


UNIVERSITY OF BRESCIA

PERSONALDATA

WeAreProject






Finanziato dall'Unione Europea
NextGenerationEU

*"Progetto Finanziato dall'Unione europea – NextGenerationEU"*

Dottorato in Ingegneria Civile, Ambientale, Della Cooperazione Internazionale e di Matematica, University of Brescia, Brescia, Italy

Curriculum in Urban Planning and Mobility

---

Scientific Disciplinary Sector: CEAR-03/B

XXXVIII Cycle

# Multi-Scale Crash Risk Estimation Using Naturalistic Driving Data: A Statistical and AI-Based Framework

**Advisor :**

Prof. Giulio Maternini & Prof.Benedetto Barabino

**Co-Advisor:**

Eng. Roberto Ventura(Ph.D.)

**PhD Candidate:**
Nuhamin Gezehagne Assefa

**Table of Contents**

**1. CHAPTER 1** ..... 1

**INTRODUCTION** ..... 1

1.1 Research Context and Problem Statement ..... 1

1.2 Research Objectives ..... 3

1.3 Research Contributions ..... 3

1.4 Thesis Structure ..... 4

2. CHAPTER 2 ..... 7

STATE OF THE ART ..... 7

2.1. Outline of topics ..... 7

2.2. Research Question Development ..... 10

2.2.1. Main Research Question ..... 10

2.2.2. Sub-questions ..... 10

2.3. Research Methods ..... 10

2.4. Summary results ..... 12

2.4.1. General statistics about the review ..... 12

2.4.2. Characteristics of the records ..... 14

2.5. Current knowledge base ..... 16

2.5.1. Naturalistic Driving Studies (NDS) and methodological advancements ..... 18

2.5.2. Crash Risk Estimation and Targeted Safety Applications Using Naturalistic Driving Studies 21

2.5.3. Assessing Driving Safety: The Impact of Driver Behavior and Environmental Factors 26

2.5.4. NDSs for Enhanced Incident Management and Traffic Safety ..... 29

2.6. Research Trends, Gaps and Future research orientations ..... 33

2.6.1. Research Trends ..... 33

2.6.2. Gaps in the current state of the art ..... 34

2.6.3. Future research orientations ..... 35

3. CHAPTER 3 ..... 38

METHODOLOGY ..... 38

3.1. Methodological Framework Overview ..... 38

3.1.1 Behavioral safety theory ..... 38
3.1.2 Risk-based road safety assessment ..... 39
3.2. Structure of the Methodology Chapter ..... 42
3.3. Phase 1: Telematics Data Engineering and Behavioral Event Construction ..... 44
3.3.1. Step 1 - Data Collection and Pre-processing ..... 44
3.3.2. Step 2 – Extracting and Geo-referencing Safety Critical Events (SCEs) ..... 52
3.3.3. Step 3 Event–Road–Crash Integration ..... 59
3.3.4. Step 4 Exposure Estimation ..... 62
3.4. Phase 2 – Parallel Modelling of Crash Frequency and Crash Severity ..... 64
3.4.1. Econometric Modelling of Crash Frequency (Negative Binomial Model) ..... 66
3.4.2. Model Structure, Functional Form, and Treatment of Exposure ..... 68
3.4.3. Econometric Modelling of Crash Severity (Binomial Logit Model) ..... 71
3.4.4. Neural Network-Based Modelling of Crash Frequency and Severity ..... 74
3.5. Phase 3: Integrated Risk Index Development, Normalization, and Spatial Risk Mapping 80
3.5.1. Exposure Treatment and Risk Formulation ..... 81
3.5.2. Composite Risk Index Construction ..... 82
3.5.3. Risk Normalization and Road Network Ranking ..... 84
3.5.4. Spatial Risk Mapping and GIS-Based Decision Support ..... 86
4. CHAPTER 4 ..... 89
REAL CASE EXPERIMENT ..... 89
4.1. Case study description ..... 89
4.2. Phase 1 Results – Telematics Data Processing and Safety-Critical Event Extraction ..... 92
4.2.1. Data collection and preparation ..... 92
4.2.2. Safety-Critical Event Extraction ..... 96
4.2.3. Event–Road–Crash Integration and Multi-Layer Data Assembly ..... 99
4.2.4. Results of Traffic Exposure Estimation and Scaling ..... 101
4.3. Phase 2 Results: Parallel Modelling of Crash Frequency and Crash Severity ..... 104
4.3.1. Crash Frequency Modelling Results ..... 105
4.3.2. Crash Severity Modelling Results ..... 110
4.3.3. Integrated Interpretation of Frequency and Severity Streams ..... 118
4.4. Phase 3: Integrated Risk Index Development, Normalization, and Spatial Risk Mapping 122
4.4.1. Composite Crash Risk Index Estimation ..... 122

4.4.2. Econometric Model–Based Composite Risk ..... 122
4.4.3. FNN-Based Composite Risk ..... 123
4.4.4. Risk Normalization and Road Network Ranking Results ..... 124
4.4.5. Econometric Model–Based Risk Classification ..... 125
4.4.6. FNN-Based Risk Classification ..... 125
4.4.7. Spatial Distribution of Composite Crash Risk Across the Road Network ..... 126
4.4.8. Econometric Model–Based Spatial Risk Distribution ..... 127
4.4.9. FNN-Based Spatial Risk Distribution ..... 129
5. CHAPTER 5 ..... 132
DISCUSSION ..... 132
5.1. Behavioral indicators as central determinants of crash risk ..... 132
5.2. The role of infrastructure, context, and exposure in shaping crash occurrence and consequences ..... 135
5.3. Comparing econometric and artificial intelligence models ..... 138
5.4. Integrated Interpretation of Crash Frequency and Severity Streams ..... 138
5.5. Interpretation of the Composite Crash Risk Index at Segment and Trip Levels ..... 147
6. CHAPTER 6 ..... 152
CONCLUSIONS, CONTRIBUTIONS, AND FUTURE RESEARCH DIRECTIONS ..... 152
6.1. Addressing the Research Gaps ..... 152
6.2. Key Findings and Methodological Contribution ..... 153
6.3. Practical Implications for Road Safety and Route Choice ..... 153
6.4. Limitations and Future Research Directions ..... 155
REFERENCES ..... 157
APPENDICES ..... 168
Appendix A: detailed time-series visualizations illustrating the threshold-based detection process ..... 168
Appendix B :Categorized Explanatory Variables Table ..... 169
Appendix C (Figures C1–C2 and Tables C1-C2). ..... 171
Appendix D (Tables D1–D2). ..... 173
Appendix E (Tables E1–E2). ..... 174

