash likelihood as well as discomfort level may differ for ILC and ULC. To prove this, we have estimated the crash risk corresponding to ILC and ULC, using ACT, TE-ACT, and TI-ACT.

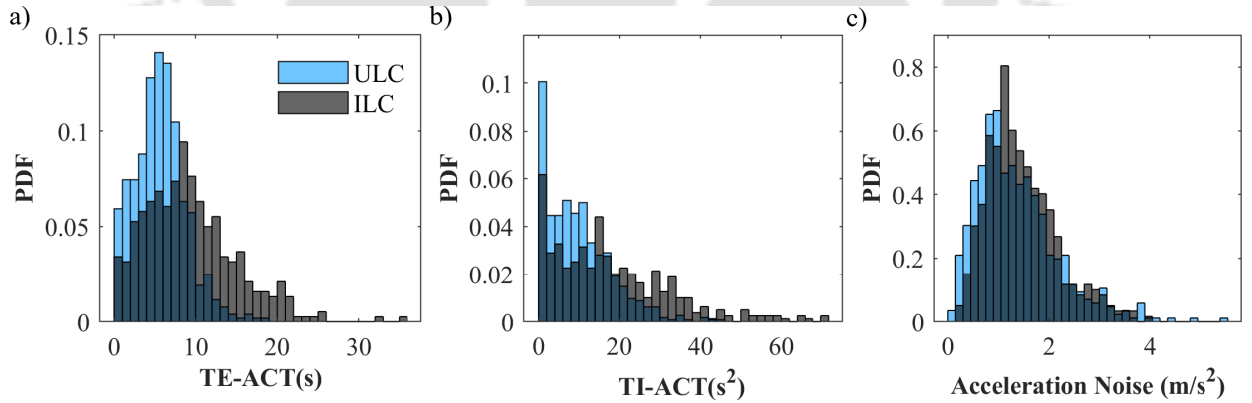


Figure 6.6: Histogram of a) TE-ACT; b) TI-ACT; and c) Acceleration Noise

TE-ACT estimates the duration for which the LC vehicle is exposed to a safety critical scenario during the course of LC. Whereas, TI-ACT gives an idea about the severity of a potential crash during LC. A higher TE-ACT or TI-ACT value indicates a higher crash exposure or crash severity, respectively.

The discomfort during LC was quantified using the Acceleration Noise (AN), which is one of the commonly used measures of driver discomfort [295]. The acceleration noise was estimated

using the following equation:

$$AN = \sqrt{\left(\frac{1}{n}(a_i - \bar{a})^2\right)} \quad (6.7)$$

Where, a_i is the acceleration of LC vehicle at i^{th} instance, and $\bar{a}$ is the average acceleration of the vehicle. A higher AN indicates a higher discomfort.

TE-ACT, TI-ACT, and AN distributions were compared to demonstrate the difference in risk and driver comfort associated with ILC and ULC. Figure 6.6a, b, & c show the distribution of ILC, ULC, and AN, respectively. In the case of ILC, the figure clearly shows that the probability corresponding to longer exposure periods to an unsafe situation is relatively higher (Figure 6.6a). Similarly, in the case of ILC, the vehicle is more likely to undergo a severe crash (Figure 6.6b). Using the maximum log-likelihood criterion, the best fitting distributions to all these attributes were identified. It was found that the TE-ACT follows Nakagami distribution for both ILC and ULC. Similarly, TI-ACT follows Generalized Pareto Distribution, for both ILC and ULC, and AN follows Gamma distribution for both ILC and ULC. Table 6.1 shows the fit statistics along with the parameters. Though the distributions were similar for all the attributes corresponding ILC and ULC, the parameters were different. This indicates that the risk and discomfort probabilities associated to ILC and ULC are different. Therefore, while probabilistic estimation of overall risk during LC activity, as in the study of Park et. al., [296], it essential to consider different probabilities for ILC and ULC.

Further, the Kruskal-Wallis ANOVA test was conducted to highlight the difference between ILC and ULC in terms of crash exposure, crash severity, and driver discomfort. The test hypothesis is as stated below:

H_0 : The samples come from the same population

Table 6.2 shows the test statistics. At 95% confidence, the Kruskal-Wallis ANOVA test proves that the safety and comfort during ILC and ULC are significantly different. Moreover, the average TE-ACT value for ILC and ULC indicates that, during ILC the vehicles undergo a prolonged exposure to risk compared to ULC. Similarly, average TI-ACT values indicate that the crash severity is significantly higher for ILC compared to ULC.

Although the distributions for AN corresponding to ILC and ULC are statistically different, their average values do not give any insight into the driver discomfort levels. Therefore, the present study considered another measure to quantify the driver's discomfort during LC. Notably, a part of the discomfort during LC could have resulted from excessive braking or steering. Braking events can be quantified as the number of sign changes in the longitudinal acceleration profile. Similarly, the steering events can be quantified as the number of sign changes in the lateral acceleration profile. Ideally, in the case of ULC, it was expected to have a single sign change for both the lateral and longitudinal accelerations. While for ILC, there can be multiple sign changes in both lateral and longitudinal acceleration profiles, due to the driver's evasive actions to reduce the perceived risk. Figure 6.7a presents the average number of sign changes for lateral and longitudinal acceleration profiles corresponding to ILC and ULC. Similarly, Figure 6.7b presents the modes

Table 6.1: Probability Distribution of various LC attributes capturing safety and comfort

Attribute	LC Type	Distribution	NLogL	BIC	AIC	AICC	Parameters
Acceleration Noise	ILC	Gamma	618.45	1249.59	1240.89	1240.92	shape (a) = 2.90 scale (b) = 0.48
	ULC		389.48	790.93	782.96	782.99	shape (a) = 4.33 scale (b) = 0.34
TE-ACT	ILC	Nakagami	1164.74	2341.37	2333.48	2333.52	shape (μ) = 0.76 scale (ω) = 117.92
	ULC		1321.83	2656.19	2647.66	2647.68	shape (μ) = 0.83 scale (ω) = 41.76
TI-ACT	ILC	Generalized Pareto	1534.74	3087.43	3075.47	3075.53	shape (k) = -0.24 scale (σ) = 22.19 threshold (θ) = 2.22E-15
	ULC		1862.82	3744.69	3731.64	3731.68	shape (k) = -0.21 scale (σ) = 11.74 threshold (θ) = 2.22E-15
Duration	ILC	Log Normal	1412.90	2838.50	2829.90	2829.90	log location (μ) = 2.57 log scale (σ) = 0.31
	ULC		880.01	1776.20	1768.00	1768.00	log location (μ) = 2.10 log scale (σ) = 0.25

Table 6.2: The Kruskal-Wallis ANOVA test statistics

Attribute	LC Type	Average	Standard Deviation	Mean Rank	Sum Rank	Chi-Square	p-Value
TE-ACT	ILC	9.3 s	5.6 s	560.58	213579.5	110.09	9.37E-26
	ULC	5.7 s	3.1 s	375.79	197291.5		
TI-ACT	ILC	17.9 s ²	14.5 s ²	583.08	232067.5	81.91	1.43E-19
	ULC	9.8 s ²	8.1 s ²	417.60	238867.5		
AN	ILC	1.6 m/s ²	0.7 m/s ²	513.05	204195	6.53	0.01
	ULC	1.4 m/s ²	0.8 m/s ²	466.33	266740		
Duration	ILC	13.43 s	5.26 s	611.90	307175.5	392.54	2.31E-87
	ULC	9.63 s	3.06 s	264.65	108240.5		

corresponding to the number of sign changes. It is evident from these figures that, during ILC, the vehicle undergoes a relatively higher number of sign changes, for both the lateral and longitudinal accelerations. This indicates that during ILC, there would be a frequent braking and steering to ensure safe and comfortable LC operation. Nevertheless, it is most likely that the ULC to have a single sign change.

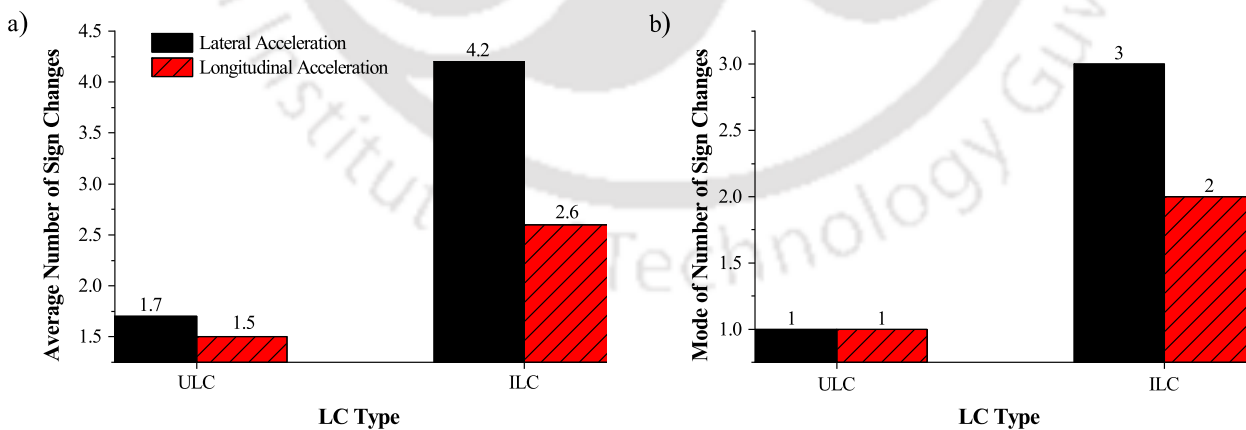


Figure 6.7: a) Average number of sign changes in the lateral and longitudinal acceleration profiles; b) Mode of number of sign changes in the lateral and longitudinal acceleration profiles

6.1.4.2 Duration of ILC and ULC

Further, the LC durations corresponding to both ILC and ULC were compared. Many studies have shown that LC duration follows a Log-Normal distribution [194, 195, 197]. In line with the literature, this study found that the LC duration for both ILC and ULC follows Log-Normal distribution, though with different parameters. The duration histogram and a probability plot with 95% confidence bound are shown in Fig. 6.8c and d, respectively for ULC and ILC. The duration distribution parameters of ULC and ILC and the test statistics are shown in Table 6.1 & 6.2, respectively. From the analysis, it is evident that the LC duration for ILC and ULC are statistically different. Therefore, it is inevitable to model the duration corresponding to ILC and ULC separately.

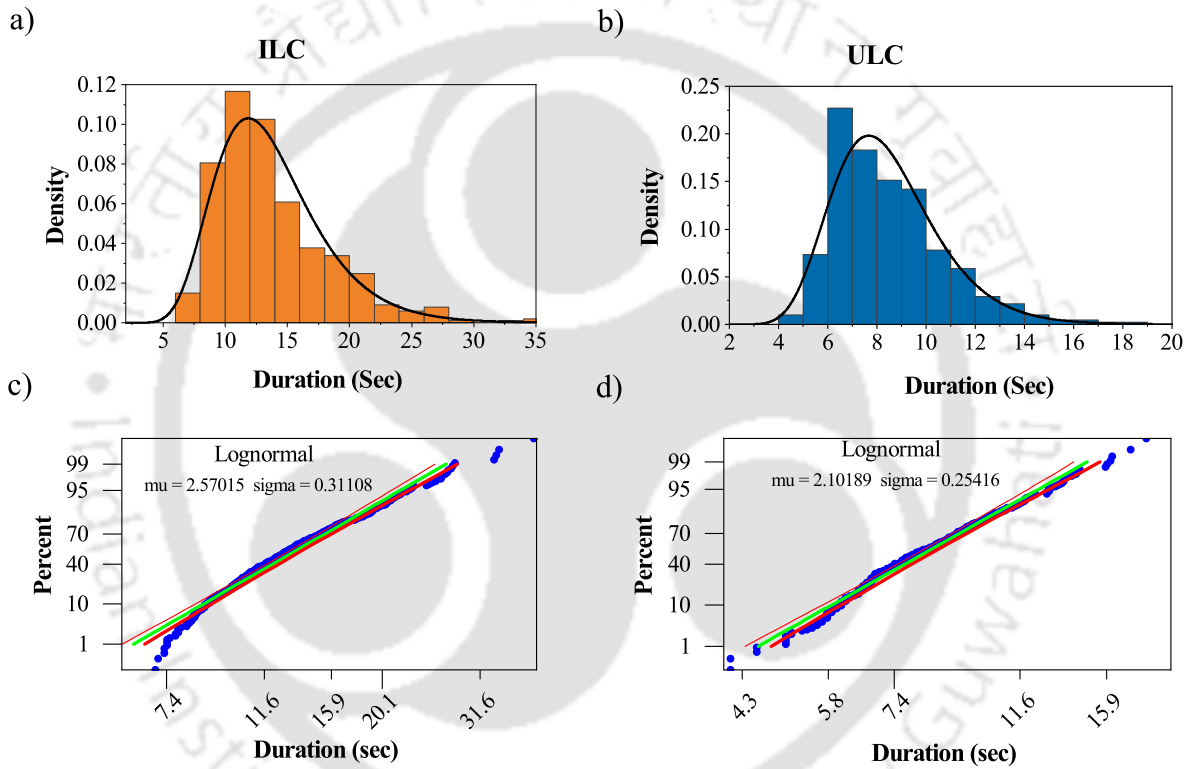


Figure 6.8: LC duration distribution; a) ILC; b) ULC; c) probability plot of Log-Normal distribution of ILC duration; d) probability plot of Log-Normal distribution of ULC duration

Lane-Change Duration Modeling Since the duration is non-negative, the present study has considered the Log-Linear function to model the LC duration.

$$\ln(\Delta t_{LC}) = \beta \mathbf{X} + \epsilon \quad (6.8)$$

where,

Δt_{LC} = Lane change duration

$\mathbf{X}$ = Vector of explanatory variables

β = Coefficient vector

ϵ = Error term

Now, the main challenge is the selection of the explanatory variables $\mathbf{X}$ to develop the duration model. It is evident from Fig. 6.9 that the LC vehicle has a direct interaction with multiple vehicles (leader and follower vehicles in the current and target lanes). Table 6.3 shows the characteristics of the surrounding vehicles commonly considered in the LC studies. The descriptive statistics of the variables presented in this table indicate a wide range of conditions under which the LC maneuvers were observed.

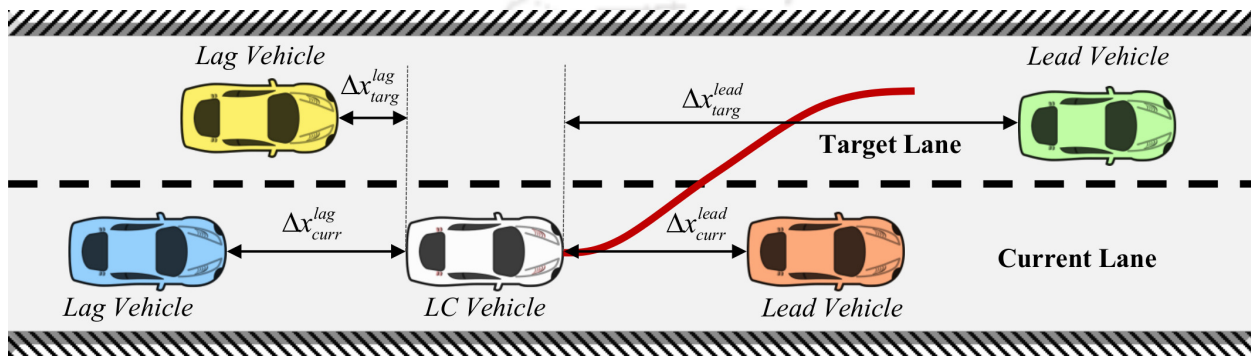


Figure 6.9: Schematic diagram of a lane-changing process

This study considers the stepwise multiple linear regression to identify the most significant variables for the LC duration modeling. The p-value of an F-statistic was used to test the models with and without a potential explanatory variable. The multi-collinearity of the explanatory variables was also checked using the Variable Inflation Factor (VIF). The most commonly used rule of thumb, $VIF \geq 10$, was used to avoid the possible cases of multi-collinearity [297]. Considering the most significant explanatory variables, a multiple linear regression was performed to estimate the model parameters. Table 6.4 presents the test statistics and the model parameters corresponding to ULC, ILC, and combined LC duration models. ANOVA was performed to test the significance of these models. The test results show that all the models are statistically significant at 1% significance.

Though a few explanatory variables are common among the models, each model is different, as indicated by the standardized coefficients. In the case of ULC, the gap between the lead vehicles in the current and target lanes contributes more to the prediction. This variable has a positive weight in the model, indicating that if there is a more gap available in the target lane, the LC vehicle takes relatively more time in the case of ULC. Such a behavior could be attributed to the driver's choice to perform a more comfortable LC, in terms of steering, with minimal/no interruption from the lead vehicle in the target lane. However, the lag vehicle's speed in the target lane is negatively correlated with the LC duration. This indicates that the driver would be cautious about the lag vehicle's speed in the target lane, thus completes LC operation quickly. Similarly, the lag vehicle's acceleration in the current lane also shows a negative correlation with the ULC duration. This could be due to the LC vehicle's perception of a relatively higher risk of being in the current lane

Table 6.3: Descriptive statistics of variables influencing LC duration for ULC and ILC

Variables	Mean		Std		Median		Min		Max	
	ULC	ILC	ULC	ILC	ULC	ILC	ULC	ILC	ULC	ILC
LC duration(Δt_{LC} (s))	9.63	13.43	3.06	5.26	9.30	12.55	4.10	5.60	25.70	40.50
Speed of LC vehicle(v_{LC} (m/s))	9.12	9.53	3.65	3.73	8.94	9.51	0.00	0.00	18.00	21.00
Desired speed of LC vehicle($v_{d,LC}$ (m/s))	12.83	13.21	3.80	3.72	12.33	12.96	4.60	4.58	26.77	24.80
Normalized speed of LC vehicle($v_{LC}^{normalized}$)	0.71	0.72	0.21	0.21	0.74	0.75	0.00	0.00	1.66	1.23
Acceleration of LC vehicle(a_{LC} (m/s ²))	0.15	0.12	0.56	0.56	0.16	0.10	-1.86	-1.54	2.00	2.65
Speed of lead vehicle in the current lane(v_{lead}^{curr} (m/s))	9.07	9.44	3.58	3.67	9.01	9.29	0.00	0.00	21.31	21.48
Acceleration of lead vehicle in the current lane(a_{lead}^{curr} (m/s ²))	0.06	0.02	0.49	0.54	0.07	0.04	-1.58	-2.20	2.09	1.55
Spacing with lead vehicle in the current lane(Δx_{lead}^{curr} (m))	6.22	6.59	3.50	3.21	5.44	5.92	1.47	1.01	36.65	33.25
Relative speed with lead vehicle in the current lane(Δv_{lead}^{curr} (m/s))	0.02	0.03	0.39	0.49	0.01	0.03	-2.14	-2.51	1.94	2.05
TTC with the lead vehicle in the current lane(TTC_{lead}^{curr} (m/s))	76.75	77.12	161.21	220.86	29.96	28.98	4.35	0.80	1405.00	2633.00
Speed of lag vehicle in the current lane (v_{lag}^{curr})	8.89	9.14	3.49	3.59	8.74	8.99	0.00	0.02	17.11	17.92
Acceleration of lag vehicle in the current lane (a_{lag}^{curr})	0.13	0.13	0.48	0.49	0.08	0.08	-1.83	-1.67	2.56	1.89
Spacing with lag vehicle in the current lane (Δx_{lag}^{curr})	6.84	7.14	3.31	3.19	6.30	6.49	0.79	1.65	23.61	22.67
Relative speed with lag vehicle in the current lane(Δv_{lag}^{curr})	0.07	0.12	0.38	0.37	0.09	0.11	-1.30	-1.06	1.26	1.50
TTC with lag vehicle in the current lane (TTC_{lag}^{curr})	74.97	68.79	179.74	150.71	25.05	27.15	0.62	2.71	1333.50	1614.75
Speed of lead vehicle in the target lane (v_{lead}^{targ})	10.35	10.20	3.63	3.73	10.41	10.08	0.00	0.34	25.30	20.00
Acceleration of lead vehicle in the target lane (a_{lead}^{targ})	0.17	0.18	0.50	0.51	0.16	0.16	-1.66	-2.20	1.92	1.89
Spacing with the lead vehicle in the target lane (Δx_{lead}^{targ})	5.26	6.39	4.77	6.23	3.87	4.78	0.01	0.00	24.19	44.92
Relative speed with the lead vehicle in the target lane (Δv_{lead}^{targ})	-0.38	-0.20	0.74	0.77	-0.35	-0.26	-5.94	-2.85	1.85	3.51
TTC with the lead vehicle in the target lane (TTC_{lead}^{targ})	39.89	51.50	152.49	372.33	9.27	10.42	0.01	0.00	2160.00	6878.00
Speed of lag vehicle in the target lane (v_{lag}^{targ})	9.22	9.23	3.66	3.62	9.20	9.36	0.00	0.00	19.66	19.58
Acceleration of lag vehicle in the target lane (a_{lag}^{targ})	0.21	0.19	0.53	0.45	0.16	0.15	-4.23	-1.52	2.76	1.75
Spacing with lag vehicle in the target lane (Δx_{lag}^{targ})	6.55	7.15	4.53	5.41	5.68	5.92	0.06	0.01	28.46	40.83
Relative speed with the lag vehicle in the target lane (Δv_{lag}^{targ})	-0.03	0.09	0.82	0.81	0.00	0.07	-4.10	-2.72	2.27	3.01
TTC with lag vehicle in the target lane(TTC_{lag}^{targ})	42.76	49.78	141.31	161.86	12.45	13.66	0.04	0.01	1749.00	1739.00
Gap in the target lane (G^{targ})	10.98	11.37	5.52	7.03	9.86	9.66	0.04	0.09	34.72	54.90
Relative speed of lead and lag vehicles in target lane (Δv_{targ})	0.34	0.29	0.55	0.60	0.31	0.27	-1.94	-1.76	3.32	3.06
Gap between lead vehicles in the current and target lanes(ΔG^{lead})	1.62	1.68	5.76	7.31	2.29	2.35	-16.79	-38.82	31.75	30.06

Table 6.4: Statistics of LC duration models

Model	Variables	Unstandardized		Standardized	t	Sig.	Collinearity Statistics	
		Coefficients		Coefficients			Tolerance	VIF
		B	Std. Err.	Beta				
ULC	(Constant)	2.215	.038		59.036	.000		
	a_{LC}	-.079	.024	-.167	-3.335	.001	.858	1.166
	Δx_{lead}^{curr}	-.030	.006	-.443	-4.662	.000	.238	4.193
	Δx_{lead}^{targ}	.035	.006	.664	5.467	.000	.146	6.845
	a_{lag}^{curr}	.057	.027	.108	2.129	.034	.834	1.199
	v_{lag}^{targ}	-.016	.003	-.225	-4.670	.000	.924	1.082
	ΔG^{lead}	.032	.006	.744	5.616	.000	.123	8.154
ILC	(Constant)	2.750	.044		62.371	.000		
	v_{LC}	-.037	.004	-.437	-8.571	.000	.619	1.614
	Δx_{lag}^{targ}	.012	.003	.203	4.159	.000	.676	1.479
	Δv_{lag}^{targ}	.051	.020	.130	2.525	.012	.610	1.640
	a_{lead}^{targ}	-.078	.026	-.126	-3.029	.003	.927	1.079
	ΔG^{lead}	.007	.003	.164	2.848	.005	.487	2.054
	Δx_{lead}^{targ}	.012	.003	.220	3.420	.001	.389	2.573
	G^{targ}	-.007	.003	-.143	-2.209	.028	.387	2.585
LC	Δx_{lag}^{curr}	.010	.005	.099	2.192	.029	.788	1.270
	(Constant)	2.579	.036		71.787	0.000		
	v_{lag}^{targ}	-.030	.003	-.299	-8.828	.000	.789	1.267
	Δx_{lag}^{targ}	.014	.003	.192	4.380	.000	.474	2.109
	a_{lag}^{targ}	-.073	.023	-.099	-3.125	.002	.907	1.102
	ΔG^{lead}	.014	.004	.237	3.869	.000	.242	4.133
	Δx_{lead}^{targ}	.025	.004	.371	6.774	.000	.303	3.301
	ΔG^{targ}	-.011	.003	-.181	-3.243	.001	.291	3.437
	Δx_{lead}^{curr}	-.010	.005	-.094	-2.161	.031	.476	2.102
	a_{LC}	-.062	.023	-.090	-2.685	.007	.803	1.245
	Δv_{lead}^{curr}	-.064	.027	-.075	-2.341	.019	.878	1.139

compared to the target lane, therefore performs the LC quickly.