## List of Tables

Table 2-1: Characteristics of the records ..... vi
Table 2-2: Studies on Naturalistic Driving Studies (NDS) and methodological advancements ..... 18
Table 2-3: Studies on ..... 23
Table 2-4: Key Safety Factors Identified in Naturalistic Driving Studies ..... 27
Table 2-5: Studies on Assessing Driving Safety ..... 27
Table 2-6: Studies on NDSs for Enhanced Incident Management and Traffic Safety ..... 31
Table 3-1: Five level crash risk ranking scale based on quartile thresholds and IQR extension ..... 87
Table 4-1: Threshold Values for Extraction of SCEs. ..... 96
Table 4-2: Negative Binomial Regression Results for Crash Frequency Estimation at Segment and Trip Levels ..... 106
Table 4-3: Comparison of Crash Frequency Model Performance ..... 110
Table 4-4: Binary Logistic Regression Results for Crash Severity Estimation at Segment and Trip Levels ..... 113
Table 4-5: Comparison of Crash Severity Model Performance ..... 117
Table 4-6: Descriptive Statistics of the Composite Crash Risk Index (Econometric Model) ..... 123
Table 4-7: Descriptive Statistics of the Composite Crash Risk Index (FNN Model) ..... 124
Table 4-8: Quantile-Based Classification of Normalized Crash Risk (Econometric Model) ..... 125
Table 4-9: Quantile-Based Classification of Normalized Crash Risk (FNN Model) ..... 126

## List of Figures

Figure 2-1: PRISMA 2020 flow diagram for new systematic reviews which included searches of databases, registers and other sources ..... 14
Figure 3-1: Safety Pyramid (Heinrich (1941); Bird (1984) adapted to road safety) ..... 39
Figure 3-2: European road safety target hierarchy (SUNflower pyramid) ..... 40
Figure 3-3: Conceptual analytical pathway of methodology ..... 42
Figure 3-4: Full Methodological Flowchart ..... 43
Figure 4-1: Main Road network of the Province of Brescia considered in the study ..... 91
Figure 4-2: Direction and Curvature Characteristics: (a) Heading Rose and (b) Curvature CDF. ..... 94
Figure 4-3: Total Distribution of Safety-Critical Events ..... 99

Figure 4-4: Segment-Level Traffic Exposure: (a) Reconstructed AADT Classes and (b) Raw Log Transformed Exposure ..... 102
Figure 4-5: Trip Level Exposure Distributions: (a) Mean Relative Exposure and (b) Maximum Relative Exposure ..... 103
Figure 4-6: SHAP Global Feature Importance for the FNN Model (Top 25 Variables). ..... 108
Figure 4-7: Behavioral Variable Importance Based on Permutation Analysis (FNN Model). ..... 109
Figure 4-8: Behavioral Risk Factor Importance: (a) Permutation Importance and (b) SHAP Global Importance. ..... 109
Figure 4-9: SHAP Global Feature Importance for the Segment-Level FNN Severity Model (Top 25 Variables). ..... 115
Figure 4-10: Permutation Importance for the Segment-Level FNN Severity Model (Top 20 Variables). ..... 116
Figure 4-11: Behavioral Risk Factor Importance: (a) Permutation Importance and (b) SHAP Global Importance. ..... 117
Figure 4-12: Spatial Mapping of Segment-Level Crash Risk Categories (Econometric Model) .... 128
Figure 4-13: Spatial Mapping of Trip-Level Crash Risk Categories (Econometric Model) ..... 129
Figure 4-14: Spatial Mapping of Segment-Level Crash Risk Categories (FNN Model)..... 130
Figure 4-15: Spatial Mapping of Trip-Level Crash Risk Categories (FNN Model)..... 131

## Acknowledgment

First and foremost, I would like to thank Almighty God for giving me the strength, health, wisdom, and determination to successfully complete this PhD journey. Without His guidance and blessings, this achievement would not have been possible.

Furthermore, I am deeply grateful to my parents and siblings for their endless love, encouragement, sacrifices, and unwavering support throughout my education. Their belief in me has been a constant source of motivation. In addition, I would like to thank my friends for their understanding, encouragement, and companionship during both the challenging and rewarding moments of this journey.

Moreover, I would like to express my sincere gratitude to my supervisor, Professor Benedetto Barabino, for his guidance, patience, continuous support, and valuable feedback throughout this research. I am also thankful to Professor Giulio Maternini for his advice, encouragement, and academic support. Likewise, I would like to thank Eng. Roberto Ventura and Eng. Martina Carra for their collaboration, technical assistance, and insightful discussions, which greatly contributed to the success of this work.

Finally, this thesis was carried out as part of the PhD program at the Department of Civil, Environmental, Architectural Engineering and Mathematics (DICATAM) of the University of Brescia. The research was funded by the European Union – Next Generation EU and conducted in collaboration with FASTERNET S.r.l., now part of PERSONAL DATA S.r.l. Therefore, I would like to sincerely acknowledge the University of Brescia, the partner company, and all those who supported me throughout this research journey.