In the case of ILC, the LC vehicle's speed contributing more to the prediction. Moreover, the LC vehicle's speed correlated negatively to the LC duration. This can be explained with the following hypothetical situation. When a LC vehicle changes lanes with a slower speed, the lag vehicle in the target lane may not yield, assuming that the LC vehicle enters the target lane only after his/her passage. This would lead to a longer LC duration since the maneuver involves evasive actions to avoid conflicts. But, when the LC vehicle is maneuvering at a higher speed, the lag vehicle would yield to avoid a potential crash; thus, the vehicle can complete the maneuver in a lesser time period. The speed difference with the lag vehicle in the target lane is another critical variable that positively affects the ILC duration. It can be inferred that, as the speed difference increases, the lag vehicle driver would be more confident about his/her passing, thereby delaying the LC maneuver. Another important observation is the impact of the lead vehicle's acceleration in the target lane. It is to be noted that the lead vehicle's acceleration in the target lane has a negative impact on the LC duration. If the lead vehicle in the target lane decelerates at a rate that can cause a potential rear-end collision, the LC vehicle might get interrupted. To avoid an unsafe situation, the LC vehicle is forced to perform evasive actions, leading to a higher LC duration. A major observation from this modeling is that, in the case of ILC, the lag vehicles' kinematics in the target lane mainly control the LC duration. Therefore, it is essential to give more importance to the interaction with the lag vehicle in the target lane while LC trajectory planning for an autonomous system, with proper sensing of the lag vehicle's kinematics.

In the combined model, the gap with the leader vehicle in the target lane contributes more to the model. Most of the variables in the combined model are part of either the ULC model or the ILC model, except the lag vehicle's acceleration in the target lane and the relative speed with the lead vehicle in the current lane. RMSE corresponding to the combined model is relatively high, indicating a more extensive scattering of the response variable. To better understand the model performance, the relative importance of the model variables was compared through sensitivity analysis.

Sensitivity analysis gives an idea about how the predicted LC duration's uncertainty can be divided and allocated to different sources of uncertainty in the explanatory variables. Changing one variable at a time (OAT) to see its effect on the output is one of the easiest and most popular ways to perform sensitivity analysis [298]. The other variables are held at their baseline (nominal) while one input variable is changed, and the dependent variable's change is then evaluated. Fig. 6.10 shows the results of the sensitivity analysis. It is evident from the figure that the response variable's sensitivity in the duration models corresponding to ILC, ULC, and combined cases is not identical. The combined model, which is commonly used in the existing practices, leads to an unrealistic prediction of LC duration for the ULC maneuver. For example, the explanatory variables common to all three models, such as the gap with the leader vehicle in the target lane (Δx_{lead}^{targ}) (Fig. 6.10c), and the gap between the lead vehicle in the current and target lanes (ΔG^{lead}) (Fig. 6.10f), affect the LC duration in a differential manner. Therefore, it can be said that the combined model is biased towards the ILC condition, and is significantly away from ULC.

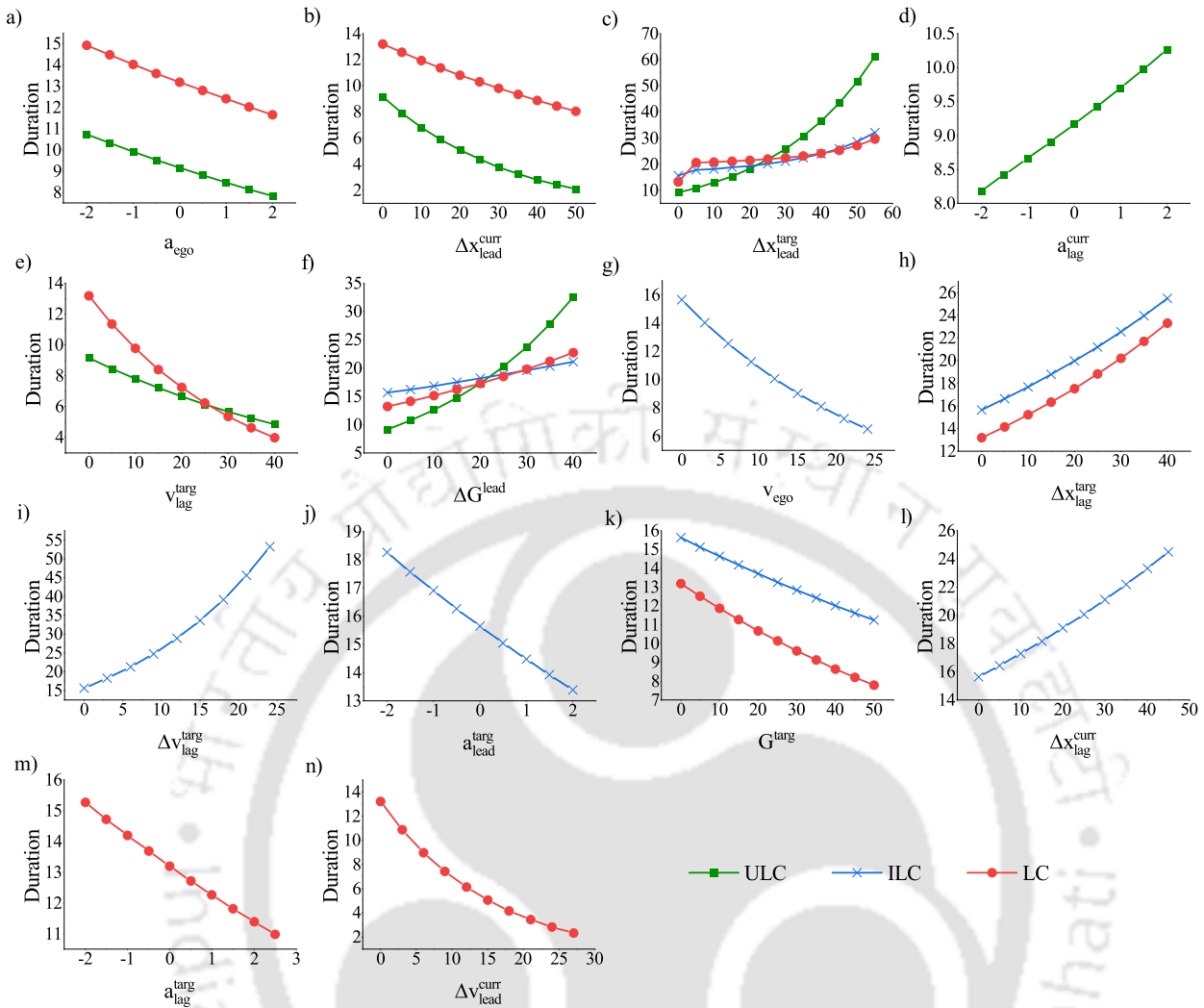


Figure 6.10: Sensitivity analysis of explanatory variables of LC duration models

6.1.5 Summary and Conclusion

To summarize, this study microscopically examined the characteristics of LC dynamics, after classifying LC into ILC and ULC. To that end, we proposed a methodology to automatically identify and classify LC operation into ILC and ULC. In addition, the characteristic differences between ILC and ULC were explored by comparing different LC attributes. We pointed out that, intermediate interruptions during an LC maneuver could increase the crash likelihood, driver discomfort, and duration of the operation.

The main conclusions drawn from this analysis are;

- The method used for detecting the LC window has a significant impact on the estimated LC duration.
- Field data indicates that the LC process was not always continuous. The surrounding vehicles interrupt the LC vehicle, which can be seen from the lateral position, speed, and acceleration

profiles. We recommend a classified analysis of LC to replicate the real field scenario in LC models.

- The interruption level in terms of extra cumulative distance traversed by an LC vehicle in the acceleration and deceleration phases can be used as a criterion for classifying LC scenarios into ILC and ULC.
- Following are the characteristics of ILC and ULC:
 - Compared to ULC, ILC has a higher crash likelihood. The average risk exposure measured using TE-ACT was found higher during ILC. On an average, during ILC, the vehicle exposed to unsafe situation for $9.3sec$, whereas during ULC it is $3.1sec$.
 - Similarly, when compared the TI-ACT values, the crash severity was also found higher for ILC. On an average, during ILC, the crash severity is $17.9s^2$, whereas for ULC it is $9.8s^2$. The crashes during ILC shows a severity level double as that of ULC.
 - During ILC, the drivers undergo a higher discomfort owing to the frequent braking and/or steering operations.
 - The time needed to complete ILC was found higher than that for ULC. Average duration of ILC is $13.43sec$, whereas for ULC it is $9.63sec$. The duration distributions corresponding to ILC and ULC were statistically different. Since duration is an important variable in most LC models, separate duration models for ILC and ULC needs to be considered for a better representation of the ground truth.
 - While probabilistic estimation of overall risk during LC activity, it is essential to consider corresponding risk and discomfort probabilities for ILC and ULC.
 - From the duration models developed in this study, explanatory variables for ILC and ULC were found to be different. Also, the sensitivity of common explanatory variables in ILC and ULC models was not the same.
 - A major observation from the LC duration modeling is that, the lag vehicles' kinematics in the target lane mainly control the LC duration for ILC. Therefore, it is essential to give more importance to the interaction with the lag vehicle in the target lane while LC trajectory planning for an autonomous system, with proper sensing of the lag vehicle's kinematics.

The results from this research illustrate the need for separate analyses of ILC and ULC dynamics in LC studies. Such consideration would enhance LC simulations' accuracy, collision alert systems, and the performance of human-like autonomous LCs.

6.2 Analysis of Powered Two Wheeler Safety in Urban Environment Using ACT

The increasing use of PTWs poses a variety of road safety concerns. It is evident from Fig. 6.11 that the VRUs are involved in around 49% of crashes, of which 47% involve PTWs. In the recent past, most of the developing countries' urban centers are flooded by PTWs, which significantly alters the traffic stream's expected behavior. The challenges related to the integration of PTWs into the traffic system are complex and varied. Addressing the safety and mobility issues associated with the PTW riders necessitates a comprehensive investigation to recognize the means for effective action. These include considering the current risk exposures for riders and the severity analysis of the potential crashes.

This section presents an approach to proactively assess the safety of PTWs in an urban environment, using the trajectory data and several surrogate safety indicators. The analyses presented in this section provide insight into the potential PTW crash pattern in an urban environment. We evaluated the prominent crash types among PTWs, and their frequency, severity, and exposure time. Besides, we investigated the crash generation process and different evasive actions taken by the PTWs.

6.2.1 Data

For this analysis, we considered the trajectory data collected from a busy urban mid-block section with mixed traffic and weak lane discipline. The mid-block is located in the Christian Basti area of Guwahati city, Assam, India. Fig. 6.12 shows the study area. The traffic video data were collected using a fixed camera setup mounted over a vantage point. Later the trajectories were extracted from the traffic videos using SAVETRAX. All the trajectories were processed to remove any noise embedded during the data collection and extraction process, using the method proposed in this research work. A total of 6291 vehicle trajectories were analyzed, out of which 3297 were PTWs.

6.2.2 Results

The proposed approach was applied to all the vehicle trajectories to estimate the ACT profiles. The conflict circumstances involving PTWs were separated to compare PTW safety with respect to other modes. Fig. 6.13 displays ACT-histograms for PTWs and for the other vehicle classes, separately. In order to distinguish safe and unsafe scenarios, an ACT threshold of 3 sec was selected. It can be seen from Fig. 6.13 that the total number of instances at which the ACT profile goes below the threshold value is almost double for PTWs compared to all other transport modes together. It is also evident that regardless of the ACT threshold, PTW crashes have a higher probability than all the other crash occurrences involving other vehicles types together. From the analysis, the PTWs were found to be responsible for 68.4% of the total occurrences of unsafe events (Fig. 6.13). Moreover, it corroborates with the actual field scenario in the study area, where PTWs constitute

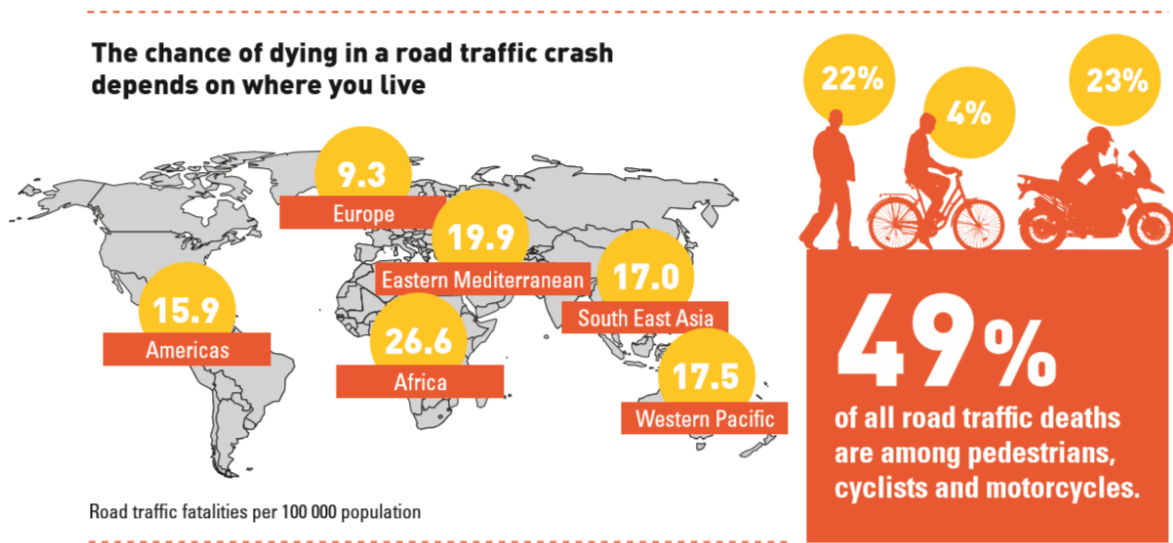


Figure 6.11: The number of fatalities among Vulnerable Road Users around the world (Source: World Health Organization [6])



Figure 6.12: Study area

58.5% of the total road traffic fatalities [299].

6.2.2.1 Analysis of Time of Evasive Action (TEA)

In order to highlight how effectively the ACT captures the unsafe driving scenario, we presented the ACT profiles corresponding to some of the PTWs in Fig. 6.14. To distinguish safe and unsafe scenarios, we set an ACT threshold (ACT^*) of 3sec, and all the instances that go below the threshold value (corresponding to unsafe situations) were represented using a solid line (Fig. 6.14). ACT profiles shown in the figure indicate a smooth recovery from an unsafe situation through performing some evasive actions such as braking or steering with comfort. However, in some cases the ACT profile jumps from unsafe region to safe region (see Fig. 6.14g), indicating an immediate recovery through sudden braking or cut-in action to avoid a potential crash. It is evident from these figures that the ACT could be used to assess the vehicles' evasive actions to avoid crashes comprehensively.

Further, to better understand the evasive actions taken by PTWs during an unsafe scenario, the acceleration profile was compared with the ACT profile. Fig. 6.15 shows the profiles corresponding to ACT, speed, and acceleration for two PTWs. This figure indicates that the vehicles perform evasive actions to avoid unsafe scenarios or to minimize the severity of a potential conflict. The instances when the acceleration profile changes its sign were identified and projected to ACT profile. The acceleration and deceleration instances are shown with different colors for better understanding (Fig. 6.15). It is evident from the figure that the vehicles decelerate at every unsafe situation. A comparison of the instances at which the vehicle enters an unsafe situation ($ACT < ACT^*$) and the start of deceleration has shown that the vehicles react to unsafe situations well before meeting them. In order to capture the driver's response pattern to unsafe situations, we used Time of Evasive Action (TEA). TEA for all PTWs and cars was calculated to see how early a PTW user detects the risk, compared to cars. Fig. 6.16 shows the TEA-histogram corresponding to PTW and cars. Table 6.5 shows the descriptive statistics of TEA estimated for cars and PTWs. It can be seen from

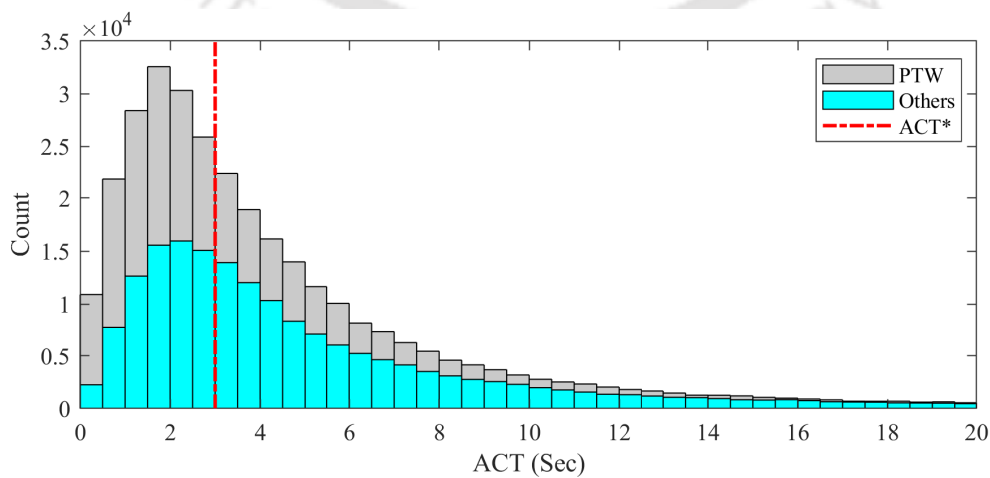


Figure 6.13: ACT-histogram of PTWs and other vehicles

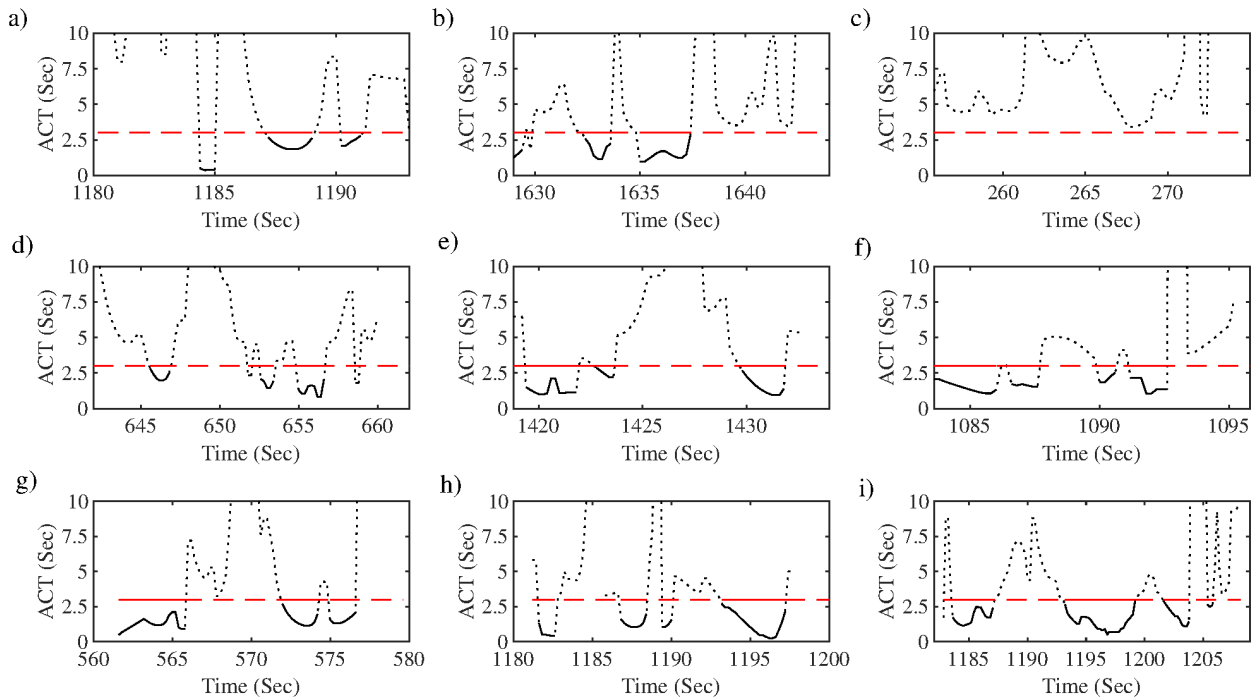


Figure 6.14: ACT profile of PTWs (black line). The thick black line portions in the profile indicate the unsafe region. The red line shows the ACT threshold;

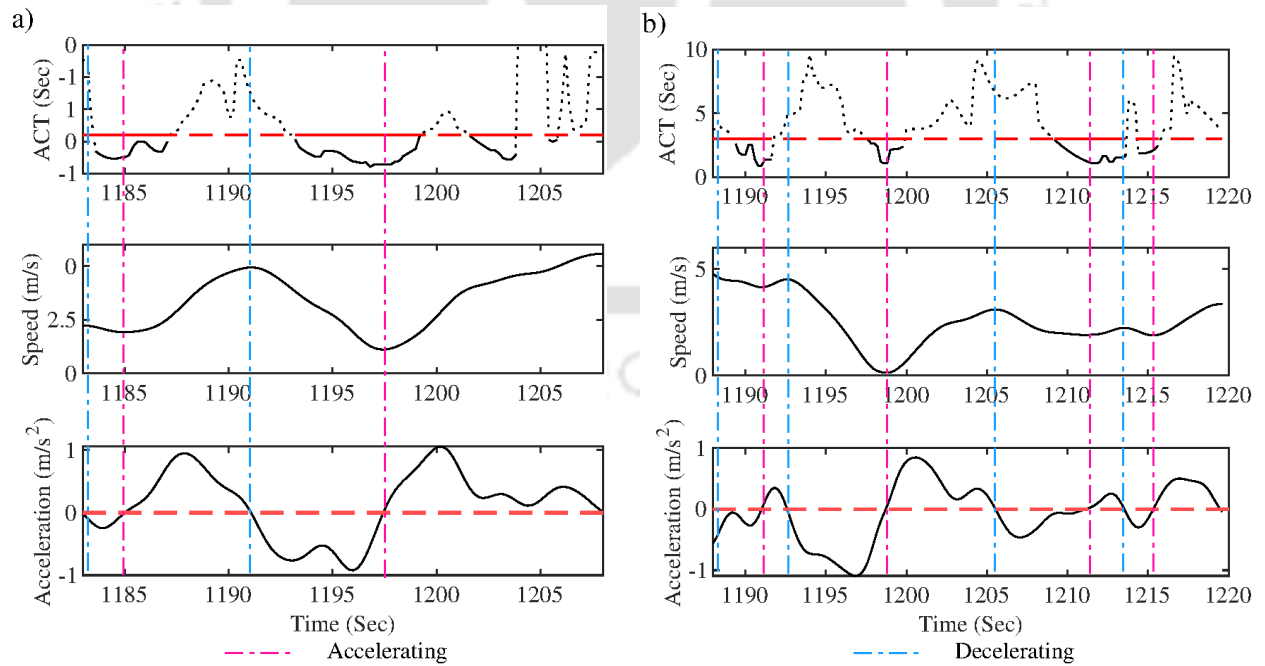


Figure 6.15: Comparison of ACT, speed, and acceleration profiles; (a) and (b) are the profiles corresponding to different PTWs

the table that, compared to PTWs, cars respond to an unsafe situation early, which indicates the risk-taking behavior of PTWs. Further, we verified whether the difference in the mean TEA for cars and PTWs is statistically significant. The t -test result shown in Table 6.6 indicates that they are statistically different.

6.2.2.2 Analyses of Crash Exposure and Severity with an Emphasis on the Importance of ACT^*

Another critical aspect of safety analysis is identifying the severity and exposure time to a potential crash scenario. Considering the ACT profile and the Eqs. 5.17&5.18, the TE-ACT and TI-ACT were estimated for PTWs and the other vehicle classes, separately. It is evident from Fig. 6.17a that, PTWs are exposed to a higher TE-ACT compared to the other vehicle classes. For example, the 75th percentile of crash exposure time for PTWs is 8sec, whereas, for other vehicle classes, it is 6sec. The comparison of TE-ACT reveals that the average duration for PTWs exposed to unsafe circumstances is relatively higher. This implies the risk-taking behavior of PTW's vis-à-vis the other vehicle types. Similarly, the TI-ACT was analyzed to assess the severity levels of PTWs and the other vehicle types. Fig. 6.17b shows that PTWs have a relatively higher TI-ACT values compared to the other vehicle classes. This indicates that the proportion of PTWs undergoing severe crashes is higher compared to the other vehicle classes. For instance, the proportion of PTWs having TI-ACT value higher than 10 s^2 is 51.7%, whereas it is 40.9% for the other vehicle classes.

Defining a threshold value for the safety indicator to distinguish safe and unsafe situations is vital for surrogate safety studies. Several studies have reported that the traffic conflict evaluation results may vary significantly depending on the preferred threshold of surrogate indicator [300]. In the past studies, the selection of threshold value was either by mapping the actual crash occurrences with the estimated one or adapting the threshold value generally considered in the literature, such as 3 sec. Besides, studies have highlighted the need to consider different thresholds for different conflicting vehicle pairs [301]. It is also essential to consider different threshold values to estimate the likelihood of different vehicle classes to involve in a collision of the same severity level or with the same crash exposure. However, the effect of threshold value on the crash exposure and severity for different vehicle classes is not well understood. Here, we investigated the variation of crash

Table 6.5: Descriptive statistics of TEA estimated for cars and PTWs

	N	Mean	SD	SEM	Median
PTW	2529	9.14	4.11	0.08	8.6
Car	1775	9.70	4.47	0.11	9
Difference		-0.55		0.13	
Overall	4304	9.37	4.27	0.07	8.8

*Standard Error of Mean (SEM) of difference is computed under the condition that equal variance is assumed.

Table 6.6: Statistical comparison of TEA corresponding to PTWs and Cars using two sample t-test

Condition	<i>t</i> -Statistic	DF	Prob> t
Equal Variance Assumed	-4.21	4302	2.63E-5
Equal Variance NOT Assumed (Welch Correction)	-4.15	3613.91	3.44E-5

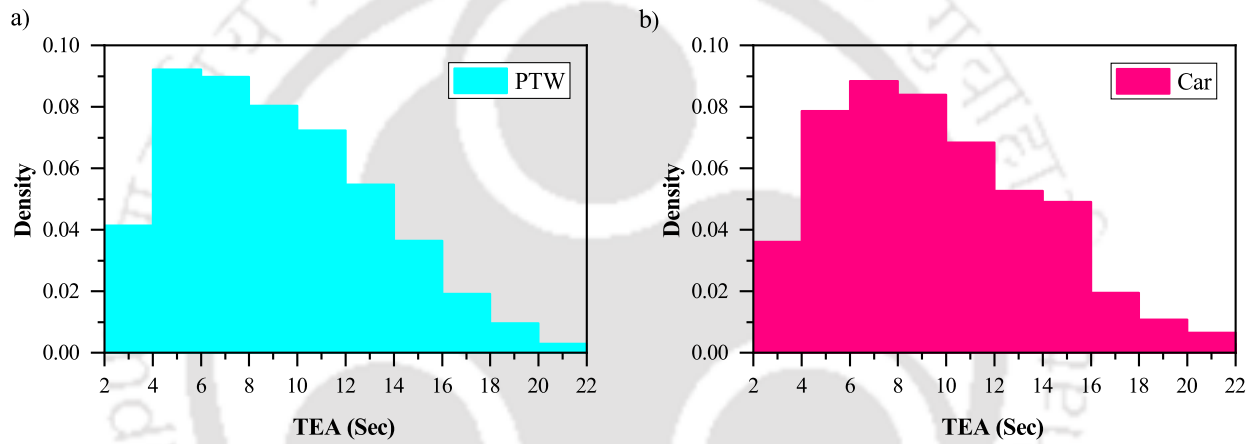


Figure 6.16: Statistical comparison of TEA for cars and PTWs using two sample t-test

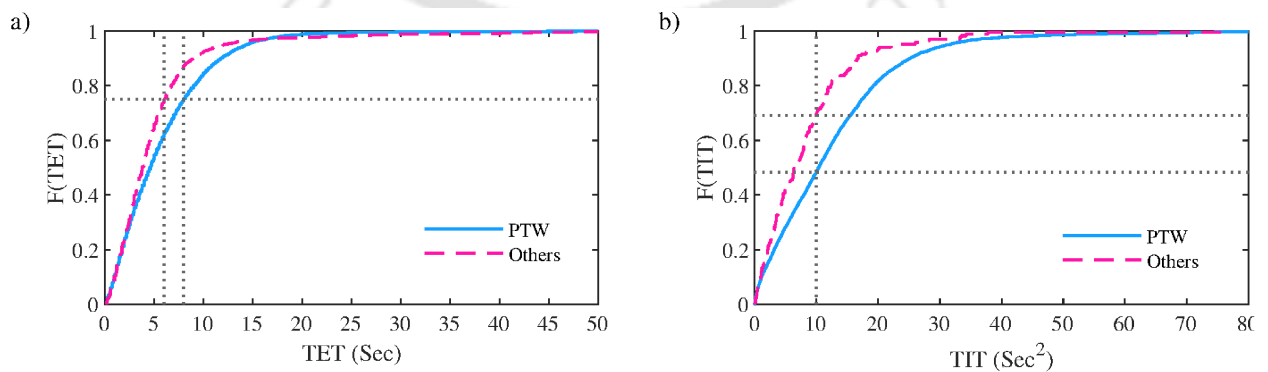


Figure 6.17: Comparison of collision risk of PTWs and other vehicle classes; a) Duration of exposition to safety critical situation (TE-ACT); b) severity of the collision (TI-ACT)

severity and exposure for different vehicle classes with the threshold value of ACT. Fig. 6.18 shows the impact of the threshold values of ACT (ACT^*) while analyzing the PTW safety compared to the other vehicle classes. Fig. 6.18a&b show the variation of average TE-ACT and TI-ACT with the ACT threshold value. At a given ACT threshold, the average TE-ACT and TI-ACT for PTWs and other vehicle classes show a clear difference. Also, this difference increases with ACT^* . For instance, it can be seen from Fig. 6.18a that the average exposure of PTWs corresponding to a threshold of 3sec was 4.6sec, whereas, for other vehicle classes, it is 3.2sec. Similarly, from Fig. 6.18b, the average severity level for PTWs at 3sec threshold is 10.5sec, whereas, for the other vehicle classes, it is 7.9sec. For comparing crash exposure and severity for different vehicle types single threshold could be used. Nevertheless, for comparing the frequency of crashes with similar severity and exposure, one needs to choose a different threshold for different vehicle classes. For example, to get the frequency of crashes at a 10sec severity level, one has to choose a threshold of 2.8sec for PTW and 3.2sec for the other vehicle classes. Similarly, for comparing crash frequency at a crash exposure level of 4sec, ACT^* should be 2.5sec for PTWs and 3.75sec for the other vehicle classes. The above analysis clearly indicates the need to consider different threshold values for different vehicle classes to estimate the crash likelihood corresponding to a particular crash exposure or severity. Considering a single threshold value for all vehicle types, on the other hand, is useful while comparing the crash severity or exposure among the vehicle types. The procedure presented here can be utilized to estimate the likelihood of a vehicle class being involved in a crash with a certain degree of severity or exposure compared to other vehicle classes. In this study, therefore, we highlight the necessity to take enough care while selecting AC threshold values considering the purpose of the analysis.

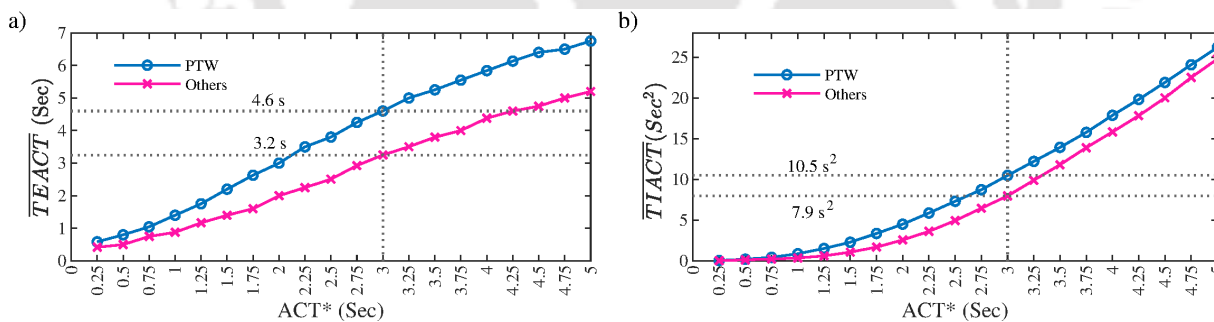


Figure 6.18: The impact of threshold value (ACT^*) on TE-ACT and TI-ACT

6.2.2.3 Comparison of Various Crash Types

As mentioned earlier, ACT enables crash type-specific traffic safety analyses as it reveals the collision type along with the safety indicator. To understand the prominent crash type among the PTWs, we evaluated the frequency of various types of collisions. Fig. 6.19a shows the relative frequency of angled, rear-end, and side-swipe collisions for PTWs. It has been observed that the side-swipe collision is the most prominent collision type for PTWs in the study area compared to the other collision types. Out of all potential collisions estimated in the study area, 57% were side-swipe. The proportions of rear-end and angled collisions were found to be 36%, and 7%,

Table 6.7: Descriptive statistics of the crash exposure and severity estimated using TTC and ACT

		N	Mean	SD	SEM*	Median
Exposure	TTC	3126	2.68	2.39	0.04	2.044
	ACT	3126	9.85	11.56	0.21	9
Severity	TTC	3126	15.31	17.48	0.31	8.81
	ACT	3126	30.65	32.01	0.57	28.93

*Standard Error of Mean (SEM) of difference is computed under the condition that equal variance is assumed.

respectively. This finding is commensurate with the crash type-wise frequency reported by several other studies [302]. The filtering behavior of PTWs in the urban environment, as reported by Clabaux et. al., [230], can be attributed to the higher sideswipe collision probability in the study area. Further, we investigated the severity and crash exposure for different crash types, and presented in Fig. 6.19b&c, respectively. It was found that the side-swipe collision has higher severity and exposure compared to other crashes in the case of PTW. Manifestly, ACT enables crash-specific proactive safety analysis, thereby allowing for the implementation of appropriate mitigation strategies for high-probability crash types.

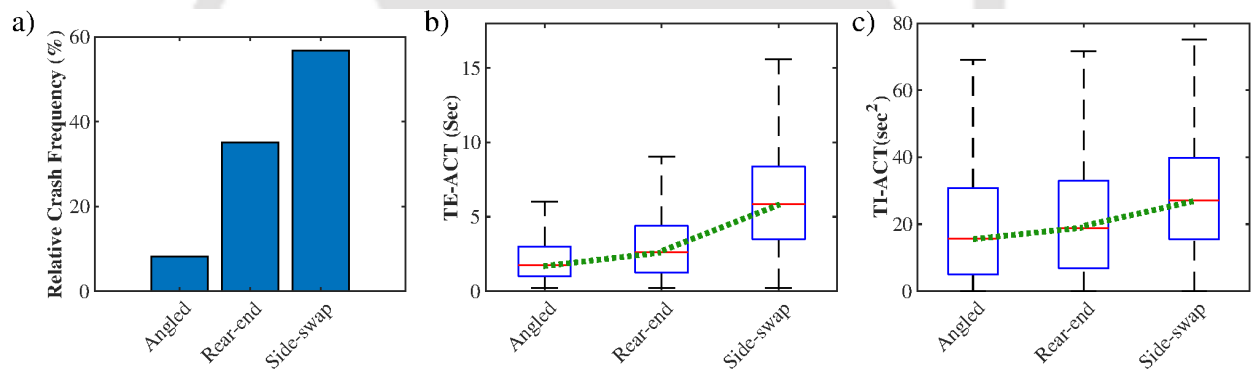


Figure 6.19: Analysis of the occurrence of different crash types for PTWs; a) Frequency of different crash types; b) Exposure time to different crash types; c) Severity of different crash types

6.2.3 Summary and Conclusion

In this section, a detailed investigation of PTW safety in an urban environment was performed. The major findings from various analyses are:

1. The likelihood of PTWs encountering unsafe situations is double compared to other vehicle types. It was observed that the PTWs are responsible for 68.4% of the total occurrences of unsafe events.
2. The response pattern of PTWs and cars towards an unsafe situation is statistically different.

The cars respond to an unsafe situation relatively early, indicating a relatively higher risk-taking behavior of PTWs.

3. The duration for which the PTWs are exposed to the unsafe situation is higher than the other vehicle classes. On average, there is a difference of 2 sec in the average exposure time between PTWs and other vehicle classes.
4. A higher proportion of the PTWs is showing higher crash severity compared to the other vehicle classes. 51.7% of PTW crashes are high severity ($TI - ACT > 10s^2$), whereas for all the other vehicle classes together, it is 40.9%.
5. For analyzing the crash likelihood of specific severity or exposure across vehicle types, or road facilities, it is essential to consider different threshold values.
6. 57% of the potential PTW collisions were found to be side-swipe collisions. Whereas the rear-end and angled collisions are 36% and 7%, respectively. This study has also found that the side-swipe collisions are more severe for PTWs. Similarly, the PTWs are exposed to side-swipe collisions for an extended period, compared to the other collision types.

6.3 Proactive Safety Assessment at 3D Road Geometries

Horizontal curves allow for a smooth transition between two tangent parts of a highway. They allow vehicles to change their travel directions gradually. However, the horizontal curves often raise safety concerns, as the driver may misperceive the geometry, thereby committing mistakes [233,234]. Such an error is more prone to occur if the horizontal curve is superimposed with a vertical alignment. Therefore, understanding the risk factors associated with the 3-Dimensional road geometry is crucial for deploying suitable safety countermeasures. Existing research has attempted to investigate the safety concerns associated with horizontal curves considering the geometry, speed, and historical crash data. However, they were limited to a relatively small-scale investigations due to the difficulties of collecting sufficient field data.

According to crash data from the various parts of the world collected at various road facilities, horizontal curves are one of the most vulnerable locations for crashes. Though there were studies that estimated the crash likelihood at curve sections, most studies depended on the historical crash data. Moreover, a crash type-specific study is limited due to the insufficient data corresponding to each crash type, therefore combine all the crash types during the safety assessment. This study investigates crash risk at horizontal curves proactively, using vehicle trajectory data collected at various horizontal curves superimposed with vertical alignment. A classified analysis of all the crashes was performed to understand the most prominent crash types on the 3D geometries. Further, the correlation between the crash likelihood and geometric variable was performed. This study also identifies the point of high crash likelihood on a curve, which is essential while designing and implementing crash mitigation measures.

6.3.1 Data

In order to understand the likelihood of different crash types at complex road geometries, the traffic videos were collected from a two-lane undivided rural highway (Shillong Bypass connecting NH-40 and NH-44) passing through hilly terrain. The highway was spanning over a 48km stretch, along which there is a maximum elevation difference of 461.93m. The traffic videos from nine curves were continuously monitored for one hour each. A fixed camera setup was used for this purpose. The vehicle trajectories were extracted from all the videos using SAVETRAX along with the other kinematic variables, such as position, speed, and acceleration. A total of 1762 vehicles were tracked from all the locations, out of which 1067 were trucks, 535 were cars, 132 were light commercial vehicles, and 21 were motorbikes. The geometric details of the each curve under study was collected from National Highway Authority of India (NHAI). Table 6.8 presents the geometric design features of the selected curves.

Apart from the trajectory and geometry data, the historical crash data was also collected from the study area. The crash reports were obtained from the Guwahati regional office of NHAI, corresponding to December 2013-September 2018 period. The crash report is a police-prepared narrative of crashes that occurred at the study location. It includes all pertinent information re-



Figure 6.20: Some of the curve sections where data is collected

Table 6.8: Geometry of the curves considered in the study

Curve no.	Speed (Kmph)	Direction	Type	R (m)	L_c (m)	e	Δ	DC	Trans. Length (m)	$L_{at}(m)$	$L_{et}(m)$	G_1 (%)	G_2 (%)	G_3 (%)	G_4 (%)	G_5 (%)	L_v	K-value	% HC in VC	$L_0(m)$
127	50	Right	Up-grade	100	33.90	0.10	45.21	17.47	45	63.68	53.49	4.75	4.75	4.75	4.75	4.75	0	0.00	0.00	0.00
171	50	Left	Up-grade	90	78.81	0.10	82.00	19.41	50	50.82	95.70	5.50	5.50	5.50	5.50	3.56	0	0.00	0.00	0.00
179	50	Left	Down-grade	200	58.70	0.06	23.98	8.74	25	37.55	14.22	-	-	-	-	-	0	0.00	0.00	0.00
												6.00	6.00	6.00	6.00	6.00				
183	50	Right	Hog	80	58.81	0.10	81.51	21.84	55	14.35	103.33	0.50	-	-	-	-	150	23.08	39.21	0.70
													1.45	2.75	4.05	6.00				
189	50	Right	Up-grade	100	56.53	0.10	58.17	17.47	45	24.98	98.68	4.60	4.60	4.60	4.60	4.60	0	0.00	0.00	0.00
221	50	Left	Down-grade	80	64.11	0.10	85.30	21.84	55	123.48	54.93	-	-	-	-	-	0	0.00	0.00	0.00
												4.80	4.80	4.80	4.80	4.80				
227	40	Right	Sag	80	54.63	0.09	67.78	21.84	40	1.00	5.00	-	-	-	-	-	60	150.00	91.06	2.08
												3.90	3.88	3.68	3.50	3.50				
238	50	Right	Down-grade	80	113.41	0.10	120.61	21.84	55	123.21	43.75	-	-	-	-	-	0	0.00	0.00	0.00
												5.00	5.00	5.00	5.00	5.00				
245	50	Left	Hog	150	83.01	0.07	43.17	11.65	30	262.09	44.72	-	-	-	-	-	60	26.67	100.00	1.68
												2.50	2.50	3.70	4.75	4.75				

R = Radius; L_c = Length of the curve; e = Superelevation; Δ = Deflection angle; DC = Degree of curve; L_{at} = Length of approach tangent; L_{et} = Length of exit tangent; G_1 = Gradient at approach tangent; G_2 = Gradient at start of curve; G_3 = Gradient at middle of the curve; G_4 = Gradient at the end of curve; G_5 = Gradient at exit tangent; L_v = Length of vertical curve; K -Value = Rate of vertical curve; %VC in HC = Percentage of vertical curve in a horizontal curve; L_0 = Distance between the mid-points of horizontal and vertical curves

garding the crash, such as the vehicle involved, the reason of the crash, the severity of the injuries, and the driver demography. The classified crash occurrence at the study location is presented in Figure 6.21a. It is evident that the most common crash types in the study area are the Run-Off-

Road, angled, side-swipe, and rear-end crashes. Nevertheless, these are the most common crash types that typically occur on horizontal curves. In the study area, most crashes were non-injury or minor (see Figure 6.21b); however, the frequency is considerably high. Moreover, as shown in Figure 6.21c, the vehicle class involved in the majority of crashes was truck. Note that the information contained in the crash report was not adequate for a detailed review of various crash events as the number of incidents for certain forms of the crash is not sufficient for statistical modeling.

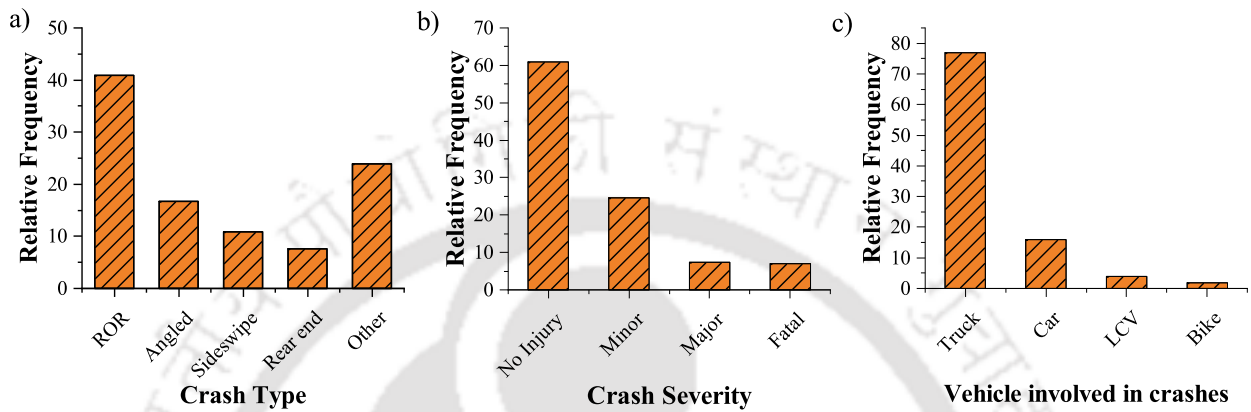


Figure 6.21: Details of the reported crash events for the study area from December 2013 to September 2018 corresponding to curves; a) Relative frequency of the occurrence of different crash types; b) Severity of crashes; c) Vehicles involved in crashes

6.3.2 Results

Considering the extracted vehicle trajectories, ACT was estimated for all conflicting vehicles using Eq.5.3. All conflicts were classified into angled, rear-end, side-swipe, and ROR crashes. Using the ACT profile, the duration of exposure to a crash as well as the severity was also investigated for each crash type, separately using equations 5.18&5.17, respectively.

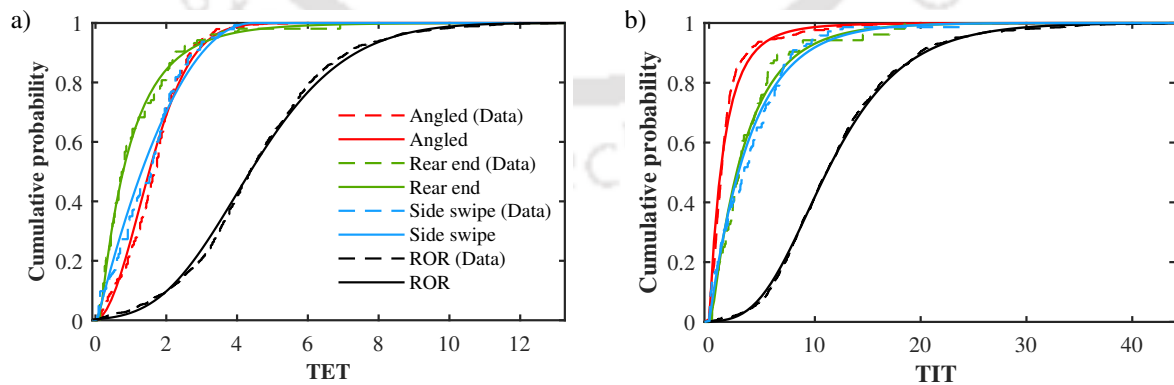


Figure 6.22: Cumulative probability density crash exposure and crash severity for different crash types; a) TET; b) TIT

Considering the estimated ACT, TE-ACT, and TI-ACT values, a comprehensive assessment of the crash frequency, duration of exposure to crash, and crash severity was performed. Figure 6.22

shows the cumulative probability density of crash exposure and crash severity of all curves together. It is evident that both the crash severity and crash exposure are significantly high in the case of ROR crashes, compared to all other crash types. It is, however, supports the crash pattern observed in the study area. The mean duration of exposure to a ROR crash was 4.56sec, whereas for rear-end, side-swipe, and angled collisions, its 1.15s, 1.48s, and 1.63s, respectively. The mean severity of the crashes was found as $12.14s^2$, $3.81s^2$, $3.99s^2$, and $1.81s^2$, respectively for ROR, rear-end, side-swipe, and angled collisions. It is therefore evident that ROR crashes are having higher probability of occurrence with a higher severity.

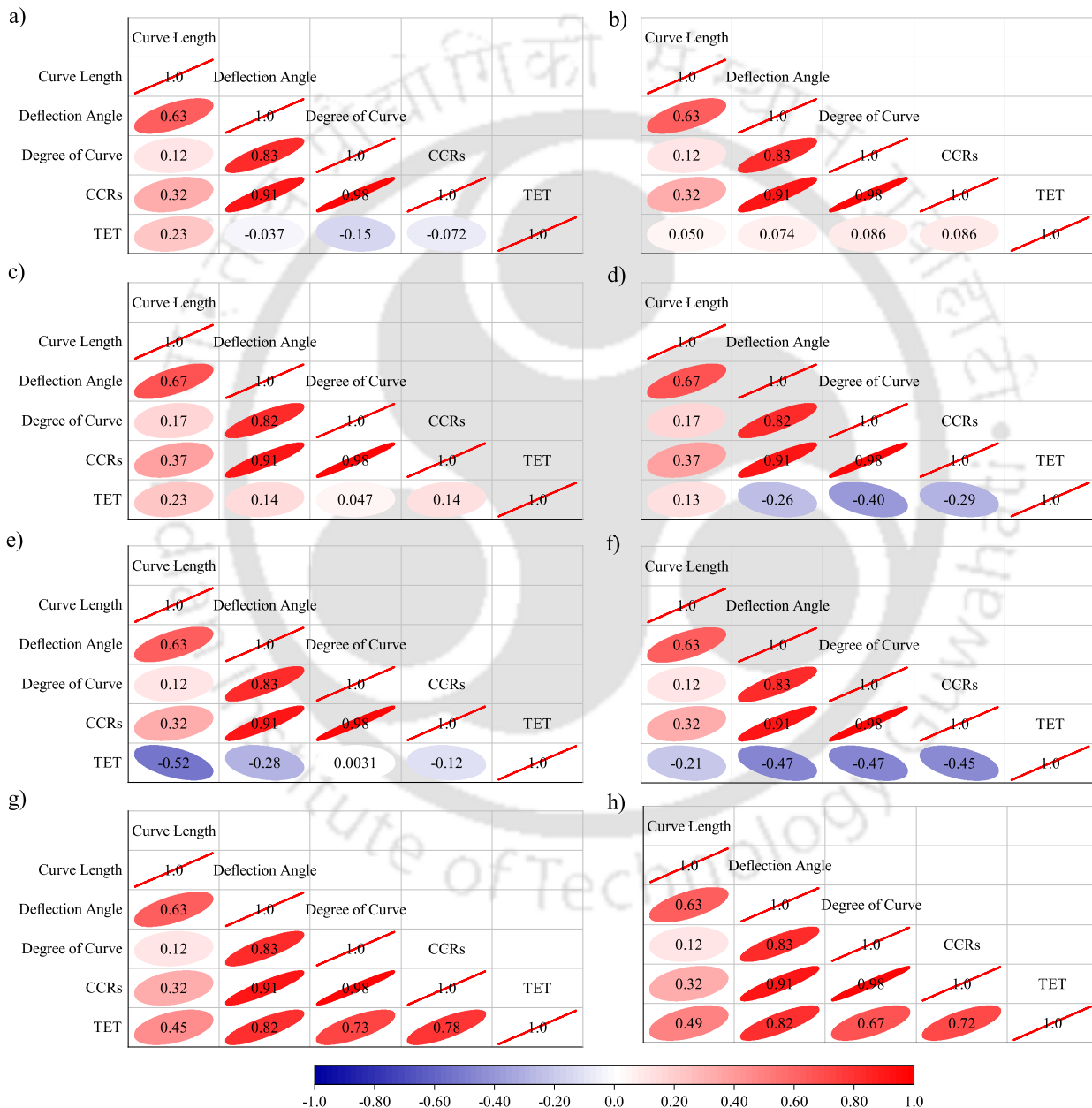


Figure 6.23: Correlation of TE-ACT and TI-ACT with geometric variables; a) Angled collision, TE-ACT; b) Angled collision, TI-ACT; c) Rear-end, TE-ACT; d) Rear-end, TI-ACT; e) Side-swipe, TE-ACT; f) Side-swipe, TI-ACT; g) ROR, TE-ACT; h) ROR, TI-ACT

A correlation analysis was performed to see how different crash types were correlated to the curve geometry. For the analysis, we have considered four geometric variables such as curve length, deflection angle, degree of curve, curvature change rate. The correlation of the crash exposure with the geometric variables and the correlation of crash severity with the geometric variables were analyzed and presented in Figure 6.23. Correlation analysis demonstrates some interesting facts regarding the relationship of curve geometry and crash type. Apparently, the curve geometric variables have dissimilar impact on different crash types. It can be seen that the severity and exposure to ROR crashes are highly correlated to all the geometric variables positively. Whereas, in the case of angled collision, no significant correlation with geometric variables was observed. The severity of a rear-end crash was found negatively correlated to the degree of the curve. The duration of exposure to side-swipe collision was found negatively correlated to the curve length, but the severity is negatively correlated to deflection angle, degree of curve, and CCRs. Evidently, geometric redesign to alleviate a particular crash type may lead to the occurrence of other crash types. Therefore, the geometric design practice should consider optimal trade-off of the occurrences of all the crash types, for safe design of the alignment. As the improvement of geometric elements is not always the only strategy to ensure safety, it would be more appropriate to add adequate protective measures (e.g., crash barriers) or installing warning systems to minimize the crashes or their severity.