## Abstract

Road traffic crashes remain a major global public health challenge, causing approximately 1.35 million deaths annually and generating substantial social and economic costs. Traditional road safety analyses rely primarily on historical crash records to identify hazardous locations; however, such approaches are inherently reactive and often constrained by underreporting, spatial inaccuracies, and the rarity of crash events. Moreover, conventional safety performance models may not fully capture the complex interactions among driver behavior, roadway infrastructure, traffic exposure, and environmental conditions. Recent advances in vehicle telematics and sensing technologies have enabled the collection of Naturalistic Driving Data (NDD), creating new opportunities for proactive road safety assessment. Nevertheless, the integration of NDD into comprehensive crash risk modelling frameworks remains limited, particularly for crash severity analysis.

To address these gaps, this dissertation develops a behavior-driven and exposure-adjusted crash risk estimation framework that integrates Naturalistic Driving Data with roadway infrastructure characteristics, environmental conditions, and traffic exposure variables. Safety-Critical Events (SCEs), including sudden braking, rapid acceleration, harsh turning, lane departures, and speeding, are extracted from vehicle trajectories and used as behavioral indicators of elevated crash risk. These indicators are combined with infrastructural and contextual variables to model crash frequency and crash severity at both segment and trip levels, thereby capturing complementary dimensions of safety: segment-level analysis supports infrastructure-based hotspot screening, while trip-level analysis captures behavioral risk dynamics during individual journeys.

The framework is applied to the road network of the Province of Brescia, Northern Italy, using telematics data collected from approximately 1,230 instrumented vehicles, generating about 730,000 geo-referenced observations (455,787 valid observations after preprocessing), integrated with crash records, roadway inventory, and traffic exposure datasets. A hybrid modelling strategy combining econometric and artificial intelligence methods is employed. For crash frequency, the segment-level Negative Binomial model achieved MAE ≈ 0.48, RMSE ≈ 1.07, and pseudo-$R^2$ ≈ 0.065, while the Feedforward Neural Network (FNN) improved prediction accuracy (MAE ≈ 0.37; RMSE ≈ 0.82). At trip level, econometric models yielded RMSE ≈ 15.05 and MAE ≈ 3.40, whereas FNN models substantially improved performance (RMSE ≈ 3.82; MAE ≈ 2.52; $R^2$ ≈ 0.31). For crash severity, Binary Logistic Regression demonstrated good predictive capability at segment level (AUC ≈ 0.81; balanced accuracy ≈ 0.67; Brier score ≈ 0.18), while the FNN classifier substantially improved performance (AUC ≈ 0.95; balanced accuracy ≈ 0.89; Brier score ≈ 0.06). At trip level, the Binary Logit model achieved AUC ≈ 0.80 and balanced accuracy ≈ 0.74, while the trip-level FNN model achieved AUC ≈ 0.81 with Brier score ≈ 0.16. Across both scales, speeding consistently emerged as the strongest determinant of severe crash outcomes, together with lane departure, vehicle instability, roadway configuration, and collision type. Explainable AI techniques, including SHAP values and permutation importance, were used to identify dominant predictors.

The modelling outputs are integrated into a Composite Crash Risk Index combining frequency, severity, and exposure components. Risk scores are normalized, classified, and spatially mapped using Geographic Information Systems to identify high-risk roadway segments and hazardous driving trajectories. The results demonstrate that behavioral indicators derived from Naturalistic Driving Data significantly enhance the identification of hazardous conditions beyond traditional crash-based analyses. Practically, the developed risk index provides a decision-support tool for infrastructure prioritization, targeted enforcement, route-risk screening, and real-time safer route guidance, allowing users or navigation systems to compare routes not only

by travel time or monetary cost, but also by predicted safety risk. Overall, this research contributes a unified and scalable framework for behavior-informed road safety assessment and network-level risk management.

**Keywords:** Road safety; Naturalistic Driving Data; Crash risk modelling; Safety-Critical Events; Artificial intelligence.

**Abstract (Italian)**

Gli incidenti stradali continuano a rappresentare una delle principali sfide globali per la salute pubblica, causando circa 1,35 milioni di decessi all'anno e generando notevoli costi sociali ed economici. Le analisi tradizionali sulla sicurezza stradale si basano principalmente sui dati storici relativi agli incidenti per individuare i punti pericolosi; tuttavia, tali approcci sono intrinsecamente reattivi e spesso limitati dalla mancata segnalazione degli incidenti, da imprecisioni spaziali e dalla rarità degli eventi. Inoltre, i modelli convenzionali di valutazione della sicurezza potrebbero non cogliere appieno le complesse interazioni tra il comportamento dei conducenti, le infrastrutture stradali, l'esposizione al traffico e le condizioni ambientali. I recenti progressi nella telematica dei veicoli e nelle tecnologie di rilevamento hanno consentito la raccolta di dati di guida naturalistici (NDD), creando nuove opportunità per una valutazione proattiva della sicurezza stradale. Tuttavia, l'integrazione dei dati NDD in modelli completi di valutazione del rischio di incidente rimane limitata, in particolare per l'analisi della gravità degli incidenti.

Per colmare queste lacune, la presente tesi sviluppa un quadro di riferimento per la stima del rischio di incidente basato sul comportamento e corretto in base all'esposizione, che integra i dati di guida naturalistica con le caratteristiche dell'infrastruttura stradale, le condizioni ambientali e le variabili relative all'esposizione al traffico. Gli eventi critici per la sicurezza (SCE), tra cui frenate brusche, accelerazioni rapide, curve brusche, uscite di corsia e eccesso di velocità, vengono estratti dalle traiettorie dei veicoli e utilizzati come indicatori comportamentali di un elevato rischio di incidente. Questi indicatori vengono combinati con variabili infrastrutturali e contestuali per modellare la frequenza e la gravità degli incidenti sia a livello di tratto che di viaggio, cogliendo così dimensioni complementari della sicurezza: l'analisi a livello di tratto supporta lo screening degli hotspot basato sulle infrastrutture, mentre l'analisi a livello di viaggio rileva le dinamiche di rischio comportamentale durante i singoli spostamenti.

Il quadro metodologico è stato applicato alla rete stradale della Provincia di Brescia, nel Nord Italia, utilizzando dati telematici raccolti da circa 1.230 veicoli strumentati, generando circa 730.000 osservazioni georeferenziate (455.787 osservazioni valide dopo la pre-elaborazione), integrate con registri di incidenti, inventario stradale e set di dati sull'esposizione al traffico. È stata impiegata una strategia di modellizzazione ibrida che combina metodi econometrici e di intelligenza artificiale. Per la frequenza degli incidenti, il modello binomiale negativo a livello di segmento ha raggiunto un MAE $\approx$ 0,48, un RMSE $\approx$ 1,07 e uno pseudo-$R^2$ $\approx$ 0,065, mentre la rete neurale feedforward (FNN) ha migliorato l'accuratezza della previsione (MAE $\approx$ 0,37; RMSE $\approx$ 0,82). A livello di viaggio, i modelli

econometrici hanno prodotto un RMSE ≈ 15,05 e un MAE ≈ 3,40, mentre i modelli FNN hanno migliorato sostanzialmente le prestazioni (RMSE ≈ 3,82; MAE ≈ 2,52; $R^2$ ≈ 0,31). Per quanto riguarda la gravità degli incidenti, la regressione logistica binaria ha dimostrato una buona capacità predittiva a livello di segmento (AUC ≈ 0,81; accuratezza bilanciata ≈ 0,67; punteggio di Brier ≈ 0,18), mentre il classificatore FNN ha migliorato notevolmente le prestazioni (AUC ≈ 0,95; accuratezza bilanciata ≈ 0,89; punteggio di Brier ≈ 0,06). A livello di viaggio, il modello logit binario ha raggiunto un AUC ≈ 0,80 e un'accuratezza bilanciata ≈ 0,74, mentre il modello FNN a livello di viaggio ha raggiunto un AUC ≈ 0,81 con un punteggio di Brier ≈ 0,16. Su entrambe le scale, l'eccesso di velocità è emerso costantemente come il fattore determinante più rilevante per gli esiti gravi degli incidenti, insieme all'uscita di corsia, all'instabilità del veicolo, alla configurazione della carreggiata e al tipo di collisione. Per identificare i predittori dominanti sono state utilizzate tecniche di IA spiegabile, tra cui i valori SHAP e l'importanza di permutazione.

I risultati della modellizzazione vengono integrati in un indice composito di rischio di incidente che combina componenti relative a frequenza, gravità ed esposizione. I punteggi di rischio vengono normalizzati, classificati e mappati spazialmente utilizzando sistemi informativi geografici per identificare i tratti stradali ad alto rischio e le traiettorie di guida pericolose. I risultati dimostrano che gli indicatori comportamentali derivati dai dati di guida naturalistica migliorano significativamente l'identificazione delle condizioni di pericolo rispetto alle tradizionali analisi basate sugli incidenti. In pratica, l'indice di rischio sviluppato fornisce uno strumento di supporto decisionale per la definizione delle priorità infrastrutturali, l'applicazione mirata delle norme, lo screening dei rischi lungo i percorsi e la guida in tempo reale verso percorsi più sicuri, consentendo agli utenti o ai sistemi di navigazione di confrontare i percorsi non solo in base al tempo di percorrenza o al costo monetario, ma anche in base al rischio di sicurezza previsto. Nel complesso, questa ricerca contribuisce a creare un quadro unificato e scalabile per la valutazione della sicurezza stradale basata sul comportamento e la gestione del rischio a livello di rete.

**Parole chiave**: Sicurezza stradale; Dati di guida naturalistici; Modellizzazione del rischio di incidenti; Eventi critici per la sicurezza; Intelligenza artificiale.

# 1. CHAPTER 1
# INTRODUCTION

## 1.1 Research Context and Problem Statement

Sustainable mobility plays a central role in the global discussion on sustainable development. However, achieving sustainable mobility requires ensuring adequate levels of road safety. Road traffic crashes represent a major public health, social, and economic challenge worldwide. According to the World Health Organization, approximately 1.35 million people die each year due to road traffic crashes, and road injuries remain the leading cause of death among individuals aged 5–29 years (WHO, 2023). In addition to the human toll, road crashes impose substantial economic costs, estimated to range between 0.4% and 4% of national Gross Domestic Product (GDP) across countries (WHO, 2019; Wijnen et al., 2016). From a sustainable development perspective, these losses highlight the need for proactive and data-driven strategies capable of preventing crashes before they occur.

Since its establishment in 1992, the European Union (EU) has demonstrated a strong commitment to improving road safety standards across its Member States. Over the past decades, several policy initiatives and strategic frameworks have been introduced with the objective of significantly reducing road fatalities and serious injuries (ETSC, 2019). In particular, the EU has adopted ambitious targets to halve the number of road deaths and serious injuries every decade, contributing to the long-term goal of Vision Zero, which aims to eliminate fatalities and serious injuries from the transport system (Passmore et al., 2019). As a result of these efforts, European roads are now considered among the safest in the world. Indeed, the number of road deaths across EU Member States decreased by approximately 37% between 2010 and 2020, while serious injuries declined by about 14% during the same period (Paramasivan et al., 2022).