The spatial configuration of crash likelihood is a crucial information while determining the location for installing effective protective measures or warning systems for safety improvements. Therefore, this study has estimated the spatial distribution of the conflict frequency along the horizontal curve. Figure 6.24 shows the spatial density of different crashes for different curve geometries. For explanation purpose, we presented two examples for spatial distribution of the density of estimated conflicts for horizontal curve with a constant grade and the other with a vertical curve. Each column in the figure represents different crash types, respectively the angular collision, rear end collision, sideswipe, and ROR crashes. The rows in the figure show different type of curves. The first and second rows represent the horizontal curve superimposed with a straight vertical grade. Whereas the third and fourth rows correspond to horizontal curve superimposed with hog/sag vertical curve.

The spatial distribution of the crash occurrences revealed that the angled collisions are more likely to occur at entry point of the curve, in the case of a horizontal curve superimposed with a straight downgrade. Similarly, for a horizontal curve with a sag curve, the vehicles at entry points of the curve have shown higher probability of angled collision. In the case of a horizontal curve superimposed with a hog curve, the vehicle leaving the summit point was found more prone to angled collision. It is evident that the likelihood of rear-end crashes is more near the middle of the curve for all the curve types. It could be due to the lower speed of the vehicle near the middle of the curve compared to the vehicles entering the curve. However, comparing to the actual crash information, the rear-end crashes is relatively less in the study area (ref. Fig. 6.25). It can be seen from the figure that the side-swipe collision is more likely to happen near the exit of a curve. This is expected due to the acceleration behavior of the vehicles while leaving a curve, and the

tendency to overtake a slow moving vehicle at the exit of the curve. The ROR crashes were more likely to occur at the entry and exit regions of a curve in the downgrade direction. This could be attributed to the higher speed of the vehicle at the entry of the curve. As shown in Figure 6.25, the comparison of the actual and estimated relative likelihood of different crash types reveals a similar trend. This indicates that the risk estimations are in line with the actual crash scenario of the study area. According to actual and expected crash events, ROR crashes are the major crash type in the study area; thus, installing adequate crash barriers or warning systems is necessary to reduce fatalities due to ROR crashes.

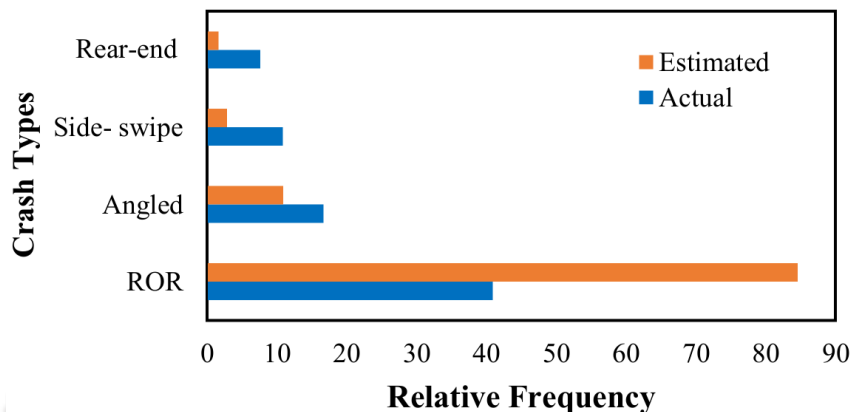


Figure 6.25: Comparison of relative likelihood of observed and estimated crashes

6.3.3 Summary

This section proactively investigated the vehicle's safety on a 3D road geometry, using the naturalistic driving data. For this purpose, this study compared the spatial distribution, the average duration of exposure to a crash situation, and the crash severity corresponding to different crash types. The correlation of different crash types with various road geometric variables was also examined. According to actual and expected crash events, ROR crashes are the major crash type in the study area; thus, installing adequate crash barriers or warning systems is necessary to reduce fatalities due to ROR crashes.

From the proactive safety analyses, the main conclusions drawn are;

1. The vehicles are more prone to ROR crashes on a horizontal curve superimposed with a vertical curve, probably due to misperception of road geometry.
2. The spatial distribution of the crashes indicates that the point of crash occurrence is not the same for all crash types.
3. The duration of exposure to crash as well as crash severity has contrasting correlation with the geometric variables for different crash types. Therefore, geometric redesign to alleviate a particular crash type may lead to the occurrence of other crash types.

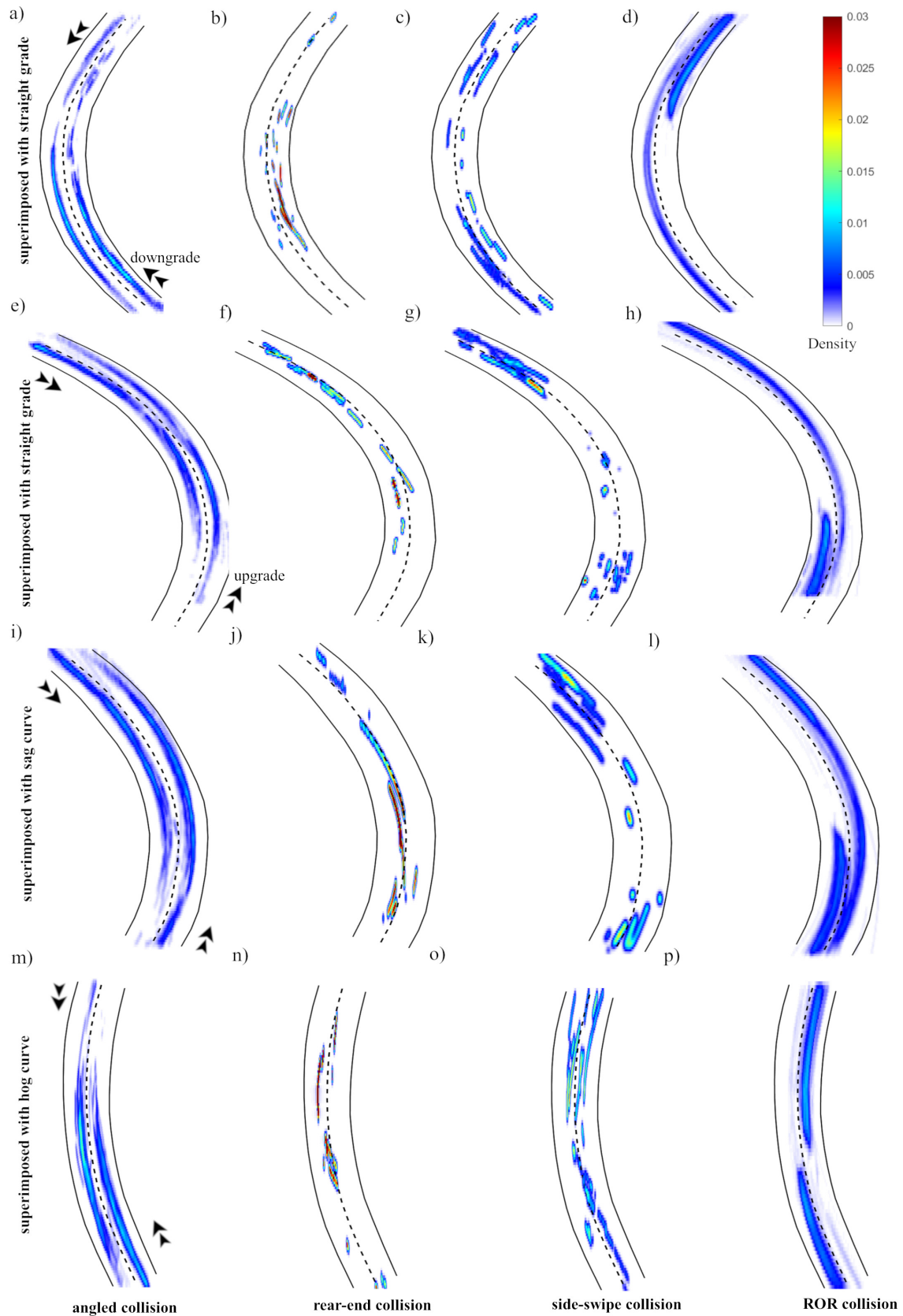


Figure 6.24: Spatial distribution of the crash frequency along the curve; first and second rows corresponding to horizontal curve superimposed with a constant grade and third and fourth corresponding to horizontal curve superimposed with vertical curve; Each column represents a crash type, such as the angled collision, rear-end collision, side-swipe, and ROR, respectively

6.4 Network-Level Proactive Safety Assessment for Identification of High-Risk Spots

As discussed earlier, the proactive safety assessment of a transportation network allows for identifying high-risk sites without relying on historical crash events. The lack of an appropriate safety indicator that can capture the transport network's overall risk was a major hurdle for such an analysis. Nevertheless, ACT brings an opportunity to perform the network-level proactive safety analysis, thereby identifying the high-risk sites. We use the recently published network-level trajectory data for performing this analysis. The details of the data and the applied pre-processing are discussed in the following subsections.

6.4.1 Data

This study utilizes the recently published network-level trajectory database, called pNEUMA data, for proactive safety analysis of a transportation network. pNEUMA is an open-source, large-scale dataset of naturalistic trajectories of half a million vehicles collected using a swarm of drones in the congested downtown area of Athens, Greece. A unique observatory of traffic stream, with a scale of the data higher than what was available until now, allows researchers from different disciplines worldwide to develop and test their models [39].

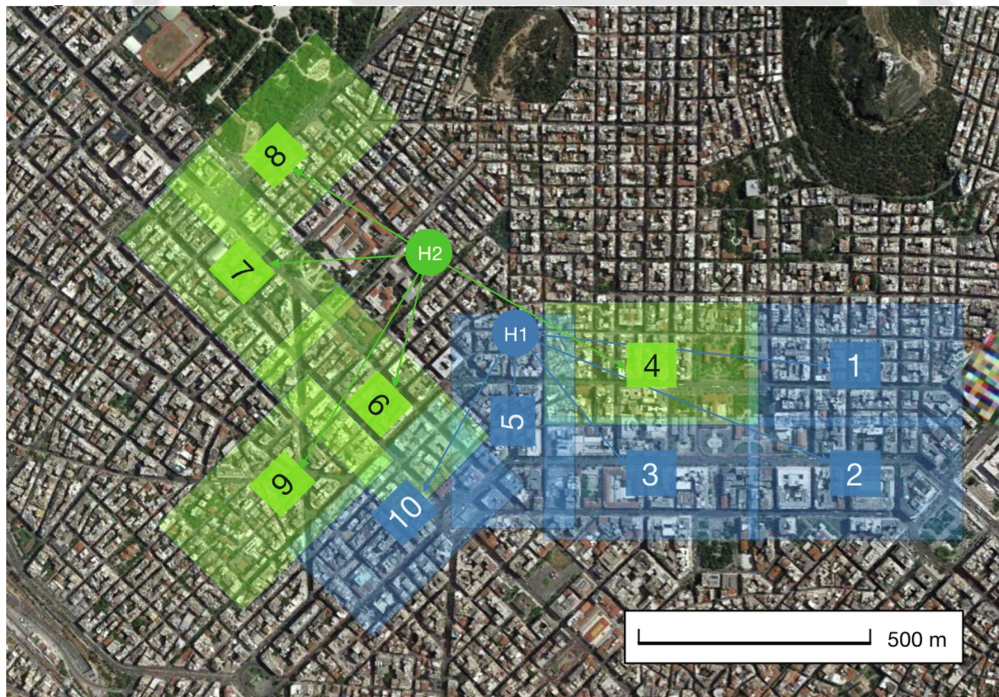


Figure 6.26: Study area and the drone coverage

A swarm of 10 drones was hovering over the study area for multiple days to record traffic streams. The total area covered by all the drones was 1.3km^2 . It covered more than 100 km-lanes of the road network, around 100 busy intersections (signalized or not), and many bus stops. Fig. 6.26 shows the study area, and the drone coverage. The pNEUMA project has outsourced the

video analysis to a start-up called DataFromSky [78]. The final data include detailed trajectories of the tracked vehicles, which are calibrated in the WGS-84 system. The time-frequency of the data was $0.04s$, which is the maximum frequency possible for the given frame rate of the video. The pNEUMA team have claimed that the detectability of the tracking algorithm was over 98.8%, and the output was manually reviewed to eradicate any false positive and false negative identifications.

To illustrate the potential of the proposed proactive safety assessment method, we have only presented the results corresponding to the sample dataset collected between 10:00-10:30 AM at the junction of Alexandras and 28is Oktovriou avenue. We have transformed the data coordinate from WGS-84 to UTM for a straight-forward analysis, taking into account the geometric zone (34 S) of the data collection site. All trajectories have been reconstructed using the methodology proposed in this study, which effectively eliminates the embedded noise. Fig. 6.27 shows the transportation network under analysis and the superimposed vehicle path data.

6.4.2 Results



Figure 6.27: pNEUMA Trajectory data

The ACT, TE-ACT, and TI-ACT were estimated using the trajectory data. Further, an ACT threshold of $3sec$ was selected to distinguish safe and unsafe situations. All instances corresponding to the unsafe scenario were projected on the spatial coordinates of the network. Later, the density of the unsafe instances across the network was plotted as a heatmap. The high-density location can be considered as the locations with a high probability of crash occurrences. Fig. 6.28 shows the result of crash frequency analysis. The heatmap of crash frequency reveals that the junctions where multiple movements intersect have a maximum number of conflict occurrences, as anticipated. Moreover, traffic congestion during the data collection period would have resulted in the highest conflict frequency near $23.733^{\circ}E$ longitude. It is evident that the traffic volume has a significant influence on the frequency of conflict situations. Therefore, consideration of only the conflict frequency may not give an actual picture of the hot spots in the study site.

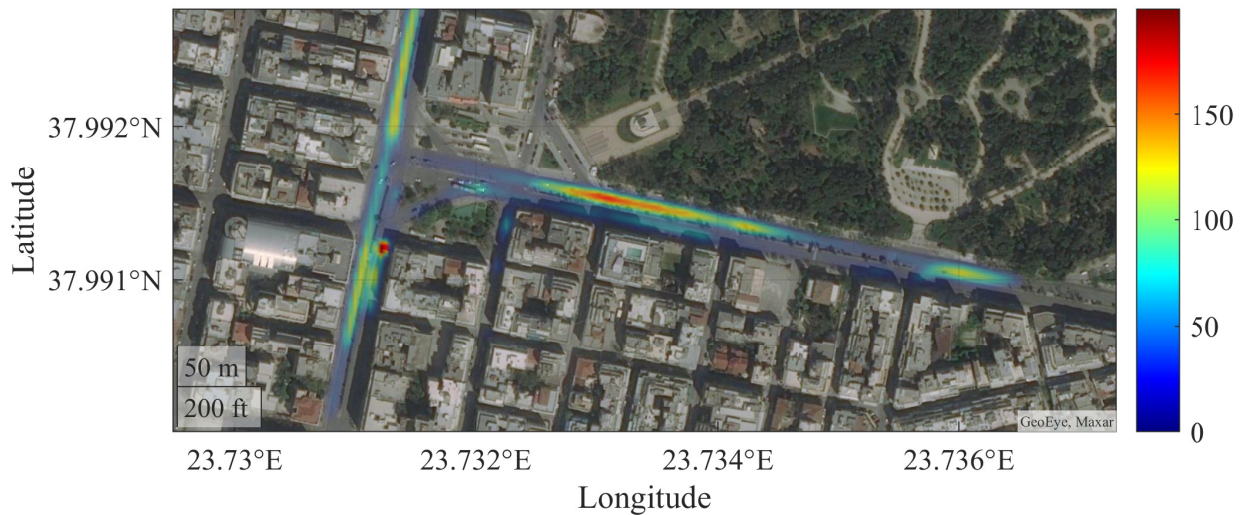


Figure 6.28: Spatial distribution of conflict frequency

In order to see how consistent is the unsafe situations, the relative exposure time to an unsafe situation was compared across different locations in the network. This analysis helps to understand the relative crash exposure across the network, on average. For this, we utilized the TE-ACT measure. All the Exposure time was projected on the spatial region, with the coordinate corresponding to the point of minimum ACT. In Fig. 6.29, the bubbles represent the unsafe situations, and the bubble size indicates the exposure time. It can be seen from the figure that some of the high crash frequency locations are having negligible crash exposure. The vehicles are showing high crash exposure near junctions, where turning movements are present.

Further, the relative spatial distribution of the crash severity was performed to see what crash severity can be expected at different locations. To identify the locations of high severity crash occurrences, the TI-ACT was used. Similar to TE-ACT, all the values of TI-ACT were projected on the network. In Fig. 6.29, the bubble color indicates the crash severity. It can be seen from the figure that the location close to junctions shows higher severity crashes.

6.4.3 Summary and Conclusions

To summarize, the suggested approach can be used to conduct a proactive network-level traffic safety evaluation. The hot-spot in the transportation network can be identified in terms of collision frequency, duration of exposure, and severity. Here, it is to be noted that, we have presented individual conflict instances and their distribution on the spatial regime. However, it is essential to integrate all the conflict scenarios for identifying the hot-spot, considering the frequency, exposure, and severity. Development of such measure is not the scope of this study, therefore kept it as the future scope.



Figure 6.29: Spatial distribution of exposure and severity of potential crashes

Chapter 7

Summary and Recommendations

7.1 Summary

This thesis put forward novel solutions to evaluate overall safety of a road transportation system. The key findings derived from the various sections of this thesis may be beneficial to related researchers and industries. In Chapter 1, we established three research objectives. In this section, we reflect on the extent to which these objectives were addressed by the findings of this thesis. Unanswered and newly identified questions are proposed for future work.

Objective 1 - Development and Standardization of Vehicle Trajectory Database: Through a rigorous literature survey presented in Chapter 2, we highlighted the demand for an advanced tool for trajectory extraction from traffic videos. In the same section, we highlighted the requirement for an appropriate methodology for removing the embedded noise in the observed trajectory data, added at several stages of data collection and extraction. These two research problems are taken as the two tasks to accomplish the first objective of this research work.

To meet the first task of the first objective, we developed a semi-automated, image-processing based vehicle trajectory extraction tool, the SAVETRAX, which allows to achieve a better quantity and quality data with minimal manual efforts. To be noted, SAVETRAX minimizes the effort required to collect the trajectory data, thereby enhancing empirical traffic flow research. Besides, SAVETRAX opens new possibilities to investigate traffic flow dynamics at complex road geometries microscopically.

The first objective's second task was to develop a trajectory reconstruction framework to eliminate the noise from the observed trajectory data. It is known that the noise in the data has a significant impact on the study result; hence, the reliability of the trajectory reconstruction technique is of prime importance. The proposed methodology is capable of handling a wide range of noises produced while collecting and extracting the trajectory data.

We believe the SAVETRAX would offer the researchers immense opportunities to perform site-specific studies by collecting a vast amount of trajectory data. The trajectory reconstruction

method would help achieving more accurate trajectory data thereby more reliable research output.

Objective 2 - Development of 2-Dimensional Surrogate Safety Indicator for General Application:

The literature review presented in Chapter 2 has revealed various limitations of the existing proactive safety analysis methods. For a comprehensive safety analysis at various road facilities, we have developed a novel surrogate safety indicator, the Anticipated Collision Time (ACT), which is capable of capturing overall safety.

ACT is independent of the collision type, road geometry, and traffic scenario. Several other safety indicators are derived from the ACT profile. The Time of Evasive Action (TEA) is one of the important contributions of this research work. This indicator allows for a more in-depth analysis of vehicle reaction patterns in the event of an unsafe situation. A thorough understanding of vehicle reaction patterns during unsafe situations is needed to design reliable control strategies for autonomous vehicles and improve collision warning systems. Two other measures, TE-ACT and TI-ACT, were extracted from the ACT profile to capture crash exposure and severity, respectively. All the measures together give a clearer picture of crash occurrences in terms of their frequency, exposure time, severity, and evasive actions.

Objective 3 - Proactive Safety Assessment at Various Elements of the Transport Network: With this objective, we aimed to demonstrate the potential of the different tools and measures developed as part of this research. We have conducted four case studies, which are: *i*) safety during an LC process; *ii*) PTW safety in an urban environment; *iii*) vehicle's safety at 3D road geometries; and, *iv*) a network-level safety assessment.

In the LC study, we microscopically examined LC dynamics after classifying LC into ILC and ULC. To that end, the present study proposed a methodology to identify and classify LC operations into ILC and ULC, automatically. Besides, the characteristic differences between ILC and ULC were explored by comparing different LC attributes. We pointed out that, intermediate interruptions during an LC maneuver could increase the crash likelihood, driver discomfort, and duration of the operation. The results from this research illustrate the need for a separate analyses of ILC and ULC dynamics in LC studies. Such a consideration would enhance the accuracy of LC simulations, collision alert systems, and human-like autonomous LCs. Besides, incorporating the finding of this study into LC models would allow more dynamic LC modeling.

The analyses of PTW safety in an urban environment have revealed that the likelihood of PTWs encountering unsafe situations is double compared to other vehicle types. The response pattern of PTWs and cars towards an unsafe situation is statistically different. The cars respond to an unsafe situation relatively early, indicating the risk-taking behavior of PTWs. The duration for which the PTWs are exposed to an unsafe situation was higher than the other vehicle classes. Similarly, the severity of the crashes involving PTWs is higher than the crashes involving any other vehicle classes. It is important to set different threshold values for different vehicle types when comparing the likelihood of a crash with specific severity or exposure time across vehicle types. This study has shown that most crashes involving PTWs are the side-swipe collisions, and have more exposure

time and severity.

The safety assessment of the vehicles traversing a horizontal curve superimposed with vertical alignment has shown that the most probable crash type on such geometry is the ROR. Corresponding to different crash types, the spatial distribution of the crash occurrences, the average duration of exposure to a crash situation, and the crash severity were compared. The correlation of different crash types with various road geometric variables was also examined. The correlation analyses have revealed that the design characteristics stimulating a crash occurrence vary across crash types. Therefore, improving the geometric design to alleviate one particular crash type might boost other crash types. According to actual and expected crash events, ROR crashes are the major crash type in the study area; thus, installing adequate crash barriers or warning systems is necessary to reduce fatalities due to such crashes.

The network-level safety assessment has shown that the hot-spot identification process needs to integrate the frequency, exposure, and severity of crash occurrences for a more efficient ranking of the hot spots. In this analysis, we highlighted that the high-frequency conflict locations might not be the location of higher severity crashes or the locations where the vehicles are exposed to a crash risk for a more extended period.

7.2 Major Contributions

The main contributions of this research work are:

1. Developed an image-based trajectory data extraction tool, which can be used for extracting traffic data from any road geometry
2. Developed a method to reconstruct the trajectory data to remove the embedded noise. The proposed approach does not cause any over/under-smoothing. Therefore, the critical incidents are not misinterpreted, which is crucial for proactive safety assessments.
3. Developed a 2D SSM, termed as ACT, which can be used independent of crash type and road facility
4. Developed criteria to perform crash-type specific, crash-severity specific, and crash-exposure specific proactive safety analyses
5. Developed a new safety indicator, Time-of-Evasive-Action (TEA), which captures the response pattern of vehicles toward an unsafe situation.
6. Developed method to automatically classify the LC operation into ILC & ULC
7. Proposed a criteria to identify the high-risk sites on a transportation network

7.3 Future Scope

This research opens tremendous future research possibilities to understand various traffic dynamics and crash generation processes. SAVETRAX and the trajectory reconstruction methods enable collecting vast trajectory data at various transportation facilities, allowing for a comprehensive understanding of various traffic dynamics.

To produce a more accurate trajectory database, further research could go into, but not limited to; *i*) a detailed investigation into the nature of the error observed in the trajectory data collected using the image processing based techniques, using different camera platforms. *ii*) Improvements to the missing data retrieval process that could handle prolonged-occlusion scenarios. *iii*) Formulation of appropriate camera stabilization technique that could minimize the heavy tailedness of the noise to some extent.

Some of the specific research problems for the future research work are:

1. Proactive PTW safety assessment for Advanced Rider Assistance Systems (ARAS) design
2. Simulator experiments to prove the realism of the proposed surrogate safety measures
3. Development of geometric design consistency evaluation method based on proactive safety measures
4. Hot-Spot Identification (HSID) on a transportation network using proactive safety measures
5. Interrupted lane change modeling for Autonomous LC applications

References

- [1] OECD, [Road injury accidents : Road fatalities](#) (2020).
URL <https://stats.oecd.org/Index.aspx?QueryId=54701>
- [2] World Health Organization, Global action plan on physical activity 2018-2030: more active people for a healthier world, World Health Organization, 2019.
- [3] C. R. Munigety, V. Vicraman, T. V. Mathew, [Semiautomated tool for extraction of microlevel traffic data from videographic survey](#), Transportation Research Record: Journal of the Transportation Research Board 2443 (1) (2014) 88–95. doi: 10.3141/2443-10.
URL <https://doi.org/10.3141/2443-10>
- [4] B. Coifman, L. Li, [A critical evaluation of the next generation simulation \(ngsim\) vehicle trajectory dataset](#), Transportation Research Part B: Methodological 105 (2017) 362 – 377. doi:<https://doi.org/10.1016/j.trb.2017.09.018>.
URL <http://www.sciencedirect.com/science/article/pii/S0191261517300838>
- [5] M. Montanino, V. Punzo, Trajectory data reconstruction and simulation-based validation against macroscopic traffic patterns, Transportation Research Part B: Methodological 80 (2015) 82–106. doi:[10.1016/j.trb.2015.06.010](https://doi.org/10.1016/j.trb.2015.06.010).
- [6] World Health Organization, [Global status report on road safety 2018](#) (2018).
URL <https://www.who.int/publications/i/item/9789241565684>
- [7] World Health Organization, About the global burden of disease (gbd) project (2002).
- [8] F. Racioppi, L. Eriksson, C. Tingvall, A. Villaveces, W. H. Organization, et al., Preventing road traffic injury: a public health perspective for europe, Tech. rep., Copenhagen: WHO Regional Office for Europe (2004).
- [9] K. Rumar, Past, present and future road safety work in ecmt (2018).
- [10] [Road Safety Performance](#), OECD, 2006. doi:[10.1787/9789282123362-ru](https://doi.org/10.1787/9789282123362-ru).
URL <https://doi.org/10.1787%2F9789282123362-ru>
- [11] C. Siddique, X. J. Ban, [State-dependent self-adaptive sampling \(SAS\) method for vehicle trajectory data](#), Transportation Research Part C: Emerging Technologies 100 (2019) 224–237. doi:[10.1016/j.trc.2019.01.018](https://doi.org/10.1016/j.trc.2019.01.018).
URL <https://doi.org/10.1016/Fj.trc.2019.01.018>
- [12] Z. Feng, Y. Zhu, [A survey on trajectory data mining: Techniques and applications](#), IEEE Access 4 (2016) 2056–2067. doi:[10.1109/access.2016.2553681](https://doi.org/10.1109/access.2016.2553681).
URL <https://doi.org/10.1109%2Faccess.2016.2553681>
- [13] S. P. Venthuruthiyil, M. Chunchu, [Vehicle path reconstruction using recursively ensembled low-pass filter \(RELP\) and adaptive tri-cubic kernel smoother](#), Transportation Research Part C: Emerging Technologies 120 (2020) 102847. doi: 10.1016/j.trc.2020.102847.
URL <https://doi.org/10.1016/j.trc.2020.102847>
- [14] D. Wang, T. Miwa, T. Morikawa, [Big trajectory data mining: A survey of methods, applications, and services](#), Sensors 20 (16) (2020) 4571. doi:[10.3390/s20164571](https://doi.org/10.3390/s20164571).
URL <https://doi.org/10.3390%2Fs20164571>
- [15] X. Zhao, Z. Wang, Z. Xu, Y. Wang, X. Li, X. Qu, [Field experiments on longitudinal characteristics of human driver behavior following an autonomous vehicle](#), Transportation Research Part C: Emerging Technologies 114 (2020) 205–224. doi:

- 10.1016/j.trc.2020.02.018.
URL <https://doi.org/10.1016/j.trc.2020.02.018>
- [16] C. Wei, R. Romano, N. Merat, Y. Wang, C. Hu, H. Taghavifar, F. Hajiseyedjavadi, E. R. Boer, [Risk-based autonomous vehicle motion control with considering human driver's behaviour](#), *Transportation Research Part C: Emerging Technologies* 107 (2019) 1–14. doi:<https://doi.org/10.1016/j.trc.2019.08.003>.
URL <http://www.sciencedirect.com/science/article/pii/S0968090X18311550>
- [17] V. Punzo, M. Montanino, [A two-level probabilistic approach for validation of stochastic traffic simulations: impact of drivers' heterogeneity models](#), *Transportation Research Part C: Emerging Technologies* 121 (2020) 102843. doi:[10.1016/j.trc.2020.102843](https://doi.org/10.1016/j.trc.2020.102843).
URL <https://doi.org/10.1016/j.trc.2020.102843>
- [18] N. C. Sarkar, A. Bhaskar, Z. Zheng, M. P. Miska, [Microscopic modelling of area-based heterogeneous traffic flow: Area selection and vehicle movement](#), *Transportation Research Part C: Emerging Technologies* 111 (2020) 373–396. doi:[10.1016/j.trc.2019.12.013](https://doi.org/10.1016/j.trc.2019.12.013).
URL <https://doi.org/10.1016/j.trc.2019.12.013>
- [19] M. Yang, X. Wang, M. Quddus, [Examining lane change gap acceptance, duration and impact using naturalistic driving data](#), *Transportation Research Part C: Emerging Technologies* 104 (2019) 317–331. doi:[10.1016/j.trc.2019.05.024](https://doi.org/10.1016/j.trc.2019.05.024).
URL <https://doi.org/10.1016/j.trc.2019.05.024>
- [20] W. Ma, L. Wan, C. Yu, L. Zou, J. Zheng, [Multi-objective optimization of traffic signals based on vehicle trajectory data at isolated intersections](#), *Transportation Research Part C: Emerging Technologies* 120 (2020) 102821. doi:[10.1016/j.trc.2020.102821](https://doi.org/10.1016/j.trc.2020.102821).
URL <https://doi.org/10.1016/j.trc.2020.102821>
- [21] L. Wang, M. Abdel-Aty, W. Ma, J. Hu, H. Zhong, [Quasi-vehicle-trajectory-based real-time safety analysis for expressways](#), *Transportation Research Part C: Emerging Technologies* 103 (2019) 30–38. doi:[10.1016/j.trc.2019.04.003](https://doi.org/10.1016/j.trc.2019.04.003).
URL <https://doi.org/10.1016/j.trc.2019.04.003>
- [22] F. Wu, R. E. Stern, S. Cui, M. L. D. Monache, R. Bhadani, M. Bunting, M. Churchill, N. Hamilton, R. Haulcy, B. Piccoli, B. Seibold, J. Sprinkle, D. B. Work, [Tracking vehicle trajectories and fuel rates in phantom traffic jams: Methodology and data](#), *Transportation Research Part C: Emerging Technologies* 99 (2019) 82–109. doi:[10.1016/j.trc.2018.12.012](https://doi.org/10.1016/j.trc.2018.12.012).
URL <https://doi.org/10.1016/j.trc.2018.12.012>
- [23] J. Liu, K. Han, X. M. Chen, G. P. Ong, [Spatial-temporal inference of urban traffic emissions based on taxi trajectories and multi-source urban data](#), *Transportation Research Part C: Emerging Technologies* 106 (2019) 145–165. doi:[10.1016/j.trc.2019.07.005](https://doi.org/10.1016/j.trc.2019.07.005).
URL <https://doi.org/10.1016/j.trc.2019.07.005>
- [24] K. Xie, K. Ozbay, H. Yang, C. Li, [Mining automatically extracted vehicle trajectory data for proactive safety analytics](#), *Transportation Research Part C: Emerging Technologies* 106 (2019) 61–72. doi:[10.1016/j.trc.2019.07.004](https://doi.org/10.1016/j.trc.2019.07.004).
URL <https://doi.org/10.1016/j.trc.2019.07.004>
- [25] Y. Guo, M. Essa, T. Sayed, M. M. Haque, S. Washington, [A comparison between simulated and field-measured conflicts for safety assessment of signalized intersections in australia](#), *Transportation Research Part C: Emerging Technologies* 101 (2019) 96–110. doi:[10.1016/j.trc.2019.02.009](https://doi.org/10.1016/j.trc.2019.02.009).
URL <https://doi.org/10.1016/j.trc.2019.02.009>
- [26] A. Y. Chen, Y.-L. Chiu, M.-H. Hsieh, P.-W. Lin, O. Angah, [Conflict analytics through the vehicle safety space in mixed traffic flows using UAV image sequences](#), *Transportation Research Part C: Emerging Technologies* 119 (2020) 102744. doi:[10.1016/j.trc.2020.102744](https://doi.org/10.1016/j.trc.2020.102744).
URL <https://doi.org/10.1016/j.trc.2020.102744>
- [27] L. Li, R. Jiang, Z. He, X. M. Chen, X. Zhou, [Trajectory data-based traffic flow studies: A revisit](#), *Transportation Research Part C: Emerging Technologies* 114 (2020) 225–240. doi:[10.1016/j.trc.2020.02.016](https://doi.org/10.1016/j.trc.2020.02.016).
URL <https://doi.org/10.1016/j.trc.2020.02.016>
- [28] S. Liu, J. Pu, Q. Luo, H. Qu, L. M. Ni, R. Krishnan, [VAIT: A visual analytics system for metropolitan transportation](#), *IEEE Transactions on Intelligent Transportation Systems* 14 (4) (2013) 1586–1596. doi:[10.1109/TITS.2013.2263225](https://doi.org/10.1109/TITS.2013.2263225).
URL <https://doi.org/10.1109/TITS.2013.2263225>

- [29] V. Punzo, M. T. Borzacchiello, B. Ciuffo, On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data, *Transportation Research Part C: Emerging Technologies* 19 (6) (2011) 1243–1262. doi:10.1016/j.trc.2010.12.007.
- [30] J. Yuan, Y. Zheng, C. Zhang, W. Xie, X. Xie, G. Sun, Y. Huang, T-drive: driving directions based on taxi trajectories, in: *Proceedings of the 18th SIGSPATIAL International conference on advances in geographic information systems*, ACM, 2010, pp. 99–108.
- [31] J. Yuan, Y. Zheng, X. Xie, G. Sun, *Driving with knowledge from the physical world*, in: *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '11*, ACM Press, 2011. doi:10.1145/2020408.2020462. URL <https://doi.org/10.1145/2020408.2020462>
- [32] V. Kanagaraj, G. Asaithambi, T. Toledo, T. C. Lee, Trajectory data and flow characteristics of mixed traffic, *Transportation Research Record* 2491 (1) (2015) 1–11. doi:10.3141/2491-01.
- [33] X. Wang, K. T. Ma, G. W. Ng, W. E. L. Grimson, Trajectory analysis and semantic region modeling using a nonparametric bayesian model, *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR (2008)* 1–8 doi:10.1109/CVPR.2008.4587718.
- [34] B. Donovan, D. Work, New york city taxi trip data (2010–2013) (2014).
- [35] J. C. Herrera, D. B. Work, R. Herring, X. J. Ban, Q. Jacobson, A. M. Bayen, Evaluation of traffic data obtained via gps-enabled mobile phones: The mobile century field experiment, *Transportation Research Part C: Emerging Technologies* 18 (4) (2010) 568–583.
- [36] M. O. Cruz, H. Macedo, A. Guimaraes, Grouping similar trajectories for carpooling purposes, in: *Intelligent Systems (BRACIS), 2015 Brazilian Conference on, IEEE, 2015*, pp. 234–239.
- [37] SHRP2, *The 2nd strategic highway research program naturalistic driving study dataset*, available from the SHRP 2 NDS InSight Data Dissemination website (2013). URL <https://insight.shrp2nds.us>
- [38] R. Krajewski, J. Bock, L. Kloecker, L. Eckstein, *The highD dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems*, in: *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, IEEE, 2018. doi:10.1109/itsc.2018.8569552. URL <https://doi.org/10.1109/itsc.2018.8569552>
- [39] E. Barmounakis, N. Geroliminis, *On the new era of urban traffic monitoring with massive drone data: The pNEUMA large-scale field experiment*, *Transportation Research Part C: Emerging Technologies* 111 (2020) 50–71. doi:10.1016/j.trc.2019.11.023. URL <https://doi.org/10.1016/j.trc.2019.11.023>
- [40] E. Hilbert, P. Rennie, W. Kneidl, *A sensor for control of arterials and networks*, *IEEE Transactions on Vehicular Technology* 29 (2) (1980) 208–215. doi:10.1109/t-vt.1980.23842. URL <https://doi.org/10.1109/t-vt.1980.23842>
- [41] P. Michalopoulos, *Vehicle detection video through image processing: the autoscope system*, *IEEE Transactions on Vehicular Technology* 40 (1) (1991) 21–29. doi:10.1109/25.69968. URL <https://doi.org/10.1109/25.69968>
- [42] S. Li, H. Yu, K. Yang, J. Zhang, R. Bin, *Video-based traffic data collection system for multiple vehicle types*, *IET Intelligent Transport Systems* 8 (2) (2014) 164–174. doi:10.1049/iet-its.2012.0099. URL <https://doi.org/10.1049/iet-its.2012.0099>
- [43] V. Kastinaki, M. Zervakis, K. Kalaitzakis, A survey of video processing techniques for traffic applications, *Image and Vision Computing* 21 (4) (2003) 359–381. doi:10.1016/s0262-8856(03)00004-0.
- [44] Q. Wang, S. Huang, H. bin Zhang, Q. Yang, J. jun Zhang, *Traffic incident detection system based on video image processing*, *Journal of Computer Applications* 28 (7) (2008) 1886–1889. doi:10.3724/sp.j.1087.2008.01886. URL <https://doi.org/10.3724/sp.j.1087.2008.01886>
- [45] J. Oh, J. Min, M. Kim, H. Cho, *Development of an automatic traffic conflict detection system based on image tracking technology*, *Transportation Research Record: Journal of the Transportation Research Board* 2129 (1) (2009) 45–54. doi:

- 10.3141/2129-06.
URL <https://doi.org/10.3141/2129-06>
- [46] T. T. Nguyen, P. Krishnakumari, S. C. Calvert, H. L. Vu, H. van Lint, [Feature extraction and clustering analysis of highway congestion](#), *Transportation Research Part C: Emerging Technologies* 100 (2019) 238–258. doi:10.1016/j.trc.2019.01.017.
URL <https://doi.org/10.1016/j.trc.2019.01.017>
- [47] P. Shen, [Flow detection based on traffic video image processing](#), *Journal of Multimedia* 8 (5) (oct 2013). doi:10.4304/jmm.8.5.519-526.
URL <https://doi.org/10.4304/jmm.8.5.519-526>
- [48] G. Zhang, R. P. Avery, Y. Wang, [Video-based vehicle detection and classification system for real-time traffic data collection using uncalibrated video cameras](#), *Transportation Research Record: Journal of the Transportation Research Board* 1993 (1) (2007) 138–147. doi:10.3141/1993-19.
URL <https://doi.org/10.3141/1993-19>
- [49] P. Suvin, C. Mallikarjuna, Modified generalized definitions for the traffic flow characteristics under heterogeneous, no-lane disciplined traffic streams, *Transportation Research Procedia* 34 (2018) 75–82. doi:10.1016/j.trpro.2018.11.016.
URL <https://doi.org/10.1016/j.trpro.2018.11.016>
- [50] M. Guerrieri, R. Mauro, T. Tollazzi, [Turbo-roundabout: Case study of driver behavior and kinematic parameters of light and heavy vehicles](#), *Journal of Transportation Engineering, Part A: Systems* 145 (6) (2019) 05019002. doi:10.1061/jtepbs.0000241.
URL <https://doi.org/10.1061/jtepbs.0000241>
- [51] M. Guerrieri, R. Mauro, G. Parla, T. Tollazzi, [Analysis of kinematic parameters and driver behavior at turbo roundabouts](#), *Journal of Transportation Engineering, Part A: Systems* 144 (6) (2018) 04018020. doi:10.1061/jtepbs.0000129.
URL <https://doi.org/10.1061/jtepbs.0000129>
- [52] T. D. Chu, T. Miwa, T. Morikawa, [Discrete choice models for gap acceptance at urban expressway merge sections considering safety, road geometry, and traffic conditions](#), *Journal of Transportation Engineering, Part A: Systems* 143 (7) (2017) 04017025. doi:10.1061/jtepbs.0000053.
URL <https://doi.org/10.1061/jtepbs.0000053>
- [53] L. Xing, J. He, Y. Li, Y. Wu, J. Yuan, X. Gu, [Comparison of different models for evaluating vehicle collision risks at upstream diverging area of toll plaza](#), *Accident Analysis & Prevention* 135 (2020) 105343. doi:10.1016/j.aap.2019.105343.
URL <https://doi.org/10.1016/j.aap.2019.105343>
- [54] A. Kathuria, P. Vedagiri, [Evaluating pedestrian vehicle interaction dynamics at un-signalized intersections: A proactive approach for safety analysis](#), *Accident Analysis & Prevention* 134 (2020) 105316. doi:10.1016/j.aap.2019.105316.
URL <https://doi.org/10.1016/j.aap.2019.105316>
- [55] C. Jiang, R. Qiu, T. Fu, L. Fu, B. Xiong, Z. Lu, [Impact of right-turn channelization on pedestrian safety at signalized intersections](#), *Accident Analysis & Prevention* 136 (2020) 105399. doi:10.1016/j.aap.2019.105399.
URL <https://doi.org/10.1016/j.aap.2019.105399>
- [56] C. Wang, C. Xu, Y. Dai, [A crash prediction method based on bivariate extreme value theory and video-based vehicle trajectory data](#), *Accident Analysis & Prevention* 123 (2019) 365–373. doi:10.1016/j.aap.2018.12.013.
URL <https://doi.org/10.1016/j.aap.2018.12.013>
- [57] L. Xing, J. He, M. Abdel-Aty, Q. Cai, Y. Li, O. Zheng, [Examining traffic conflicts of up stream toll plaza area using vehicles' trajectory data](#), *Accident Analysis & Prevention* 125 (2019) 174–187. doi:10.1016/j.aap.2019.01.034.
URL <https://doi.org/10.1016/j.aap.2019.01.034>
- [58] Y. Guo, T. Sayed, M. H. Zaki, [Exploring evasive action-based indicators for PTW conflicts in shared traffic facility environments](#), *Journal of Transportation Engineering, Part A: Systems* 144 (11) (2018) 04018065. doi:10.1061/jtepbs.0000190.
URL <https://doi.org/10.1061/jtepbs.0000190>
- [59] E.-J. Kim, H.-C. Park, S.-W. Ham, S.-Y. Kho, D.-K. Kim, [Extracting vehicle trajectories using unmanned aerial vehicles in congested traffic conditions](#), *Journal of Advanced Transportation* 2019 (2019) 1–16. doi:10.1155/2019/9060797.
URL <https://doi.org/10.1155/2019/9060797>

- [60] P. St-Aubin, N. Saunier, L. Miranda-Moreno, [Large-scale automated proactive road safety analysis using video data](#), Transportation Research Part C: Emerging Technologies 58 (2015) 363–379. doi:10.1016/j.trc.2015.04.007. URL <https://doi.org/10.1016/j.trc.2015.04.007>
- [61] A. van Beinum, E. Broekman, H. Farah, W. Schakel, F. Wegman, S. Hoogendoorn, [Critical assessment of microscopic simulation models for simulating turbulence around motorway ramps](#), Journal of Transportation Engineering, Part A: Systems 146 (2) (2020) 04019066. doi:10.1061/jtepbs.0000296. URL <https://doi.org/10.1061/jtepbs.0000296>
- [62] J. C. Herrera, D. B. Work, R. Herring, X. J. Ban, Q. Jacobson, A. M. Bayen, [Evaluation of traffic data obtained via GPS-enabled mobile phones: The mobile century field experiment](#), Transportation Research Part C: Emerging Technologies 18 (4) (2010) 568–583. doi:10.1016/j.trc.2009.10.006. URL <https://doi.org/10.1016/j.trc.2009.10.006>
- [63] V. Kovvali, V. Alexiadis, L. Zhang, P. Length, [Video-Based Vehicle Trajectory Data Collection Video-Based Vehicle Trajectory Data Collection](#), TRB 86th Annual Meeting Compendium of Papers CD-ROM (2007) 1–18. URL <https://trid.trb.org/view/801154>
- [64] S. P. Venthuruthiyil, M. Chunchu, Trajectory reconstruction using locally weighted regression: a new methodology to identify the optimum window size and polynomial order, Transportmetrica A: Transport Science 14 (10) (2018) 881–900. doi:10.1080/23249935.2018.1449032.
- [65] H. Dinh, H. Tang, A video processing system for automated traffic data collection of gap size for roundabouts, in: 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), 2017, pp. 1–4. doi:10.1109/CCECE.2017.7946597.
- [66] B. Coifman, D. Beymer, P. McLauchlan, J. Malik, A real-time computer vision system for vehicle tracking and traffic surveillance, Transportation Research Part C: Emerging Technologies 6 (4) (1998) 271–288. doi:10.1016/S0968-090X(98)00019-9.
- [67] W. Zhang, G. Jordan, V. Livshits, [Generating a vehicle trajectory database from time-lapse aerial photography](#), Transportation Research Record: Journal of the Transportation Research Board 2594 (1) (2016) 148–158. doi:10.3141/2594-18. URL <https://doi.org/10.3141/2594-18>
- [68] E. N. Barmounakis, E. I. Vlahogianni, J. C. Golias, A. Babinec, How accurate are small drones for measuring microscopic traffic parameters? (jul 2019). doi:10.1080/19427867.2017.1354433.
- [69] M. A. Khan, W. Ectors, T. Bellemans, Y. Ruichek, A. U. H. Yasar, D. Janssens, G. Wets, Unmanned Aerial Vehicle-based Traffic Analysis: A Case Study to Analyze Traffic Streams at Urban Roundabouts, in: Procedia Computer Science, Vol. 130, 2018, pp. 636–643. doi:10.1016/j.procs.2018.04.114.
- [70] J. Lee, Z. Zhong, K. Kim, B. Dimitrijevic, B. Du, S. Gutesa, Examining the applicability of small quadcopter drone for traffic surveillance and roadway incident monitoring, in: 94th Transportation Research Board Annual Meeting, Vol. 134, Washington, DC, 2015, pp. 635–646.
- [71] X. Gu, M. Abdel-Aty, Q. Xiang, Q. Cai, J. Yuan, [Utilizing UAV video data for in-depth analysis of drivers' crash risk at interchange merging areas](#), Accident Analysis & Prevention 123 (2019) 159–169. doi:10.1016/j.aap.2018.11.010. URL <https://doi.org/10.1016/j.aap.2018.11.010>
- [72] A. van Beinum, F. Wegman, [Design guidelines for turbulence in traffic on dutch motorways](#), Accident Analysis & Prevention 132 (2019) 105285. doi:10.1016/j.aap.2019.105285. URL <https://doi.org/10.1016/j.aap.2019.105285>
- [73] A. van Beinum, H. Farah, F. Wegman, S. Hoogendoorn, [Driving behaviour at motorway ramps and weaving segments based on empirical trajectory data](#), Transportation Research Part C: Emerging Technologies 92 (2018) 426–441. doi:10.1016/j.trc.2018.05.018. URL <https://doi.org/10.1016/j.trc.2018.05.018>
- [74] S. Kaufmann, B. S. Kerner, H. Rehborn, M. Koller, S. L. Klenov, Aerial observations of moving synchronized flow patterns in over-saturated city traffic, Transportation Research Part C: Emerging Technologies 86 (2018) 393–406. doi:10.1016/j.trc.2017.11.024.