Despite these improvements, the burden of road crashes remains significant. According to Eurostat, more than 22,000 people lost their lives in road crashes across the European Union in 2019, and approximately 1.23 million injuries were recorded (Eurostat, 2021). Furthermore, the notable reduction in road fatalities observed in 2020 cannot be fully attributed to structural safety improvements. Instead, it largely resulted from the reduction in mobility associated with COVID-19 travel restrictions (Paramasivan et al., 2022). Consequently, most EU Member States did not achieve the target of halving road deaths between 2010 and 2020, highlighting the need for renewed efforts and innovative approaches to improve road safety performance.

In response to these challenges, the European Commission adopted the Strategic Action Plan for Road Safety in 2018, promoting the adoption of the Safe System approach (ETSC, 2019). This approach recognizes that road crashes are not solely caused by human error but rather emerge from the interaction between multiple components of the transportation system, including road users, vehicles, infrastructure, traffic conditions, and environmental factors (Ronan, 2020; Zeng et al., 2024). Therefore, improving road safety requires a systemic perspective that accommodates human fallibility while reducing the likelihood of severe crash outcomes through coordinated interventions across the transport system.

Within this framework, road infrastructure safety has been identified as a key component of crash prevention strategies. Studies indicate that roadway design and roadside conditions contribute to more than 30% of road crashes (Elvik et al., 2009). Consequently, improving road infrastructure safety requires analytical approaches capable of systematically evaluating safety conditions across road networks and identifying locations where safety interventions are most needed. Traditional road safety analyses have relied primarily on historical crash data to identify hazardous locations and prioritize safety interventions. However, crash databases often suffer from underreporting, spatial inaccuracies, and temporal limitations, particularly for minor crashes and near-miss events (Lord & Mannering, 2010). As a result, reactive crash-based analyses may fail to detect emerging safety issues or risky driving conditions before crashes occur.

Recent advances in vehicle sensing technologies and telematics systems provide new opportunities to overcome these limitations. In particular, Naturalistic Driving Data (NDD) enable detailed observation of real-world driving behavior under normal traffic conditions (Dingus et al., 2006). These datasets capture high-resolution information on vehicle trajectories and driver behavior, allowing the identification of Safety-Critical Events (SCEs) such as sudden braking, rapid acceleration, sharp turning, lane departures, and speeding. Such events can be interpreted as behavioral indicators of elevated crash risk and are increasingly used as surrogate safety measures because they occur more frequently than crashes and therefore provide richer information about risky driving conditions (Khattak et al., 2023; Li et al., 2024).

Despite the growing availability of telematics data and Naturalistic Driving Data, their integration into network-level crash risk assessment frameworks remains limited. Many existing studies analyze infrastructure characteristics or driver behavior separately, without integrating these elements within

a unified analytical framework. Furthermore, methodological approaches often rely either on econometric models, which provide interpretable statistical relationships, or machine learning techniques, which can capture complex nonlinear patterns but may lack interpretability. Another important limitation concerns the role of traffic exposure, which represents a fundamental component of crash risk. While crash risk is conceptually defined as a function of crash occurrence, crash severity, and exposure, many existing studies model crash frequency and severity independently, and exposure is rarely incorporated as an explicitly modeled component of risk.

To address these limitations, this research develops a behavior-driven crash risk estimation framework that integrates Naturalistic Driving Data, behavioral safety indicators, roadway infrastructure characteristics, environmental conditions, and traffic exposure variables within a unified analytical structure. The proposed methodology combines econometric (here used interchangeably with statistical) modeling approaches and artificial neural networks to capture both interpretable statistical relationships and complex nonlinear interactions among risk factors. In this dissertation, the terms econometric and statistical are used interchangeably to refer to regression-based statistical modeling techniques applied for crash risk estimation. The predicted crash frequency and crash severity outcomes are subsequently integrated into a Composite Crash Risk Index, enabling the assessment and ranking of safety conditions across the road network.

The framework is applied to the road network of the Province of Brescia (Italy) using telematics data collected from instrumented vehicles together with infrastructure, environmental, and traffic datasets. Through this case study, the research demonstrates how behavioral driving patterns, roadway characteristics, and contextual conditions jointly influence crash risk and how these insights can support proactive, data-driven road safety management and network-level safety assessment.

## 1.2 Research Objectives

The primary objective of this research is to develop and validate a methodological framework for estimating crash risk across a road network by integrating behavioural indicators derived from Naturalistic Driving Data with roadway infrastructure characteristics, environmental conditions, and traffic exposure variables. The framework is designed to support proactive and network-wide road safety assessment by jointly considering crash occurrence, crash severity, and behavioural risk precursors within a unified analytical structure applied at both segment level and trip level.

To achieve this primary objective, the following specific objectives were pursued:

1. To develop procedures for extracting behavioural safety indicators and Safety-Critical Events (SCEs) from Naturalistic Driving Data, enabling the identification of proactive surrogate measures of crash risk at both roadway-segment and individual-trip scales.
2. To model crash occurrence (frequency) at segment level and trip level using econometric and machine-learning approaches that account for roadway characteristics, behavioural indicators, contextual factors, and traffic exposure conditions.
3. To model crash severity at both analytical levels by examining the influence of behavioural, infrastructural, environmental, and crash-related factors associated with serious road safety outcomes.
4. To integrate crash frequency and crash severity predictions into a Composite Crash Risk Index, capable of representing overall safety performance across the road network and across individual driving trajectories in a unified and operational form.
5. To apply the proposed framework for network-wide crash risk assessment, trip-level behavioral risk evaluation, spatial risk mapping, and decision support, identifying high-risk road segments and unsafe driving trajectories for proactive intervention planning.