- [75] M. A. Khan, W. Ectors, T. Bellemans, D. Janssens, G. Wets, **Unmanned aerial vehicle-based traffic analysis: Methodological framework for automated multivehicle trajectory extraction**, *Transportation Research Record* 2626 (1) (2017) 25–33. doi:10.3141/2626-04.
URL <http://journals.sagepub.com/doi/10.3141/2626-04>
- [76] TRAZER, **Traffic Analysis & Classification - TRAZER** | Kritikal Solutions.
URL <https://kritikalsolutions.com/products/trazer/>
- [77] intuvisiotech, **Traffic Video Analytics**.
URL https://www.intuvisiotech.com/intuvisionVA{_}solutions/intuvisionVA{_}traffic
- [78] DataFromSky, **DataFromSky Deep traffic video analysis - DataFromSky**.
URL <https://datafromsky.com/>
- [79] Traffic Survey Solutions, **Traffic Survey Solutions**.
URL <https://deltasoftwareolutions.net/>
- [80] TrafficVision, **TrafficVision**.
URL <http://www.trafficvision.com/>
- [81] GoodVision, **GoodVision Video Insights - Advanced Traffic Analytics Platform**.
URL <https://goodvisionlive.com/>
- [82] N. Noy, A. Noy, **Let go of your data**, *Nature Materials* 19 (1) (2019) 128–128. doi:10.1038/s41563-019-0539-5.
URL <https://doi.org/10.1038/s41563-019-0539-5>
- [83] J. Boughner, **A lack of funding**, *Nature* 460 (7255) (2009) 655–655. doi:10.1038/nj7255-655b.
URL <https://doi.org/10.1038/nj7255-655b>
- [84] C. Mallikarjuna, A. Phanindra, K. R. Rao, **Traffic data collection under mixed traffic conditions using video image processing**, *Journal of Transportation Engineering* 135 (4) (2009) 174–182. doi:10.1061/(ASCE)0733-947X(2009)135:4(174).
- [85] V. Alexiadis, J. Colyar, J. Halkias, R. Hranac, G. McHale, **The next generation simulation program**, *Institute of Transportation Engineers. ITE Journal* 74 (8) (2004) 22.
- [86] D. Pal, M. Chunchu, **Smoothing of vehicular trajectories under heterogeneous traffic conditions to extract microscopic data**, *Canadian Journal of Civil Engineering* 45 (6) (2018) 435–445. doi:10.1139/cjce-2017-0452.
- [87] G. Guido, V. Gallelli, D. Rogano, A. Vitale, **Evaluating the accuracy of vehicle tracking data obtained from Unmanned Aerial Vehicles**, *International Journal of Transportation Science and Technology* 5 (3) (2016) 136–151. doi:10.1016/j.ijtst.2016.12.001.
- [88] K. Kanistras, G. Martins, M. J. Rutherford, K. P. Valavanis, **Survey of unmanned aerial vehicles (UAVs) for traffic monitoring**, in: *Handbook of Unmanned Aerial Vehicles*, Springer Netherlands, 2014, pp. 2643–2666. doi:10.1007/978-90-481-9707-1_122.
- [89] G. S. K. Fung, **Camera calibration from road lane markings**, *Optical Engineering* 42 (10) (2003) 2967. doi:10.1117/1.1606458.
URL <https://doi.org/10.1117/1.1606458>
- [90] X. C. He, **New method for overcoming ill-conditioning in vanishing-point-based camera calibration**, *Optical Engineering* 46 (3) (2007) 037202. doi:10.1117/1.2714991.
URL <https://doi.org/10.1117/1.2714991>
- [91] R. Tsai, **A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf TV cameras and lenses**, *IEEE Journal on Robotics and Automation* 3 (4) (1987) 323–344. doi:10.1109/jra.1987.1087109.
URL <https://doi.org/10.1109/jra.1987.1087109>
- [92] T. Schoepflin, D. Dailey, **Dynamic camera calibration of roadside traffic management cameras**, in: *Proceedings. The IEEE 5th International Conference on Intelligent Transportation Systems, IEEE, 2002*. doi:10.1109/itsc.2002.1041183.
URL <https://doi.org/10.1109/itsc.2002.1041183>

- [93] L. L. Wang, W.-H. Tsai, [Camera calibration by vanishing lines for 3-d computer vision](#), IEEE Transactions on Pattern Analysis and Machine Intelligence 13 (4) (1991) 370–376. doi:10.1109/34.88572. URL <https://doi.org/10.1109/34.88572>
- [94] T.-C. Lee, J. W. Polak, M. G. H. Bell, [Trajectory Extractor User Manual Version 1.0](#) (2008). URL <http://myweb.ncku.edu.tw/~jtclee/bike/iccts01155.pdf>
- [95] H. Wei, C. Feng, E. Meyer, J. Lee, [Video-capture-based approach to extract multiple vehicular trajectory data for traffic modeling](#), Journal of Transportation Engineering 131 (7) (2005) 496–505. doi:10.1061/(asce)0733-947x(2005)131:7(496). URL [https://doi.org/10.1061/\(ASCE\)0733-947X\(2005\)131:7\(496\)](https://doi.org/10.1061/(ASCE)0733-947X(2005)131:7(496))
- [96] J. Apeltauer, A. Babinec, D. Herman, T. Apeltauer, [Automatic vehicle trajectory extraction for traffic analysis from aerial video data](#), ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-3/W2 (2015) 9–15. doi:10.5194/isprsarchives-xl-3-w2-9-2015. URL <https://doi.org/10.5194/isprsarchives-xl-3-w2-9-2015>
- [97] A. Babinec, Automatic vehicle trajectory extraction from aerial video data, 2015.
- [98] A. Kesting, M. Treiber, [Calibrating car-following models by using trajectory data](#), Transportation Research Record: Journal of the Transportation Research Board 2088 (1) (2008) 148–156. doi:10.3141/2088-16. URL <https://doi.org/10.3141/2088-16>
- [99] C. Thiemann, M. Treiber, A. Kesting, [Estimating acceleration and lane-changing dynamics from next generation simulation trajectory data](#), Transportation Research Record: Journal of the Transportation Research Board 2088 (1) (2008) 90–101. doi:10.3141/2088-10. URL <https://doi.org/10.3141/2088-10>
- [100] M. Montanino, V. Punzo, [Making NGSIM data usable for studies on traffic flow theory](#), Transportation Research Record: Journal of the Transportation Research Board 2390 (1) (2013) 99–111. doi:10.3141/2390-11. URL <https://doi.org/10.3141/2390-11>
- [101] V. Punzo, D. J. Formisano, V. Torrieri, Nonstationary kalman filter for estimation of accurate and consistent car-following data, Transportation Research Record: Journal of the Transportation Research Board 1934 (1) (2005) 2–12. doi:10.1177/0361198105193400101.
- [102] T. Toledo, H. N. Koutsopoulos, K. I. Ahmed, [Estimation of vehicle trajectories with locally weighted regression](#), Transportation Research Record: Journal of the Transportation Research Board 1999 (1) (2007) 161–169. doi:10.3141/1999-17. URL <https://doi.org/10.3141/1999-17>
- [103] M. Treiber, A. Kesting, Microscopic calibration and validation of car-following models – a systematic approach, Procedia - Social and Behavioral Sciences 80 (2013) 922–939. doi:10.1016/j.sbspro.2013.05.050.
- [104] V. Punzo, M. T. Borzacchiello, B. Ciuffo, Estimation of vehicle trajectories from observed discrete positions and next-generation simulation program (ngsim) data, in: TRB 2009 Annual Meeting, Citeseer, 2009, pp. 1–17.
- [105] Y. L. Murphey, R. Milton, L. Kiliaris, Driver's style classification using jerk analysis, in: 2009 IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems, 2009, pp. 23–28. doi:10.1109/CIVVS.2009.4938719.
- [106] S. H. Hamdar, H. S. Mahmassani, Driver car-following behavior: From discrete event process to continuous set of episodes, Tech. rep. (2008).
- [107] T. Hiraoka, T. Kunitatsu, O. Nishihara, H. Kumamoto, Modeling of driver following behavior based on minimum-jerk theory, in: Proc. 12th World Congress ITS, 2005.
- [108] F. Marczak, C. Buisson, New filtering method for trajectory measurement errors and its comparison with existing methods, Transportation Research Record: Journal of the Transportation Research Board 2315 (1) (2012) 35–46. doi:10.3141/2315-04.
- [109] M. Rafati Fard, A. Shariat Mohaymany, M. Shahri, A new methodology for vehicle trajectory reconstruction based on wavelet analysis, Transportation Research Part C: Emerging Technologies 74 (2017) 150–167. doi:10.1016/j.trc.2016.11.010.
- [110] W. Xin, J. Hourdos, P. G. Michalopoulos, Vehicle trajectory collection and processing methodology and its implementation, Tech. rep. (2008).

- [111] S. Hoogendoorn, R. G. Hoogendoorn, W. Daamen, Wiedemann revisited, *Transportation Research Record: Journal of the Transportation Research Board* 2260 (1) (2011) 152–162. doi:10.3141/2260-17.
- [112] M. A. Khan, W. Ectors, T. Bellemans, D. Janssens, G. Wets, [UAV-based traffic analysis: A universal guiding framework based on literature survey](#), *Transportation Research Procedia* 22 (2017) 541–550. doi:10.1016/j.trpro.2017.03.043.
URL <https://doi.org/10.1016/j.trpro.2017.03.043>
- [113] K. M. Lynch, N. Marchuk, M. L. Elwin, [Chapter 22 - digital signal processing](#), in: K. M. Lynch, N. Marchuk, M. L. Elwin (Eds.), *Embedded Computing and Mechatronics with the PIC32*, Newnes, Oxford, 2016, pp. 341–374. doi:<https://doi.org/10.1016/B978-0-12-420165-1.00022-6>.
URL <https://www.sciencedirect.com/science/article/pii/B9780124201651000226>
- [114] L. B. Meuleners, A. H. Lee, C. Haworth, [Emergency presentations by vulnerable road users: implications for injury prevention](#), *Injury Prevention* 12 (1) (2006) 12–14. doi:10.1136/ip.2005.010389.
URL <https://doi.org/10.1136/ip.2005.010389>
- [115] C. Johnsson, A. Laureshyn, T. D. Ceunynck, [In search of surrogate safety indicators for vulnerable road users: a review of surrogate safety indicators](#), *Transport Reviews* 38 (6) (2018) 765–785. doi:10.1080/01441647.2018.1442888.
URL <https://doi.org/10.1080/01441647.2018.1442888>
- [116] M. Li, Z. Li, C. Xu, T. Liu, [Short-term prediction of safety and operation impacts of lane changes in oscillations with empirical vehicle trajectories](#), *Accident Analysis & Prevention* 135 (2020) 105345. doi:10.1016/j.aap.2019.105345.
URL <https://doi.org/10.1016/j.aap.2019.105345>
- [117] U. Durrani, C. Lee, [Calibration and validation of psychophysical car-following model using driver's action points and perception thresholds](#), *Journal of Transportation Engineering, Part A: Systems* 145 (9) (2019) 04019039. doi:10.1061/jtepbs.0000264.
URL <https://doi.org/10.1061/jtepbs.0000264>
- [118] H. Mohammed, A. Y. Bigazzi, T. Sayed, [Characterization of bicycle following and overtaking maneuvers on cycling paths](#), *Transportation Research Part C: Emerging Technologies* 98 (2019) 139–151. doi:10.1016/j.trc.2018.11.012.
URL <https://doi.org/10.1016/j.trc.2018.11.012>
- [119] J. Tian, H. Zhang, M. Treiber, R. Jiang, Z.-Y. Gao, B. Jia, [On the role of speed adaptation and spacing indifference in traffic instability: Evidence from car-following experiments and its stochastic model](#), *Transportation Research Part B: Methodological* 129 (2019) 334–350. doi:10.1016/j.trb.2019.09.014.
URL <https://doi.org/10.1016/j.trb.2019.09.014>
- [120] Y. Wei, K. Li, K. Tang, [Trajectory-based identification of critical instantaneous decision events at mixed-flow signalized intersections](#), *Accident Analysis & Prevention* 123 (2019) 324–335. doi:10.1016/j.aap.2018.11.019.
URL <https://doi.org/10.1016/j.aap.2018.11.019>
- [121] M. Saifuzzaman, Z. Zheng, M. M. Haque, S. Washington, [Understanding the mechanism of traffic hysteresis and traffic oscillations through the change in task difficulty level](#), *Transportation Research Part B: Methodological* 105 (2017) 523–538. doi:10.1016/j.trb.2017.09.023.
URL <https://doi.org/10.1016/j.trb.2017.09.023>
- [122] M. Zhou, X. Qu, X. Li, [A recurrent neural network based microscopic car following model to predict traffic oscillation](#), *Transportation Research Part C: Emerging Technologies* 84 (2017) 245–264. doi:10.1016/j.trc.2017.08.027.
URL <https://doi.org/10.1016/j.trc.2017.08.027>
- [123] M. H. Zaki, T. Sayed, K. Shaaban, [Use of drivers' jerk profiles in computer vision-based traffic safety evaluations](#), *Transportation Research Record* 2434 (1) (2014) 103–112. arXiv:<https://doi.org/10.3141/2434-13>, doi:10.3141/2434-13.
URL <https://doi.org/10.3141/2434-13>
- [124] S. Ahn, S. Vadlamani, J. Laval, [A method to account for non-steady state conditions in measuring traffic hysteresis](#), *Transportation Research Part C: Emerging Technologies* 34 (2013) 138–147. doi:10.1016/j.trc.2011.05.020.
URL <https://doi.org/10.1016/j.trc.2011.05.020>
- [125] O. Bagdadi, A. Várhelyi, [Jerky driving—an indicator of accident proneness?](#), *Accident Analysis & Prevention* 43 (4) (2011) 1359–1363. doi:10.1016/j.aap.2011.02.009.
URL <https://doi.org/10.1016/j.aap.2011.02.009>

- [126] A. Bovik, [The Essential Guide to Image Processing](#), Elsevier, 2009. doi:10.1016/b978-0-12-374457-9.x0001-7.
URL <https://doi.org/10.1016/b978-0-12-374457-9.x0001-7>
- [127] C. F. Daganzo, [Traffic flow theory](#), in: Fundamentals of Transportation and Traffic Operations, Emerald Group Publishing Limited, 1997, pp. 66–160. doi:10.1108/9780585475301-004.
URL <https://doi.org/10.1108/9780585475301-004>
- [128] B. S. Kerner, [Introduction to Modern Traffic Flow Theory and Control](#), Springer Berlin Heidelberg, 2009. doi:10.1007/978-3-642-02605-8.
URL <https://doi.org/10.1007/978-3-642-02605-8>
- [129] W. Leutzbach, [Introduction to the Theory of Traffic Flow](#), Springer Berlin Heidelberg, 1988. doi:10.1007/978-3-642-61353-1.
URL <https://doi.org/10.1007/978-3-642-61353-1>
- [130] L. Elefteriadou, [An Introduction to Traffic Flow Theory](#), Springer New York, 2014. doi:10.1007/978-1-4614-8435-6.
URL <https://doi.org/10.1007/978-1-4614-8435-6>
- [131] J. A. Laval, C. F. Daganzo, [Lane-changing in traffic streams](#), Transportation Research Part B: Methodological 40 (3) (2006) 251–264. doi:10.1016/j.trb.2005.04.003.
URL <https://doi.org/10.1016/j.trb.2005.04.003>
- [132] W. J. Schakel, V. L. Knoop, B. van Arem, [Integrated lane change model with relaxation and synchronization](#), Transportation Research Record: Journal of the Transportation Research Board 2316 (1) (2012) 47–57. doi:10.3141/2316-06.
URL <https://doi.org/10.3141/2316-06>
- [133] J. A. Laval, L. Leclercq, [Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model](#), Transportation Research Part B: Methodological 42 (6) (2008) 511–522. doi:10.1016/j.trb.2007.10.004.
URL <https://doi.org/10.1016/j.trb.2007.10.004>
- [134] T.-Q. Tang, S. C. Wong, H.-J. Huang, P. Zhang, [Macroscopic modeling of lane-changing for two-lane traffic flow](#), Journal of Advanced Transportation 43 (3) (2009) 245–273. doi:10.1002/atr.5670430302.
URL <https://doi.org/10.1002/atr.5670430302>
- [135] W.-L. Jin, [A kinematic wave theory of lane-changing traffic flow](#), Transportation Research Part B: Methodological 44 (8-9) (2010) 1001–1021. doi:10.1016/j.trb.2009.12.014.
URL <https://doi.org/10.1016/j.trb.2009.12.014>
- [136] D. D. Salvucci, A. Liu, [The time course of a lane change: Driver control and eye-movement behavior](#), Transportation Research Part F: Traffic Psychology and Behaviour 5 (2) (2002) 123–132. doi:10.1016/s1369-8478(02)00011-6.
URL [https://doi.org/10.1016/s1369-8478\(02\)00011-6](https://doi.org/10.1016/s1369-8478(02)00011-6)
- [137] J. M. Molera, F. C. Martínez, J. A. Cuesta, R. Brito, [Theoretical approach to two-dimensional traffic flow models](#), Physical Review E 51 (1) (1995) 175–187. doi:10.1103/physreve.51.175.
URL <https://doi.org/10.1103/physreve.51.175>
- [138] M. Herty, A. Tosin, G. Visconti, M. Zanella, [Hybrid stochastic kinetic description of two-dimensional traffic dynamics](#), SIAM Journal on Applied Mathematics 78 (5) (2018) 2737–2762. doi:10.1137/17m1155909.
URL <https://doi.org/10.1137/17m1155909>
- [139] M. Herty, A. Fazekas, G. V. and, [A two-dimensional data-driven model for traffic flow on highways](#), Networks & Heterogeneous Media 13 (2) (2018) 217–240. doi:10.3934/nhm.2018010.
URL <https://doi.org/10.3934/nhm.2018010>
- [140] J.-W. Zeng, Y.-S. Qian, S.-B. Yu, X.-T. Wei, [Research on critical characteristics of highway traffic flow based on three phase traffic theory](#), Physica A: Statistical Mechanics and its Applications 530 (2019) 121567. doi:10.1016/j.physa.2019.121567.
URL <https://doi.org/10.1016/j.physa.2019.121567>
- [141] V. Papathanasopoulou, C. Antoniou, [Flexible car-following models for mixed traffic and weak lane-discipline conditions](#), European Transport Research Review 10 (2) (jun 2018). doi:10.1186/s12544-018-0338-0.
URL <https://doi.org/10.1186/s12544-018-0338-0>

- [142] M. Wang, S. P. Hoogendoorn, W. Daamen, B. van Arem, R. Happee, [Game theoretic approach for predictive lane-changing and car-following control](#), *Transportation Research Part C: Emerging Technologies* 58 (2015) 73–92. doi:10.1016/j.trc.2015.07.009.
URL <https://doi.org/10.1016/j.trc.2015.07.009>
- [143] D. Yang, S. Zheng, C. Wen, P. J. Jin, B. Ran, [A dynamic lane-changing trajectory planning model for automated vehicles](#), *Transportation Research Part C: Emerging Technologies* 95 (2018) 228–247. doi:10.1016/j.trc.2018.06.007.
URL <https://doi.org/10.1016/j.trc.2018.06.007>
- [144] S. Singh, [Critical reasons for crashes investigated in the national motor vehicle crash causation survey](#), Vol. 2, 2015, pp. 1–2.
URL <http://www-nrd.nhtsa.dot.gov/Pubs/812115.pdf>
- [145] C. Fu, T. Sayed, [Comparison of threshold determination methods for the deceleration rate to avoid a crash \(DRAC\)-based crash estimation](#), *Accident Analysis & Prevention* 153 (2021) 106051. doi:10.1016/j.aap.2021.106051.
URL <https://doi.org/10.1016%2Fj.aap.2021.106051>
- [146] D. Yang, K. Xie, K. Ozbay, H. Yang, [Fusing crash data and surrogate safety measures for safety assessment: Development of a structural equation model with conditional autoregressive spatial effect and random parameters](#), *Accident Analysis & Prevention* 152 (2021) 105971. doi:10.1016/j.aap.2021.105971.
URL <https://doi.org/10.1016%2Fj.aap.2021.105971>
- [147] A. P. Tarko, [Chapter 17. surrogate measures of safety](#), in: *Transport and Sustainability*, Emerald Publishing Limited, 2018, pp. 383–405. doi:10.1108/s2044-994120180000011019.
URL <https://doi.org/10.1108%2Fs2044-994120180000011019>
- [148] T. Sayed, S. Zein, [Traffic conflict standards for intersections](#), *Transportation Planning and Technology* 22 (4) (1999) 309–323. doi:10.1080/03081069908717634.
URL <https://doi.org/10.1080%2F03081069908717634>
- [149] B. Zhou, X. Wang, S. Zhang, Z. Li, S. Sun, K. Shu, Q. Sun, [Comparing factors affecting injury severity of passenger car and truck drivers](#), *IEEE Access* 8 (2020) 153849–153861. doi:10.1109/access.2020.3018183.
URL <https://doi.org/10.1109%2Faccess.2020.3018183>
- [150] L. Li, C. G. Prato, Y. Wang, [Ranking contributors to traffic crashes on mountainous freeways from an incomplete dataset: A sequential approach of multivariate imputation by chained equations and random forest classifier](#), *Accident Analysis & Prevention* 146 (2020) 105744. doi:10.1016/j.aap.2020.105744.
URL <https://doi.org/10.1016%2Fj.aap.2020.105744>
- [151] B. Wali, A. J. Khattak, T. Karnowski, [Exploring microscopic driving volatility in naturalistic driving environment prior to involvement in safety critical events—concept of event-based driving volatility](#), *Accident Analysis & Prevention* 132 (2019) 105277. doi:10.1016/j.aap.2019.105277.
URL <https://doi.org/10.1016/j.aap.2019.105277>
- [152] A. Lareshyn, T. D. Ceunynck, C. Karlsson, Å. Svensson, S. Daniels, [In search of the severity dimension of traffic events: Extended delta-v as a traffic conflict indicator](#), *Accident Analysis & Prevention* 98 (2017) 46–56. doi:10.1016/j.aap.2016.09.026.
URL <https://doi.org/10.1016/j.aap.2016.09.026>
- [153] A. Lareshyn, Å. Svensson, C. Hydén, [Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation](#), *Accident Analysis & Prevention* 42 (6) (2010) 1637–1646. doi:10.1016/j.aap.2010.03.021.
URL <https://doi.org/10.1016/j.aap.2010.03.021>
- [154] Z. Chen, X. Qin, M. R. R. Shaon, [Modeling lane-change-related crashes with lane-specific real-time traffic and weather data](#), *Journal of Intelligent Transportation Systems* 22 (4) (2017) 291–300. doi:10.1080/15472450.2017.1309529.
URL <https://doi.org/10.1080/15472450.2017.1309529>
- [155] Z. Li, S. Ahn, K. Chung, D. R. Ragland, W. Wang, J. W. Yu, [Surrogate safety measure for evaluating rear-end collision risk related to kinematic waves near freeway recurrent bottlenecks](#), *Accident Analysis & Prevention* 64 (2014) 52–61. doi:10.1016/j.aap.2013.11.003.
URL <https://doi.org/10.1016%2Fj.aap.2013.11.003>

- [156] C. Oh, S. Park, S. G. Ritchie, [A method for identifying rear-end collision risks using inductive loop detectors](#), *Accident Analysis & Prevention* 38 (2) (2006) 295–301. doi:10.1016/j.aap.2005.09.009. URL <https://doi.org/10.1016%2Fj.aap.2005.09.009>
- [157] A. Pande, M. Abdel-Aty, [Comprehensive analysis of the relationship between real-time traffic surveillance data and rear-end crashes on freeways](#), *Transportation Research Record: Journal of the Transportation Research Board* 1953 (1) (2006) 31–40. doi:10.1177/0361198106195300104. URL <https://doi.org/10.1177%2F0361198106195300104>
- [158] F. A. Mullakkal-Babu, M. Wang, H. Farah, B. van Arem, R. Happee, [Comparative assessment of safety indicators for vehicle trajectories on highways](#), *Transportation Research Record: Journal of the Transportation Research Board* 2659 (1) (2017) 127–136. doi:10.3141/2659-14. URL <https://doi.org/10.3141%2F2659-14>
- [159] A. Arun, M. M. Haque, A. Bhaskar, S. Washington, T. Sayed, [A systematic mapping review of surrogate safety assessment using traffic conflict techniques](#), *Accident Analysis & Prevention* 153 (2021) 106016. doi:10.1016/j.aap.2021.106016. URL <https://doi.org/10.1016%2Fj.aap.2021.106016>
- [160] L. Zheng, K. Ismail, X. Meng, [Traffic conflict techniques for road safety analysis: open questions and some insights](#), *Canadian Journal of Civil Engineering* 41 (7) (2014) 633–641. doi:10.1139/cjce-2013-0558. URL <https://doi.org/10.1139/cjce-2013-0558>
- [161] J. R. Ward, G. Agamennoni, S. Worrall, A. Bender, E. Nebot, [Extending time to collision for probabilistic reasoning in general traffic scenarios](#), *Transportation Research Part C: Emerging Technologies* 51 (2015) 66–82. doi:10.1016/j.trc.2014.11.002. URL <https://doi.org/10.1016/j.trc.2014.11.002>
- [162] N. Nadimi, D. R. Ragland, A. M. Amiri, [An evaluation of time-to-collision as a surrogate safety measure and a proposal of a new method for its application in safety analysis](#), *Transportation Letters* 12 (7) (2019) 491–500. doi:10.1080/19427867.2019.1650430. URL <https://doi.org/10.1080%2F19427867.2019.1650430>
- [163] S. Liu, X. Wang, O. Hassanin, X. Xu, M. Yang, D. Hurwitz, X. Wu, [Calibration and evaluation of responsibility-sensitive safety \(RSS\) in automated vehicle performance during cut-in scenarios](#), *Transportation Research Part C: Emerging Technologies* 125 (2021) 103037. doi:10.1016/j.trc.2021.103037. URL <https://doi.org/10.1016%2Fj.trc.2021.103037>
- [164] C. Hydén, L. Linderholm, [The swedish traffic-conflicts technique](#), in: *International Calibration Study of Traffic Conflict Techniques*, Springer Berlin Heidelberg, 1984, pp. 133–139. doi:10.1007/978-3-642-82109-7_12. URL https://doi.org/10.1007/978-3-642-82109-7_12
- [165] S. S. Mahmud, L. Ferreira, M. S. Hoque, A. Tavassoli, [Application of proximal surrogate indicators for safety evaluation: A review of recent developments and research needs](#), *IATSS Research* 41 (4) (2017) 153–163. doi:10.1016/j.iatssr.2017.02.001. URL <https://doi.org/10.1016%2Fj.iatssr.2017.02.001>
- [166] Y. Luo, Y. Xiang, K. Cao, K. Li, [A dynamic automated lane change maneuver based on vehicle-to-vehicle communication](#), *Transportation Research Part C: Emerging Technologies* 62 (2016) 87–102. doi:10.1016/j.trc.2015.11.011. URL <https://doi.org/10.1016/j.trc.2015.11.011>
- [167] S. I. Guler, M. Menendez, L. Meier, [Using connected vehicle technology to improve the efficiency of intersections](#), *Transportation Research Part C: Emerging Technologies* 46 (2014) 121–131. doi:10.1016/j.trc.2014.05.008. URL <https://doi.org/10.1016/j.trc.2014.05.008>
- [168] M. Sepulcre, J. Gozalvez, J. Hernandez, [Cooperative vehicle-to-vehicle active safety testing under challenging conditions](#), *Transportation Research Part C: Emerging Technologies* 26 (2013) 233–255. doi:10.1016/j.trc.2012.10.003. URL <https://doi.org/10.1016/j.trc.2012.10.003>
- [169] R. Yu, Y. Wang, Z. Zou, L. Wang, [Convolutional neural networks with refined loss functions for the real-time crash risk analysis](#), *Transportation Research Part C: Emerging Technologies* 119 (2020) 102740. doi:10.1016/j.trc.2020.102740. URL <https://doi.org/10.1016/j.trc.2020.102740>

- [170] M. Hossain, M. Abdel-Aty, M. A. Quddus, Y. Muromachi, S. N. Sadeek, [Real-time crash prediction models: State-of-the-art, design pathways and ubiquitous requirements](#), *Accident Analysis & Prevention* 124 (2019) 66–84. doi:10.1016/j.aap.2018.12.022.
URL <https://doi.org/10.1016/j.aap.2018.12.022>
- [171] Å. Svensson, C. Hydén, Estimating the severity of safety related behaviour, *Accident Analysis & Prevention* 38 (2) (2006) 379–385. doi:10.1016/j.aap.2005.10.009.
- [172] M. M. Minderhoud, P. H. Bovy, Extended time-to-collision measures for road traffic safety assessment, *Accident Analysis & Prevention* 33 (1) (2001) 89–97. doi:10.1016/S0001-4575(00)00019-1.
- [173] J. Wang, Q. Zhang, D. Zhao, Y. Chen, [Lane change decision-making through deep reinforcement learning with rule-based constraints](#), in: 2019 International Joint Conference on Neural Networks (IJCNN), IEEE, 2019. doi:10.1109/ijcnn.2019.8852110.
URL <https://doi.org/10.1109/ijcnn.2019.8852110>
- [174] B. Mirchevska, C. Pek, M. Werling, M. Althoff, J. Boedecker, [High-level decision making for safe and reasonable autonomous lane changing using reinforcement learning](#), in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), IEEE, 2018. doi:10.1109/itsc.2018.8569448.
URL <https://doi.org/10.1109/itsc.2018.8569448>
- [175] J. Nie, X. Wan, J. Zhang, W. Ding, B. Ran, Modelling of vehicle interaction behavior during discretionary lane-changing preparation process on freeway, 2017.
- [176] J. Wu, H. Wen, W. Qi, [A new method of temporal and spatial risk estimation for lane change considering conventional recognition defects](#), *Accident Analysis & Prevention* 148 (2020) 105796. doi:10.1016/j.aap.2020.105796.
URL <https://doi.org/10.1016/j.aap.2020.105796>
- [177] A. Ji, D. Levinson, [A review of game theory models of lane changing](#), *Transportmetrica A: Transport Science* 16 (3) (2020) 1628–1647. doi:10.1080/23249935.2020.1770368.
URL <https://doi.org/10.1080/23249935.2020.1770368>
- [178] C.-J. Jin, V. L. Knoop, D. Li, L.-Y. Meng, H. Wang, [Discretionary lane-changing behavior: empirical validation for one realistic rule-based model](#), *Transportmetrica A: Transport Science* 15 (2) (2018) 244–262. doi:10.1080/23249935.2018.1464526.
URL <https://doi.org/10.1080/23249935.2018.1464526>
- [179] T. van Dijk, G. van der Heijden, [VisionSense: an advanced lateral collision warning system](#), in: IEEE Proceedings. Intelligent Vehicles Symposium, 2005., IEEE, 2005. doi:10.1109/ivs.2005.1505118.
URL <https://doi.org/10.1109/ivs.2005.1505118>
- [180] J. Li-sheng, F. Wen-ping, Z. Ying-nan, Y. Shuang-bin, H. Hai-jing, [Research on safety lane change model of driver assistant system on highway](#), in: 2009 IEEE Intelligent Vehicles Symposium, IEEE, 2009. doi:10.1109/ivs.2009.5164426.
URL <https://doi.org/10.1109%2Fivs.2009.5164426>
- [181] Z. Zheng, S. Ahn, D. Chen, J. Laval, [The effects of lane-changing on the immediate follower: Anticipation, relaxation, and change in driver characteristics](#), *Transportation Research Part C: Emerging Technologies* 26 (2013) 367–379. doi:10.1016/j.trc.2012.10.007.
URL <https://doi.org/10.1016%2Fj.trc.2012.10.007>
- [182] Z. Zheng, [Recent developments and research needs in modeling lane changing](#), *Transportation Research Part B: Methodological* 60 (2014) 16–32. doi:10.1016/j.trb.2013.11.009.
URL <https://doi.org/10.1016%2Fj.trb.2013.11.009>
- [183] J. Zheng, K. Suzuki, M. Fujita, [Predicting driver's lane-changing decisions using a neural network model](#), *Simulation Modelling Practice and Theory* 42 (2014) 73–83. doi:10.1016/j.simpat.2013.12.007.
URL <https://doi.org/10.1016%2Fj.simpat.2013.12.007>
- [184] T. Chen, X. Shi, Y. D. Wong, [Key feature selection and risk prediction for lane-changing behaviors based on vehicles' trajectory data](#), *Accident Analysis & Prevention* 129 (2019) 156–169. doi:10.1016/j.aap.2019.05.017.
URL <https://doi.org/10.1016%2Fj.aap.2019.05.017>
- [185] P. Gipps, [A model for the structure of lane-changing decisions](#), *Transportation Research Part B: Methodological* 20 (5) (1986) 403–414. doi:10.1016/0191-2615(86)90012-3.
URL [https://doi.org/10.1016/0191-2615\(86\)90012-3](https://doi.org/10.1016/0191-2615(86)90012-3)

- [186] P. Hidas, [Modelling lane changing and merging in microscopic traffic simulation](#), Transportation Research Part C: Emerging Technologies 10 (5-6) (2002) 351–371. doi:10.1016/s0968-090x(02)00026-8. URL [https://doi.org/10.1016/s0968-090x\(02\)00026-8](https://doi.org/10.1016/s0968-090x(02)00026-8)
- [187] A. Kesting, M. Treiber, D. Helbing, [General lane-changing model MOBIL for car-following models](#), Transportation Research Record: Journal of the Transportation Research Board 1999 (1) (2007) 86–94. doi:10.3141/1999-10. URL <https://doi.org/10.3141/1999-10>
- [188] S. H. Hamdar, H. S. Mahmassani, [Life in the fast lane](#), Transportation Research Record: Journal of the Transportation Research Board 2124 (1) (2009) 89–102. doi:10.3141/2124-09. URL <https://doi.org/10.3141/2124-09>
- [189] M. Keyvan-Ekbatani, V. L. Knoop, W. Daamen, [Categorization of the lane change decision process on freeways](#), Transportation Research Part C: Emerging Technologies 69 (2016) 515–526. doi:10.1016/j.trc.2015.11.012. URL <https://doi.org/10.1016/j.trc.2015.11.012>
- [190] D. L. Smith, R. Glassco, J. Chang, D. Cohen, [Feasibility of modeling lane-change performance](#), in: SAE Technical Paper Series, SAE International, 2003. doi:10.4271/2003-01-0280. URL <https://doi.org/10.4271%2F2003-01-0280>
- [191] J. L. Bascuñana, [Analysis of lane change crash avoidance](#), in: SAE Technical Paper Series, SAE International, 1995. doi:10.4271/951895. URL <https://doi.org/10.4271/951895>
- [192] G. Lu, M. Tomizuka, [Vehicle following as backup control schemes for magnet-magnetometer-based lateral guidance](#), IEEE Transactions on Control Systems Technology 13 (2) (2005) 274–285. doi:10.1109/tcst.2004.839564. URL <https://doi.org/10.1109%2Ftcst.2004.839564>
- [193] F. A. Mullakkal-Babu, M. Wang, B. van Arem, R. Happee, [Empirics and models of fragmented lane changes](#), IEEE Open Journal of Intelligent Transportation Systems 1 (2020) 187–200. doi:10.1109/ojits.2020.3029056. URL <https://doi.org/10.1109%2Fojects.2020.3029056>
- [194] T. Toledo, D. Zohar, [Modeling duration of lane changes](#), Transportation Research Record: Journal of the Transportation Research Board 1999 (1) (2007) 71–78. doi:10.3141/1999-08. URL <https://doi.org/10.3141/1999-08>
- [195] E. Balal, R. L. Cheu, T. Gyan-Sarkodie, J. Miramontes, [Analysis of discretionary lane changing parameters on freeways](#), International Journal of Transportation Science and Technology 3 (3) (2014) 277–296. doi:10.1260/2046-0430.3.3.277. URL <https://doi.org/10.1260/2046-0430.3.3.277>
- [196] W. van Winsum, D. de Waard, K. Brookhuis, [Lane change manoeuvres and safety margins](#), Transportation Research Part F: Traffic Psychology and Behaviour 2 (3) (1999) 139–149. doi:10.1016/s1369-8478(99)00011-x. URL [https://doi.org/10.1016/s1369-8478\(99\)00011-x](https://doi.org/10.1016/s1369-8478(99)00011-x)
- [197] Q. Wang, Z. Li, L. Li, [Investigation of discretionary lane-change characteristics using next-generation simulation data sets](#), Journal of Intelligent Transportation Systems 18 (3) (2014) 246–253. doi:10.1080/15472450.2013.810994. URL <https://doi.org/10.1080/15472450.2013.810994>
- [198] E. Balal, R. L. Cheu, T. Sarkodie-Gyan, [A binary decision model for discretionary lane changing move based on fuzzy inference system](#), Transportation Research Part C: Emerging Technologies 67 (2016) 47–61. doi:10.1016/j.trc.2016.02.009. URL <https://doi.org/10.1016/j.trc.2016.02.009>
- [199] Z. qiang Liu, T. Zhang, Y. fan Wang, [Research on local dynamic path planning method for intelligent vehicle lane-changing](#), Journal of Advanced Transportation 2019 (2019) 1–10. doi:10.1155/2019/4762658. URL <https://doi.org/10.1155/2019/4762658>
- [200] P. Ferrari, [The effect of driver behaviour on motorway reliability](#), Transportation Research Part B: Methodological 23 (2) (1989) 139–150. doi:10.1016/0191-2615(89)90037-4. URL [https://doi.org/10.1016/0191-2615\(89\)90037-4](https://doi.org/10.1016/0191-2615(89)90037-4)

- [201] F. Meng, J. Su, C. Liu, W.-H. Chen, [Dynamic decision making in lane change: Game theory with receding horizon](#), in: 2016 UKACC 11th International Conference on Control (CONTROL), IEEE, 2016. doi:10.1109/control.2016.7737643.
URL <https://doi.org/10.1109/control.2016.7737643>
- [202] T. Toledo, H. N. Koutsopoulos, M. E. Ben-Akiva, [Modeling integrated lane-changing behavior](#), Transportation Research Record: Journal of the Transportation Research Board 1857 (1) (2003) 30–38. doi:10.3141/1857-04.
URL <https://doi.org/10.3141/1857-04>
- [203] I. Ahmed, A. Karr, N. M. Roupail, G. Chun, S. Tanvir, [Characterizing lane changes via digitized infrastructure and low-cost GPS](#), Transportation Research Record: Journal of the Transportation Research Board 2673 (8) (2019) 298–309. doi:10.1177/0361198119841277.
URL <https://doi.org/10.1177/0361198119841277>
- [204] E. C. B. Olsen, S. E. Lee, W. W. Wierwille, M. J. Goodman, [Analysis of distribution, frequency, and duration of naturalistic lane changes](#), Proceedings of the Human Factors and Ergonomics Society Annual Meeting 46 (22) (2002) 1789–1793. doi:10.1177/154193120204602203.
URL <https://doi.org/10.1177/154193120204602203>
- [205] C. Wang, Q. Sun, Z. Li, H. Zhang, [Human-like lane change decision model for autonomous vehicles that considers the risk perception of drivers in mixed traffic](#), Sensors 20 (8) (2020) 2259. doi:10.3390/s20082259.
URL <https://doi.org/10.3390/s20082259>
- [206] A. Montella, R. de Oña, F. Mauriello, M. R. Riccardi, G. Silvestro, [A data mining approach to investigate patterns of powered two-wheeler crashes in Spain](#), Accident Analysis & Prevention 134 (2020) 105251. doi:10.1016/j.aap.2019.07.027.
URL <https://doi.org/10.1016/j.aap.2019.07.027>
- [207] E. N. Barmounakis, E. I. Vlahogianni, J. C. Golias, [Intelligent transportation systems and powered two wheelers traffic](#), IEEE Transactions on Intelligent Transportation Systems 17 (4) (2016) 908–916. doi:10.1109/tits.2015.2497406.
URL <https://doi.org/10.1109/tits.2015.2497406>
- [208] Y. Guo, T. Sayed, M. H. Zaki, [Evaluating the safety impacts of powered two wheelers on a shared roadway in China using automated video analysis](#), Journal of Transportation Safety & Security 11 (4) (2018) 414–429. doi:10.1080/19439962.2018.1447058.
URL <https://doi.org/10.1080/19439962.2018.1447058>
- [209] E. I. Vlahogianni, [Powered-two-wheelers kinematic characteristics and interactions during filtering and overtaking in urban arterials](#), Transportation Research Part F: Traffic Psychology and Behaviour 24 (2014) 133–145. doi:10.1016/j.trf.2014.04.004.
URL <https://doi.org/10.1016/j.trf.2014.04.004>
- [210] E. I. Vlahogianni, G. Yannis, J. C. Golias, [Detecting powered-two-wheeler incidents from high resolution naturalistic data](#), Transportation Research Part F: Traffic Psychology and Behaviour 22 (2014) 86–95. doi:10.1016/j.trf.2013.11.002.
URL <https://doi.org/10.1016/j.trf.2013.11.002>
- [211] E. I. Vlahogianni, G. Yannis, J. C. Golias, [Critical power two wheeler driving patterns at the emergence of an incident](#), Accident Analysis & Prevention 58 (2013) 340–345. doi:10.1016/j.aap.2012.12.026.
URL <https://doi.org/10.1016/j.aap.2012.12.026>
- [212] T.-C. Lee, J. W. Polak, M. G. H. Bell, [New approach to modeling mixed traffic containing motorcycles in urban areas](#), Transportation Research Record: Journal of the Transportation Research Board 2140 (1) (2009) 195–205. doi:10.3141/2140-22.
URL <https://doi.org/10.3141/2140-22>
- [213] R. Nair, H. S. Mahmassani, E. Miller-Hooks, [A porous flow approach to modeling heterogeneous traffic in disordered systems](#), Transportation Research Part B: Methodological 45 (9) (2011) 1331–1345. doi:10.1016/j.trb.2011.05.009.
URL <https://doi.org/10.1016/j.trb.2011.05.009>

- [214] W. H. Organization, Powered two- and three-wheeler safety: a road safety manual for decision-makers and practitioners, World Health Organization, 2017.
- [215] R. Q. Ivers, C. Sakashita, T. Senserrick, J. Elkington, S. Lo, S. Boufous, L. de Rome, [Does an on-road motorcycle coaching program reduce crashes in novice riders? a randomised control trial](#), *Accident Analysis & Prevention* 86 (2016) 40–46. doi:10.1016/j.aap.2015.10.015. URL <https://doi.org/10.1016%2Fj.aap.2015.10.015>
- [216] K.-F. Wu, L. Sasidharan, C. P. Thor, S.-Y. Chen, [Crash sequence based risk matrix for motorcycle crashes](#), *Accident Analysis & Prevention* 117 (2018) 21–31. doi:10.1016/j.aap.2018.03.022. URL <https://doi.org/10.1016%2Fj.aap.2018.03.022>
- [217] P. Puthan, N. Lubbe, J. Shaikh, B. Sui, J. Davidsson, [Defining crash configurations for powered two-wheelers: Comparing ISO 13232 to recent in-depth crash data from germany, india and china](#), *Accident Analysis & Prevention* 151 (2021) 105957. doi:10.1016/j.aap.2020.105957. URL <https://doi.org/10.1016%2Fj.aap.2020.105957>
- [218] L. T. Aarts, J. J. F. Commandeur, R. Welsh, S. Niesen, M. Lerner, P. Thomas, N. Bos, R. J. Davidse, Study on serious road traffic injuries in the eu, Brussels: European Commission (2016).
- [219] D. F. Preusser, A. F. Williams, R. G. Ulmer, [Analysis of fatal motorcycle crashes: crash typing](#), *Accident Analysis & Prevention* 27 (6) (1995) 845–851. doi:10.1016/0001-4575(95)00027-5. URL <https://doi.org/10.1016%2F0001-4575%2895%2900027-5>
- [220] M.-R. Lin, J. F. Kraus, [A review of risk factors and patterns of motorcycle injuries](#), *Accident Analysis & Prevention* 41 (4) (2009) 710–722. doi:10.1016/j.aap.2009.03.010. URL <https://doi.org/10.1016%2Fj.aap.2009.03.010>
- [221] M. M. Haque, H. C. Chin, H. Huang, [Modeling fault among motorcyclists involved in crashes](#), *Accident Analysis & Prevention* 41 (2) (2009) 327–335. doi:10.1016/j.aap.2008.12.010. URL <https://doi.org/10.1016%2Fj.aap.2008.12.010>
- [222] C.-W. Pai, [Motorcyclist injury severity in angle crashes at t-junctions: Identifying significant factors and analysing what made motorists fail to yield to motorcycles](#), *Safety Science* 47 (8) (2009) 1097–1106. doi:10.1016/j.ssci.2008.12.007. URL <https://doi.org/10.1016%2Fj.ssci.2008.12.007>
- [223] E. I. Vlahogianni, G. Yannis, J. C. Golias, [Overview of critical risk factors in power-two-wheeler safety](#), *Accident Analysis & Prevention* 49 (2012) 12–22. doi:10.1016/j.aap.2012.04.009. URL <https://doi.org/10.1016%2Fj.aap.2012.04.009>
- [224] F. Chang, M. Li, P. Xu, H. Zhou, M. Haque, H. Huang, [Injury severity of motorcycle riders involved in traffic crashes in hunan, china: A mixed ordered logit approach](#), *International Journal of Environmental Research and Public Health* 13 (7) (2016) 714. doi:10.3390/ijerph13070714. URL <https://doi.org/10.3390%2Fijerph13070714>
- [225] F. J. C. Cunto, S. Ferreira, [An analysis of the injury severity of motorcycle crashes in brazil using mixed ordered response models](#), *Journal of Transportation Safety & Security* 9 (sup1) (2016) 33–46. doi:10.1080/19439962.2016.1162891. URL <https://doi.org/10.1080%2F19439962.2016.1162891>
- [226] F. L. Mannering, V. Shankar, C. R. Bhat, [Unobserved heterogeneity and the statistical analysis of highway accident data](#), *Analytic Methods in Accident Research* 11 (2016) 1–16. doi:10.1016/j.amar.2016.04.001. URL <https://doi.org/10.1016%2Fj.amar.2016.04.001>
- [227] J. C. Milton, V. N. Shankar, F. L. Mannering, [Highway accident severities and the mixed logit model: An exploratory empirical analysis](#), *Accident Analysis & Prevention* 40 (1) (2008) 260–266. doi:10.1016/j.aap.2007.06.006. URL <https://doi.org/10.1016%2Fj.aap.2007.06.006>
- [228] J. R. Treat, N. S. Tumbas, S. T. McDonald, D. Shinar, R. D. Hume, R. E. Mayer, R. L. Stansifer, N. J. Castellan, [Tri-level study of the causes of traffic accidents: Final report. volume i: Casual factor tabulations and assessments](#) (1979). doi:10.1037/e488172008-001. URL <https://doi.org/10.1037%2Fe488172008-001>

- [229] K.-F. Wu, C. P. Thor, [Method for the use of naturalistic driving study data to analyze rear-end crash sequences](#), Transportation Research Record: Journal of the Transportation Research Board 2518 (1) (2015) 27–36. doi:10.3141/2518-04. URL <https://doi.org/10.3141%2F2518-04>
- [230] N. Clabaux, J.-Y. Fournier, J.-E. Michel, [Powered two-wheeler riders' risk of crashes associated with filtering on urban roads](#), Traffic Injury Prevention 18 (2) (2017) 182–187, pMID: 27589253. arXiv:<https://doi.org/10.1080/15389588.2016.1225298>, doi:10.1080/15389588.2016.1225298. URL <https://doi.org/10.1080/15389588.2016.1225298>
- [231] N. Clabaux, J.-Y. Fournier, J.-E. Michel, [Powered two-wheeler drivers' crash risk associated with the use of bus lanes](#), Accident Analysis & Prevention 71 (2014) 306–310. doi:10.1016/j.aap.2014.05.021. URL <https://doi.org/10.1016%2Fj.aap.2014.05.021>
- [232] D. Atalar, P. Thomas, [Powered two-wheeler crash scenario development](#), Accident Analysis & Prevention 125 (2019) 198–206. doi:10.1016/j.aap.2019.02.001. URL <https://doi.org/10.1016/j.aap.2019.02.001>
- [233] S. M. Gaweesh, M. M. Ahmed, A. V. Piccorelli, [Developing crash prediction models using parametric and nonparametric approaches for rural mountainous freeways: A case study on wyoming interstate 80](#), Accident Analysis & Prevention 123 (2019) 176–189. doi:10.1016/j.aap.2018.10.011. URL <https://doi.org/10.1016/j.aap.2018.10.011>
- [234] G. Yang, M. Ahmed, S. Gaweesh, E. Adomah, [Connected vehicle real-time traveler information messages for freeway speed harmonization under adverse weather conditions: Trajectory level analysis using driving simulator](#), Accident Analysis & Prevention 146 (2020) 105707. doi:10.1016/j.aap.2020.105707. URL <https://doi.org/10.1016/j.aap.2020.105707>
- [235] O. Bagdadi, [Assessing safety critical braking events in naturalistic driving studies](#), Transportation Research Part F: Traffic Psychology and Behaviour 16 (2013) 117–126. doi:10.1016/j.trf.2012.08.006. URL <https://doi.org/10.1016/j.trf.2012.08.006>
- [236] Y. Wang, Y. Luo, F. Chen, [Interpreting risk factors for truck crash severity on mountainous freeways in jiangxi and shaanxi, china](#), European Transport Research Review 11 (1) (may 2019). doi:10.1186/s12544-019-0366-4. URL <https://doi.org/10.1186%2Fs12544-019-0366-4>
- [237] E. Commission, [2018 road safety statistics: what is behind the figures?](#) (2019). URL https://ec.europa.eu/commission/presscorner/detail/en/MEMO_19_1990
- [238] D. J. Adminaite, G. Stipdonk, H. H & Ward, [Reducing deaths in single vehicle collisions](#) (2017).
- [239] S. G. Charlton, [The role of attention in horizontal curves: A comparison of advance warning, delineation, and road marking treatments](#), Accident Analysis & Prevention 39 (5) (2007) 873–885. doi:10.1016/j.aap.2006.12.007. URL <https://doi.org/10.1016%2Fj.aap.2006.12.007>
- [240] I. Milleville-Pennel, H. Jean-Michel, J. Elise, [The use of hazard road signs to improve the perception of severe bends](#), Accident Analysis & Prevention 39 (4) (2007) 721–730. doi:10.1016/j.aap.2006.11.001. URL <https://doi.org/10.1016%2Fj.aap.2006.11.001>
- [241] A. Calvi, [A study on driving performance along horizontal curves of rural roads](#), Journal of Transportation Safety & Security 7 (3) (2014) 243–267. doi:10.1080/19439962.2014.952468. URL <https://doi.org/10.1080%2F19439962.2014.952468>
- [242] D. Babić, T. Brijs, [Low-cost road marking measures for increasing safety in horizontal curves: A driving simulator study](#), Accident Analysis & Prevention 153 (2021) 106013. doi:10.1016/j.aap.2021.106013. URL <https://doi.org/10.1016%2Fj.aap.2021.106013>
- [243] J. Theeuwes, R. van der Horst, M. Kuiken, [Designing Safe Road Systems](#), CRC Press, 2017. doi:10.1201/9781315576732. URL <https://doi.org/10.1201%2F9781315576732>
- [244] S. L. Hallmark, N. Hawkins, O. Smadi, [Evaluation of low-cost treatments on rural two-lane curves](#), Tech. rep. (2012).
- [245] J. E. Hummer, W. Rasdorf, D. J. Findley, C. V. Zegeer, C. A. Sundstrom, [Curve collisions: Road and collision characteristics and countermeasures](#), Journal of Transportation Safety & Security 2 (3) (2010) 203–220. doi:10.1080/