Together, these objectives establish a coherent multi-scale pathway linking behavioural data extraction, predictive modelling, risk integration, and practical implementation for evidence-based road safety management. By explicitly combining segment-level infrastructure risk assessment with trip-level behavioral risk analysis, the study provides a more comprehensive understanding of crash risk dynamics than approaches limited to a single analytical scale.

### 1.3 Research Contributions

This research contributes to the advancement of crash risk analysis and proactive road safety management by addressing several key methodological and practical challenges identified in the contemporary road safety literature. Recent studies increasingly emphasize the transition from reactive crash-based analysis toward proactive, telematics-driven, and artificial intelligence-supported safety management frameworks. In particular, recent research has demonstrated the growing capability of Naturalistic Driving Data (NDD), connected vehicle systems, and machine learning approaches in identifying risky driving behaviour, predicting crash occurrence, and supporting proactive interventions. For example, Moosavi and Ramnath developed a context-aware telematics

framework for real-time driver risk estimation, while Roussou et al. applied machine learning techniques to investigate behavioural dynamics among European drivers. Similarly, Osman and Hajij demonstrated the effectiveness of advanced deep learning techniques for crash prediction, and Golestani et al. highlighted the value of interpretable machine learning approaches integrating behavioural, spatiotemporal, and weather-related risk factors using telematics big data. Other recent studies further explored telematics-based behavioural analysis, proactive crash prediction, and AI-supported road safety frameworks (Berhanu et al., 2024; Liang et al., 2025; Somvanshi et al., 2026; Yedilkhan et al., 2025; Khan et al., 2026). However, despite these important advances, much of the existing literature remains focused primarily on driver-level behavioural prediction, localized hotspot analysis, crash severity estimation, or isolated machine learning applications without jointly integrating crash frequency, crash severity, roadway infrastructure, environmental context, and traffic exposure within a unified multi-scale analytical framework.

In response to these limitations, the present study investigates how behavioral indicators derived from Naturalistic Driving Data can improve the estimation of crash risk across road networks, how econometric and artificial intelligence models compare in predicting crash frequency and crash severity, and how crash occurrence, crash severity, and traffic exposure can be integrated within a unified risk assessment framework. To address these questions, the research develops a behavior-driven crash risk estimation framework that integrates Naturalistic Driving Data, roadway infrastructure characteristics, environmental conditions, and traffic exposure variables. The proposed methodology combines econometric modeling techniques and artificial neural networks to capture both interpretable statistical relationships and complex nonlinear interactions among crash risk factors. Furthermore, the study formulates a Composite Crash Risk Index that integrates crash frequency and crash severity predictions to provide a comprehensive measure of safety performance across the road network. Unlike many previous studies that focus either on localized roadway segments or isolated behavioural prediction tasks, the proposed framework simultaneously operates at both the segment level and the trip level, thereby capturing infrastructure-related crash risk concentrations together with dynamic behavioral instability during individual driving episodes. The framework is applied to the road network of the Province of Brescia using real-world telematics and infrastructure data, demonstrating its potential as a decision-support tool capable of identifying high-risk locations and driving patterns before crashes occur. By enabling network-wide crash risk assessment and spatial risk mapping, the proposed approach supports proactive road safety

management, more effective prioritization of safety interventions, and the broader transition toward behavior-informed and AI-supported road safety systems.

## 1.4 Thesis Structure

The remainder of this thesis is structured as follows. Chapter 2 presents the state of the art by discussing the theoretical foundations of crash risk analysis, crash prediction models, surrogate safety indicators, and the use of Naturalistic Driving Data in road safety research. Chapter 3 describes the proposed methodological framework for behavior-driven crash risk estimation, including the extraction of behavioral indicators and Safety-Critical Events from telematics data, the development of econometric and artificial intelligence models for crash frequency and severity estimation, and the formulation of the Composite Crash Risk Index. Chapter 4 introduces the study area and datasets used in the analysis and presents the application of the proposed methodology to the road network of the Province of Brescia. Chapter 5 presents and discusses the results obtained from empirical analysis. Finally, Chapter 6 summarizes the main findings of the research, outlines the limitations of the study, and discusses directions for future research.

# 2. CHAPTER 2

# STATE OF THE ART

## 2.1.Outline of topics

Road traffic injuries represent one of the most critical public health and development issues worldwide. According to the World Health Organization (2019), over 1.35 million people are killed each year in road crashes, with an estimated 20–50 million more sustaining non-fatal injuries, many resulting in long-term disabilities. The burden is disproportionately high in low- and middle-income countries (LMICs), where infrastructural gaps, limited enforcement, and rapid motorization exacerbate crash risks (Yu et al., 2021). Conventional safety analyses rely heavily on historical crash reports and aggregate datasets. However, these sources often underrepresent crucial pre-crash behaviors and near-miss events, limiting their utility in proactive risk identification (Abdel-Aty & Pande, 2005; Antin et al., 2019). Traditional crash data typically treat crashes as isolated outcomes, ignoring the dynamic sequences and situational complexity that precede them (Jovanis et al., 2011). Moreover, these datasets seldom to capture Safety-Critical Events (SCEs) such as sudden braking, evasive maneuvers, or distraction episodes, which are often strong precursors to actual crashes (Bagdadi, 2013). As a result, many early warning signs and behavioral risk indicators remain undocumented, hampering the development of predictive models and real-time safety interventions.

In response to the limitations of traditional crash data, Naturalistic Driving Studies (NDS) have emerged as a transformative approach to road safety research. Unlike retrospective methods, NDS allows for the continuous, unobtrusive monitoring of drivers in real-world conditions by outfitting vehicles with advanced data acquisition systems, including GPS, accelerometers, cameras, and radar sensors (Bärgman et al., 2015). This comprehensive data collection enables researchers to observe driver behavior, vehicle motion, and environmental interactions in fine detail, capturing subtle but critical precursors to risk that traditional datasets often miss. Moreover, NDS shifts the focus from documenting crash outcomes to modeling behavioral risk patterns that lead up to safety-critical events, allowing for earlier detection and potential prevention (Arvin & Khattak, 2020). Through this transition, NDS supports the development of real-time risk detection systems, offering a proactive alternative to conventional post-crash analysis.