19439961003734880.
URL <https://doi.org/10.1080%2F19439961003734880>
- [246] B. Lewis-Evans, S. G. Charlton, [Explicit and implicit processes in behavioural adaptation to road width](#), *Accident Analysis & Prevention* 38 (3) (2006) 610–617. doi:10.1016/j.aap.2005.12.005.
URL <https://doi.org/10.1016%2Fj.aap.2005.12.005>
- [247] F. Jørgensen, T. Wentzel-Larsen, [Optimal use of warning signs in traffic](#), *Accident Analysis & Prevention* 31 (6) (1999) 729–738. doi:10.1016/S0001-4575(99)00036-6.
URL <https://doi.org/10.1016%2Fs0001-4575%2899%2900036-6>
- [248] D. SHINAR†, T. H. ROCKWELL, J. A. MALECKI‡, [The effects of changes in driver perception on rural curve negotiation*](#), *Ergonomics* 23 (3) (1980) 263–275. doi:10.1080/00140138008924739.
URL <https://doi.org/10.1080%2F00140138008924739>
- [249] N. Haghighi, X. C. Liu, G. Zhang, R. J. Porter, [Impact of roadway geometric features on crash severity on rural two-lane highways](#), *Accident Analysis & Prevention* 111 (2018) 34–42. doi:10.1016/j.aap.2017.11.014.
URL <https://doi.org/10.1016/j.aap.2017.11.014>
- [250] J. P. Gooch, V. V. Gayah, E. T. Donnell, [Quantifying the safety effects of horizontal curves on two-way, two-lane rural roads](#), *Accident Analysis & Prevention* 92 (2016) 71–81. doi:10.1016/j.aap.2016.03.024.
URL <https://doi.org/10.1016/j.aap.2016.03.024>
- [251] J. P. Gooch, V. V. Gayah, E. T. Donnell, [Safety performance functions for horizontal curves and tangents on two lane, two way rural roads](#), *Accident Analysis & Prevention* 120 (2018) 28–37. doi:10.1016/j.aap.2018.07.030.
URL <https://doi.org/10.1016/j.aap.2018.07.030>
- [252] B. Persaud, C. Lyon, T. Nguyen, [Empirical bayes procedure for ranking sites for safety investigation by potential for safety improvement](#), *Transportation Research Record: Journal of the Transportation Research Board* 1665 (1) (1999) 7–12. doi:10.3141/1665-02.
URL <https://doi.org/10.3141%2F1665-02>
- [253] W. Cheng, S. P. Washington, [Experimental evaluation of hotspot identification methods](#), *Accident Analysis & Prevention* 37 (5) (2005) 870–881. doi:10.1016/j.aap.2005.04.015.
URL <https://doi.org/10.1016%2Fj.aap.2005.04.015>
- [254] F. Mannering, [Temporal instability and the analysis of highway accident data](#), *Analytic Methods in Accident Research* 17 (2018) 1–13. doi:10.1016/j.amar.2017.10.002.
URL <https://doi.org/10.1016%2Fj.amar.2017.10.002>
- [255] X. Guo, L. Wu, Y. Zou, L. Fawcett, [Comparative analysis of empirical bayes and bayesian hierarchical models in hotspot identification](#), *Transportation Research Record: Journal of the Transportation Research Board* 2673 (7) (2019) 111–121. doi:10.1177/0361198119849899.
URL <https://doi.org/10.1177%2F0361198119849899>
- [256] W. Cheng, S. Washington, [New criteria for evaluating methods of identifying hot spots](#), *Transportation Research Record: Journal of the Transportation Research Board* 2083 (1) (2008) 76–85. doi:10.3141/2083-09.
URL <https://doi.org/10.3141%2F2083-09>
- [257] X. Guo, L. Wu, D. Lord, [Generalized criteria for evaluating hotspot identification methods](#), *Accident Analysis & Prevention* 145 (2020) 105684. doi:10.1016/j.aap.2020.105684.
URL <https://doi.org/10.1016%2Fj.aap.2020.105684>
- [258] T. Harmon, G. B. Bahar, F. B. Gross, et al., *Crash costs for highway safety analysis*, Tech. rep., United States. Federal Highway Administration. Office of Safety (2018).
- [259] R. Tay, *Prioritising road safety initiatives: fatality versus social cost.*, *Applied health economics and health policy* 1 (3) (2002) 129–134.
- [260] A. P. Afghari, M. M. Haque, S. Washington, [Applying a joint model of crash count and crash severity to identify road segments with high risk of fatal and serious injury crashes](#), *Accident Analysis & Prevention* 144 (2020) 105615. doi:10.1016/j.aap.2020.105615.
URL <https://doi.org/10.1016/j.aap.2020.105615>

- [261] G. Lamptey, S. Labi, K. C. Sinha, [Investigating the sensitivity of optimal network safety needs to key safety management inputs](#), Transportation Research Record: Journal of the Transportation Research Board 1922 (1) (2005) 52–61. doi:10.1177/0361198105192200108.
URL <https://doi.org/10.1177%2F0361198105192200108>
- [262] J. Ambros, P. Havránek, V. Valentová, Z. Křivánková, R. Striegler, [Identification of hazardous locations in regional road network – comparison of reactive and proactive approaches](#), Transportation Research Procedia 14 (2016) 4209–4217. doi:10.1016/j.trpro.2016.05.392.
URL <https://doi.org/10.1016%2Fj.trpro.2016.05.392>
- [263] B. N. Persaud, [Statistical methods in highway safety analysis](#) (2001).
URL <https://trid.trb.org/view/694935>
- [264] S. Washington, A. P. Afghari, M. M. Haque, [Detecting high-risk accident locations](#), in: Transport and Sustainability, Emerald Publishing Limited, 2018, pp. 351–382. doi:10.1108/s2044-994120180000011018.
URL <https://doi.org/10.1108/s2044-994120180000011018>
- [265] G. Lamptey, L. Singh, S. Labi, K. C. Sinha, [Systematic framework for incorporating safety in network-level transportation planning and programming](#), Journal of Transportation Engineering 136 (5) (2010) 436–447. doi:10.1061/(asce)te.1943-5436.0000070.
URL <https://doi.org/10.1061%2F%28asce%29te.1943-5436.0000070>
- [266] I. E. Sutherland, Sketchpad: a man-machine graphical communication system, Simulation 2 (5) (1964) 3–20. doi:10.1177/003754976400200514.
- [267] A. V. A. Bharat Kumar, S. P. Venthuruthiyil, M. Chunchu, A novel coordinate transformation technique for image processing based naturalistic driving data extraction from complex road geometry, Working Paper (2020).
- [268] M. Danelljan, G. Bhat, F. S. Khan, M. Felsberg, [ECO: Efficient convolution operators for tracking](#), in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017. doi:10.1109/cvpr.2017.733.
URL <https://doi.org/10.1109/cvpr.2017.733>
- [269] R. Hartley, A. Zisserman, Esimation - 2D Projective Transformations, Cambridge University Press, Cambridge, 2004. doi:10.1017/CBO9780511811685.
- [270] C. C. Lin, S. U. Pankanti, K. N. Ramamurthy, A. Y. Aravkin, Adaptive as-natural-as-possible image stitching, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 07-12-June-2015, IEEE, 2015, pp. 1155–1163. doi:10.1109/CVPR.2015.7298719.
- [271] A. Vedaldi, B. Fulkerson, Vlfeat - An open and portable library of computer vision algorithms, MM'10 - Proceedings of the ACM Multimedia 2010 International Conference 3 (1) (2010) 1469–1472. doi:10.1145/1873951.1874249.
- [272] R. Szeliski, Image alignment and stitching, in: Handbook of Mathematical Models in Computer Vision, Springer-Verlag, New York, 2006, pp. 273–292. doi:10.1007/0-387-28831-7_17.
- [273] C. A. Hall, W. Meyer, [Optimal error bounds for cubic spline interpolation](#), Journal of Approximation Theory 16 (2) (1976) 105–122. doi:10.1016/0021-9045(76)90040-x.
URL [https://doi.org/10.1016/0021-9045\(76\)90040-x](https://doi.org/10.1016/0021-9045(76)90040-x)
- [274] C. de Boor, [A Practical Guide to Splines](#), Springer New York, 1978. doi:10.1007/978-1-4612-6333-3.
URL <https://doi.org/10.1007/978-1-4612-6333-3>
- [275] M. U. A. Bromba, H. Ziegler, [Application hints for savitzky-golay digital smoothing filters](#), Analytical Chemistry 53 (11) (1981) 1583–1586. doi:10.1021/ac00234a011.
URL <https://doi.org/10.1021/ac00234a011>
- [276] A. Savitzky, M. J. E. Golay, [Smoothing and differentiation of data by simplified least squares procedures.](#), Analytical Chemistry 36 (8) (1964) 1627–1639. doi:10.1021/ac60214a047.
URL <https://doi.org/10.1021/ac60214a047>
- [277] J. Fan, I. Gijbels, [Local Polynomial Modelling and Its Applications](#), Routledge, 2018. doi:10.1201/9780203748725.
URL <https://doi.org/10.1201/9780203748725>
- [278] W. S. Cleveland, [Robust locally weighted regression and smoothing scatterplots](#), Journal of the American Statistical Association 74 (368) (1979) 829. doi:10.2307/2286407.
URL <https://doi.org/10.2307/2286407>

- [279] T. A. Dingus, S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. A. Perez, J. Hankey, D. Ramsey, S. Gupta, C. Bucher, Z. R. Doerzaph, J. Jermeland, R. R. Knipling, [The 100-car naturalistic driving study: Phase II - results of the 100-car field experiment](#) (2006). doi:10.1037/e624282011-001.
URL <https://doi.org/10.1037/e624282011-001>
- [280] Q. Lin, R. Feng, B. Cheng, J. Lai, H. Zhang, B. Mei, [Analysis of causes of rear-end conflicts using naturalistic driving data collected by video drive recorders](#), in: SAE Technical Paper Series, SAE International, 2008. doi:10.4271/2008-01-0522.
URL <https://doi.org/10.4271/2008-01-0522>
- [281] F. Feng, S. Bao, J. R. Sayer, C. Flannagan, M. Manser, R. Wunderlich, [Can vehicle longitudinal jerk be used to identify aggressive drivers? an examination using naturalistic driving data](#), Accident Analysis & Prevention 104 (2017) 125–136. doi:10.1016/j.aap.2017.04.012.
URL <https://doi.org/10.1016/j.aap.2017.04.012>
- [282] MathWorks, [Apps overview](#), accessed on 2020/01/01 (2019).
URL https://in.mathworks.com/help/matlab/creating_guis/apps-overview.html
- [283] C. Tingvall, N. Haworth, [Vision zero - an ethical approach to safety and mobility](#), in: 6th ITE International Conference Road Safety & Traffic Enforcement: Beyond 2000, 1999.
URL <https://eprints.qut.edu.au/134991/>
- [284] A. Charly, T. V. Mathew, [Estimation of traffic conflicts using precise lateral position and width of vehicles for safety assessment](#), Accident Analysis & Prevention 132 (2019) 105264. doi:10.1016/j.aap.2019.105264.
URL <https://doi.org/10.1016/j.aap.2019.105264>
- [285] E. Kim, P. Muennig, Z. Rosen, [Vision zero: a toolkit for road safety in the modern era](#), Injury Epidemiology 4 (1) (jan 2017). doi:10.1186/s40621-016-0098-z.
URL <https://doi.org/10.1186/s40621-016-0098-z>
- [286] R. B. Naumann, S. Heiny, K. R. Evenson, S. LaJeunesse, J. F. Cooper, S. Doggett, S. W. Marshall, [Organizational networks in road safety: Case studies of u.s. vision zero cities](#), Traffic Injury Prevention 20 (4) (2019) 378–385. doi:10.1080/15389588.2019.1587752.
URL <https://doi.org/10.1080/15389588.2019.1587752>
- [287] A. Dijkstra, [Assessing the safety of routes in a regional network](#), Transportation Research Part C: Emerging Technologies 32 (2013) 103–115. doi:10.1016/j.trc.2012.10.008.
URL <https://doi.org/10.1016/j.trc.2012.10.008>
- [288] K. Ozbay, H. Yang, B. Bartin, S. Mudigonda, [Derivation and validation of new simulation-based surrogate safety measure](#), Transportation Research Record: Journal of the Transportation Research Board 2083 (1) (2008) 105–113. doi:10.3141/2083-12.
URL <https://doi.org/10.3141/2083-12>
- [289] AASHTO, A Policy on Geometric Design of Highways and Streets, 7th Edition, The American Association of State Highway and Transportation Officials, 2018.
- [290] T. Sayed, M. H. Zaki, J. Autey, [Automated safety diagnosis of vehicle–bicycle interactions using computer vision analysis](#), Safety Science 59 (2013) 163–172. doi:10.1016/j.ssci.2013.05.009.
URL <https://doi.org/10.1016%2Fj.ssci.2013.05.009>
- [291] C. A. Monk, D. A. Boehm-Davis, G. Mason, J. G. Trafton, [Recovering from interruptions: Implications for driver distraction research](#), Human Factors: The Journal of the Human Factors and Ergonomics Society 46 (4) (2004) 650–663. doi:10.1518/hfes.46.4.650.56816.
URL <https://doi.org/10.1518%2Fhfes.46.4.650.56816>
- [292] D. Zeng, Z. Yu, L. Xiong, J. Zhao, P. Zhang, Z. Li, Z. Fu, J. Yao, Y. Zhou, [A novel robust lane change trajectory planning method for autonomous vehicle](#), in: 2019 IEEE Intelligent Vehicles Symposium (IV), IEEE, 2019. doi:10.1109/ivs.2019.8814151.
URL <https://doi.org/10.1109%2Fivs.2019.8814151>
- [293] H. Li, Y. Luo, J. Wu, [Collision-free path planning for intelligent vehicles based on bézier curve](#), IEEE Access 7 (2019) 123334–123340. doi:10.1109/access.2019.2938179.
URL <https://doi.org/10.1109%2Faccess.2019.2938179>

- [294] N. Berger, [Lane change path planning: with state-dependent safety constraints](#) (2018).
URL <http://resolver.tudelft.nl/uuid:03aa26dd-e7e6-4ebb-9cd3-def3e169526f>
- [295] J. Ko, R. Guensler, M. Hunter, [Analysis of effects of driver/vehicle characteristics on acceleration noise using GPS-equipped vehicles](#), *Transportation Research Part F: Traffic Psychology and Behaviour* 13 (1) (2010) 21–31. doi:10.1016/j.trf.2009.09.003.
URL <https://doi.org/10.1016%2Fj.trf.2009.09.003>
- [296] H. Park, C. Oh, J. Moon, S. Kim, [Development of a lane change risk index using vehicle trajectory data](#), *Accident Analysis & Prevention* 110 (2018) 1–8. doi:10.1016/j.aap.2017.10.015.
URL <https://doi.org/10.1016/j.aap.2017.10.015>
- [297] R. M. O'brien, [A caution regarding rules of thumb for variance inflation factors](#), *Quality & Quantity* 41 (5) (2007) 673–690. doi:10.1007/s11135-006-9018-6.
URL <https://doi.org/10.1007/s11135-006-9018-6>
- [298] J. M. Murphy, D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, D. A. Stainforth, [Quantification of modelling uncertainties in a large ensemble of climate change simulations](#), *Nature* 430 (7001) (2004) 768–772. doi:10.1038/nature02771.
URL <https://doi.org/10.1038%2Fnature02771>
- [299] MORTH, [Road accidents in india-2019](#) (2019).
URL www.morth.nic.in
- [300] L. Zheng, T. Sayed, A. Tageldin, [Before-after safety analysis using extreme value theory: A case of left-turn bay extension](#), *Accident Analysis & Prevention* 121 (2018) 258–267. doi:10.1016/j.aap.2018.09.023.
URL <https://doi.org/10.1016/j.aap.2018.09.023>
- [301] S. Das, A. K. Maurya, [Defining time-to-collision thresholds by the type of lead vehicle in non-lane-based traffic environments](#), *IEEE Transactions on Intelligent Transportation Systems* 21 (12) (2020) 4972–4982. doi:10.1109/tits.2019.2946001.
URL <https://doi.org/10.1109%2Ftits.2019.2946001>
- [302] S. Piantini, P. M. N. Baldanzini, A. Franci, M. Mangini, A. Peris, [Injury analysis of powered two-wheeler versus other-vehicle urban accidents](#), *IRC-16-102*, 2016.
URL <https://core.ac.uk/download/pdf/301572783.pdf>

List of Publications

Journal Articles:

1. **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Vehicle Path Reconstruction Using Recursively Ensembled Low-Pass Filter (RELP) and Adaptive Tri-Cubic Kernel Smoother.” *Transportation Research Part C: Emerging Technologies* 120 (November 2020): 102847. doi: [10.1016/j.trc.2020.102847](https://doi.org/10.1016/j.trc.2020.102847).
2. **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Trajectory Reconstruction Using Locally Weighted Regression: a New Methodology to Identify the Optimum Window Size and Polynomial Order.” *Transportmetrica A: Transport Science* 14, no. 10 (March 14, 2018): 881–900. doi: [10.1080/23249935.2018.1449032](https://doi.org/10.1080/23249935.2018.1449032).
3. **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Interrupted and Uninterrupted Lane Changes: A Microscopic Outlook of Lane-Changing Dynamics.” *Transportmetrica A: Transport Science*, 2021, doi: <https://doi.org/10.1080/23249935.2021.1965240>.
4. **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Anticipated Collision Time (ACT): A Two-Dimensional Surrogate Safety Indicator for Trajectory-Based Proactive Safety Assessment.” *Transportation Research Part C: Emerging Technologies*, Revision is under review.
5. **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Image-based Naturalistic Driving Data Extraction for the Critical Assessment of Operational Efficiency and Safety of Vehicles.” Working Paper.
6. Bharat Kumar, A. V. A., **Suvin P. Venthuruthiyil**, Shallam, R. D. K., and Mallikarjuna Chunchu. “A Novel Coordinate Transformation Technique for Image Processing Based Trajectory Data Collection from Complex Road Geometries.” , Working paper.

Conference Proceedings:

1. **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Reconstruction of the Vehicle Trajectories for a Meaningful Representation of the Heterogeneous, No Lane-Disciplined Traffic Stream”, *Transportation Research Board 97th Annual Meeting*, 2018, Washington DC, United States.

2. **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “SAVETRAX: A Semi-Automated Image-Processing Based Vehicle Trajectory Extractor” *Transportation Research Board 99th Annual Meeting*, 2020, Washington DC, United States.
3. **Suvin P. Venthuruthiyil**, Mallikarjuna Chunchu, “Association of Maneuverability and Safety of Vehicles in a Congested Urban Road with Weak Lane-Discipline and Vehicle Heterogeneity”, *13th International (Online) Conference on Transportation Planning and Implementation Methodologies for Developing Countries (TPMDC)*, 2020, Mumbai
4. **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Proactive Safety Assessment of 3D Road Geometries Using Naturalistic Driving Data” *Transportation Research Board 101st Annual Meeting*, 2022, Washington DC, United States.



Curriculum Vitae

Education:

2009-2013 : B.Tech from kannur University, Kerala

2014-2016 : M.Tech from Indian Institute of Technology Guwahati

Journal Articles:

1. **Suvín P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Interrupted and Uninterrupted Lane Changes: A Microscopic Outlook of Lane-Changing Dynamics.” *Transportmetrica A: Transport Science*, 2021, doi: <https://doi.org/10.1080/23249935.2021.1965240>.
2. **Suvín P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Vehicle Path Reconstruction Using Recursively Ensembled Low-Pass Filter (RELP) and Adaptive Tri-Cubic Kernel Smoother.” *Transportation Research Part C: Emerging Technologies* 120 (November 2020): 102847. doi: [10.1016/j.trc.2020.102847](https://doi.org/10.1016/j.trc.2020.102847).
3. **Suvín P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Trajectory Reconstruction Using Locally Weighted Regression: a New Methodology to Identify the Optimum Window Size and Polynomial Order.” *Transportmetrica A: Transport Science* 14, no. 10 (March 14, 2018): 881–900. doi: [10.1080/23249935.2018.1449032](https://doi.org/10.1080/23249935.2018.1449032).
4. Shallam, R. D. K., **Suvín P. Venthuruthiyil**, Mallikarjuna Chunchu, and Anjan Kumar Siddagangaiah. “Empirical Analysis of Operating Speed Performance on Undivided Hilly Roads.” *Journal of Transportation Engineering, Part A: Systems* 145, no. 8 (August 2019): 04019034. doi: [10.1061/jtepbs.0000258](https://doi.org/10.1061/jtepbs.0000258).
5. Shallam, R. D. K., **Suvín P. Venthuruthiyil**, Mallikarjuna Chunchu, and Anjan Kumar Siddagangaiah. “Operating Speed Modeling for the rural highways passing through hilly terrain.” *Journal of Transportation Engineering, Part A: Systems*, Vol. 147, Issue 5 (May 2021), doi: [10.1061/JTEPBS.0000524](https://doi.org/10.1061/JTEPBS.0000524).
6. **Suvín P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Modified Generalized Definitions for the Traffic Flow Characteristics Under Heterogeneous, No-Lane Disciplined Traffic Streams.” *Transportation Research Procedia* 34 (2018): 75–82. doi: [10.1016/j.trpro.2018.11.016](https://doi.org/10.1016/j.trpro.2018.11.016).

7. Pal, Dibyendu, **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Characteristics of Vehicular Lateral Shifts in Non-Lane-Disciplined Traffic Stream.” *Transportation Research Procedia* 48 (2020): 3245–3253. doi: [10.1016/j.trpro.2020.08.153](https://doi.org/10.1016/j.trpro.2020.08.153).
8. **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Anticipated Collision Time (ACT): A Two-Dimensional Surrogate Safety Indicator for Trajectory-Based Proactive Safety Assessment.” *Transportation Research Part C: Emerging Technologies*, Revision is under review.
9. **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Image-based Naturalistic Driving Data Extraction for the Critical Assessment of Operational Efficiency and Safety of Vehicles.” *Transportation Research Part C: Emerging Technologies*, Working paper.
10. Bharat Kumar, A. V. A., **Suvin P. Venthuruthiyil**, Shallam, R. D. K., and Mallikarjuna Chunchu. “A Novel Coordinate Transformation Technique for Image Processing Based Trajectory Data Collection from Complex Road Geometries.” *Transportation Letters*, Working paper.

Conference Proceedings:

1. **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Reconstruction of the Vehicle Trajectories for a Meaningful Representation of the Heterogeneous, No Lane-Disciplined Traffic Stream”, *Transportation Research Board 97th Annual Meeting*, 2018, Washington DC, United States.
2. **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “SAVETRAX: A Semi-Automated Image-Processing Based Vehicle Trajectory Extractor” *Transportation Research Board 99th Annual Meeting*, 2020, Washington DC, United States.
3. **Suvin P. Venthuruthiyil**, Diwakar Das, Mallikarjuna Chunchu, “Identification of the Influential Parameters of the Lane-Change Decision-Making for Autonomous Driving using Random-Forest”, 2020, *The 3rd Symposium on Management of Future Motorway and Urban Traffic Systems*, Luxembourg
4. **Suvin P. Venthuruthiyil**, Mallikarjuna Chunchu, “Association of Maneuverability and Safety of Vehicles in a Congested Urban Road with Weak Lane-Discipline and Vehicle Heterogeneity”, *13th International (Online) Conference on Transportation Planning and Implementation Methodologies for Developing Countries (TPMDC)*, 2020, Mumbai
5. **Suvin P. Venthuruthiyil**, Sapan Tiwari, Pratiksha J. Choudhury, Jerifa Zaman, Mallikarjuna Chunchu (2019), “Empirical Investigation of the Existence of Class-Wise Fundamental Diagram for a Heterogeneous Traffic Stream”, *15th World Conference on Transport Research*, 2019, Mumbai
6. **Suvin P. Venthuruthiyil**, Ankita R. Karmakar, Sabina Y. Islam, Mallikarjuna Chunchu, “Criteria for Temporal Aggregation of the Traffic Data from a Heterogeneous Traffic Stream”, *15th World Conference on Transport Research*, 2019, Mumbai

7. Dibyendu Pal, **Suvin P. Venthuruthiyil**, Mallikarjuna Chunchu, “Characteristics of Vehicular Lateral Shifts in Non-lane-disciplined Traffic Stream”, *15th World Conference on Transport Research*, 2019, Mumbai
8. **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Modified Generalized Definitions for the Traffic Flow Characteristics Under Heterogeneous, No-Lane Disciplined Traffic Streams”, *International Symposium of Transport Simulation (ISTS’18) and the International Workshop on Traffic Data Collection and its Standardization (IWTDCS’18) Emerging Transport Technologies for Next Generation Mobility*, 2018, Ehime, Japan.
9. Shallam, R. D. K., **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Impact of the Direction of the Horizontal Curves on the Operating Speed Performance of the Vehicles on Hilly Terrain”, *International Symposium of Transport Simulation (ISTS’18) and the International Workshop on Traffic Data Collection and its Standardization (IWTDCS’18) Emerging Transport Technologies for Next Generation Mobility*, 2018, Ehime, Japan.
10. **Suvin P. Venthuruthiyil**, and Mallikarjuna Chunchu. “Proactive Safety Assessment of 3D Road Geometries Using Naturalistic Driving Data” *Transportation Research Board 101st Annual Meeting*, 2022, Washington DC, United States. *Under review*