A defining advantage of NDS lies in its ability to uncover missed or undocumented risks, particularly in complex traffic environments where crashes are avoided due to evasive actions. For example, (Kluger et al. 2016) demonstrated how kinematic data and signal processing techniques (e.g., Discrete Fourier Transform) can identify sudden braking events that never escalate into crashes yet indicate high-risk behavior. Similarly, Ahmed et al. (2022) emphasizes the importance of using trajectory-level data from global NDS programs to inform the design of connected and autonomous vehicle (CAV) systems. Although the study provided a comprehensive synthesis of Naturalistic Driving Studies (NDS), with particular emphasis on driver behavior, crash causation, and implications for autonomous and driverless vehicle systems, the present study differs from that work in several important respects.

Building upon previous syntheses of Naturalistic Driving Studies (NDS), this review extends the existing knowledge base in several important ways. First, it expands the temporal scope of the literature to include studies published from 2005 to 2023, thereby capturing recent methodological developments in machine learning, deep learning, and real-time crash risk modelling. In addition, rather than relying solely on a thematic overview, the review adopts a more explicitly structured methodological typology, organizing the literature according to analytical approaches while introducing the concept of missed or undetected crash risk as a central framework for interpreting NDS-based safety research. Furthermore, it critically examines several issues that have received limited attention in earlier reviews, including the integration of crash frequency and crash severity perspectives, the geographical imbalance in the distribution of NDS studies, and the translation of NDS findings into practical safety technologies.

Recent advances in machine learning (ML) and deep learning (DL) have significantly improved the capacity to model crash risk using NDS data. For example, ensemble algorithms such as Random Forest and Support Vector Machines (SVM) have been widely used to classify driving behaviours and identify precursors to safety-critical events due to their robustness and interpretability (Das & Ahmed, 2021). Similarly, Long Short-Term Memory (LSTM) networks have gained prominence for analysing time-series driving data and capturing temporal dependencies in driver behaviour, enabling improved detection of near-crash patterns (Naji et al., 2022). These approaches allow researchers to uncover nonlinear interactions among key risk factors such as speed volatility, driver distraction, and roadway geometry that are often overlooked by traditional statistical models. Despite these methodological advances, several important limitations remain in existing NDS literature. A substantial proportion of

studies have been conducted in North America, particularly in the United States, creating a geographical imbalance that limits the generalizability of findings to other regions and traffic environments.

In addition, many investigations have been carried out at localized roadway sections rather than at a network-wide scale, which restricts the ability to evaluate safety patterns across entire transport systems. Another important limitation concerns the analytical focus of most studies, which typically emphasize the frequency of safety-critical events (SCEs) while giving comparatively little attention to crash severity and its implications for safety management and policy development. Furthermore, the practical deployment of NDS findings remains relatively underexplored, as much of the existing work has primarily supported the development of Advanced Driver Assistance Systems (ADAS) rather than broader network-level safety assessment or infrastructure planning. Finally, although several studies suggest that Safety-Critical Events (SCEs) can serve as behavioral precursors of crash risk, the empirical validation of a proportional relationship between SCE occurrence and crash frequency or severity remains limited, highlighting the need for more rigorous validation frameworks. Importantly, these challenges largely reflect gaps in primary NDS research rather than deficiencies in existing literature reviews.

Consequently, there remains a need for a systematic synthesis that extends the temporal coverage of studies, organizes the evidence according to methodological approaches including statistical modelling, machine learning, deep learning, and hybrid frameworks and critically evaluates how NDS research addresses missed crash risk, severity integration, geographical disparities, validation processes, and practical deployment challenges. To address this need, the present systematic literature review synthesizes evidence from 93 peer-reviewed articles published between 2005 and 2023, following the PRISMA 2020 framework. The review spans multiple disciplines, including transportation engineering, behavioral science, and data science, and evaluates how NDS can support early detection of crash risks, identifies emerging AI-based modelling approaches, and highlights barriers to broader adoption such as geographical bias, methodological fragmentation, and cost constraints. By doing so, this review advances beyond previous syntheses by providing an extended temporal coverage (2005–2023), a structured methodological classification of studies, and a targeted analysis of missed or undetected crash risks and crash severity considerations within NDS research. Ultimately, the study contributes to the evolving paradigm of road safety research by advocating a transition from reactive crash-based analysis toward proactive, behavior-driven risk modelling, with the aim of informing future research, guiding policy development, and supporting the integration of

NDS-derived insights into intelligent transport systems, particularly in regions where data-driven safety solutions are still emerging.

## 2.2.Research Question Development

The first step in an SLR is to define one or more research questions clearly. In this sense, the main topic of our research is to examine the existing state of the art through a shift from traditional models focusing on random events to Naturalistic Driving Studies.

### *2.2.1. Main Research Question*

How can Naturalistic Driving Studies be systematically synthesized to advance the detection and understanding of missed or undetected crash risks, and which methodological frameworks most effectively support proactive road safety assessment through driver behavior and contextual data analysis?

### *2.2.2. Sub-questions*

1. What types of safety-critical events are most frequently identified using Naturalistic Driving Studies, and how do this surrogate safety indicators complement or differ from traditional crash-based datasets?
2. How can existing studies be categorized according to methodological typologies (e.g., econometric models, machine learning, deep learning, hybrid approaches), and which approaches demonstrate the strongest potential for proactive crash risk prediction?
3. How do driver behaviour patterns and environmental factors such as speed volatility, distraction, weather, or road geometry interact within Naturalistic Driving Studies to explain missed or undetected crash risks?
4. What limitations emerge from the systematic synthesis of the literature, particularly regarding geographic imbalance, integration of crash severity outcomes, and the translation of research findings into deployable safety technologies?

## 2.3.Research Methods

This systematic literature review aims to objectively assess current knowledge and address key research questions outlined in Section 4. The review follows the Systematic Literature Review (SLR) approach, guided by the PRISMA 2020 framework developed by Moher et al. (2009) and revised by

Page et al. (2021). This methodology includes steps specified in PRISMA 2020, particularly checklist items 5 to 15, as outlined in the PRISMA Checklist, ensuring consistency and rigor across disciplines and research perspectives. Articles were selected for this systematic review based on eligibility criteria covering subject relevance, methodology, language, and publication specifics. The focus was on fields such as Civil Engineering, Transport Planning, and Transportation Engineering, which aligned closely with the research questions on crash risk analysis. Various methodologies, quantitative, theoretical, and empirical, were included to provide a broad understanding, with only English-language studies considered for accessibility. This approach aimed at capturing a comprehensive view, facilitating potential progress towards quantitative insights.

The selected articles comprised peer-reviewed journal articles and full-article proceedings (i.e., peer-reviewed conference papers published in official conference proceedings volumes, excluding abstracts, posters, or extended summaries). A computerized search strategy was adopted for efficiency, utilizing Scopus and Web of Science as the primary databases. Additionally, Google Scholar was consulted to ensure comprehensive coverage of the literature. To initiate our paper collection, we conducted keyword searches for terms such as “Missed risk”,” Undetected risk “, “Near crash”, “Vehicle Traffic Trajectory”, “Vehicle travel trajectories”, “Naturalistic Driving Studies”, “Estimation”, “machine learning”, “Data analysis”, “Driver Behavior”, and “Artificial Intelligence” and the search for papers relies on their title, abstract, and keywords. We employ multiple search terms, including various combinations of keywords. The selection criteria for these keywords are as follows: "Missed risk" or "Undetected risk" specifically target risks that have not been identified or detected. These keywords are chosen for their specificity, which helps narrow down search results to studies addressing missed or undetected risks through NDS, aligning closely with the research topic. They are directly relevant to key aspects of the study, including vehicle movement patterns (vehicle traffic trajectory or vehicle travel trajectories), methodologies for analyzing vehicle movements (naturalistic driving studies), and techniques for estimating crash risk (machine learning, data analysis, artificial intelligence). These keywords are selected because they excel at discerning patterns and correlations within complex datasets, a direction emphasized in recent studies (Naji et al., 2022; Sharmin et al., 2023; Arvin et al., 2021). Additionally, keywords such as "Near crash" and "Driver Behavior" offer valuable context for understanding factors contributing to crashes and missed crash risks on the road.

Once we established definitions of key words, we initiated our search in the designated databases. While Scopus and Web of Science retrieved data using titles, Google Scholar allowed searches based on titles and text. However, due to the large number of papers retrieved, only the title and abstract were utilized as search criteria in Google Scholar. The analysis and interpretation of literature on the issue may vary depending on the country or region, as the emphasis on the issue differs among different countries. Hence, we employ "country of affiliation" (indicating scientists' workplace, not nationality) and "country where research is conducted" as geographical indicators derived from the literature under review.

The search spanned from 2005 to 2023 to coincide with the period of widespread adoption and recognition of the naturalistic driving study (NDS) methodology, which serves as the primary approach for this research. The 100-Car Naturalistic Driving Study, conducted by the Virginia Tech Transportation Institute, is a notable instance of naturalistic driving research dating back to the late 1990s. This study stands out as it provides valuable insights into driver behavior and road safety without solely relying on past accidents. However, while there were a few studies conducted in the late 1990s, they were excluded from consideration due to limitations in certainty assessment. These studies predominantly relied on quantitative methods, and overall, the inclusion criteria were designed to encompass a broad range of literature relevant to understanding missed crash risk in road networks through NDS, incorporating diverse methodologies, data sources, and geographical contexts.

## 2.4.Summary results

### *2.4.1. General statistics about the review*

The initial search began in January 2023, utilizing databases and registries, as well as employing methods like Reference Mining, which involves reviewing the reference lists of relevant papers to identify additional studies ("backward citation searching"). After applying the eligibility criteria, a total of 1426 articles were identified. Following the removal of duplicates, 1389 articles remained eligible for screening. The screening process, depicted in Figure 2.1, proceeded according to PRISMA’s statement. During the first screening, based on the title and abstract, a significant number of articles (1156) were excluded due to a lack of relevance to the article's purpose. The initial search included articles containing keywords in their title, keywords, or abstract, resulting in a large volume initially. Exclusion criteria included irrelevant topics, focus on transportation modes other than road vehicles, or unavailability, leading to a total of 233 articles eligible for the second screening.

Before initially screening the 233 articles by reviewing their introductions and conclusions, 25 papers were eliminated because of a lack of full articles, and 111 articles were excluded due to insufficient relevance or methodological flaws. This exclusion was guided by pre-determined eligibility criteria, such as the misalignment with the core research questions (e.g. the studies not using Naturalistic Driving Study data or not considering crash risk or safety-critical events) or methodological reasons, such as the absence of data transparency, ambiguity of modeling frameworks, or empirical validation, which would undermine the quality and reproducibility of research. Subsequently, a more thorough screening of the remaining 102 articles, involving reading each article in its entirety, was conducted. Nine articles were further excluded as they diverged from the chosen topics or failed to provide potential or observed results or methodologies relevant to the basic research questions. Ultimately, 93 articles were considered eligible for analysis, comprising 85 journals and 8 from conference papers.

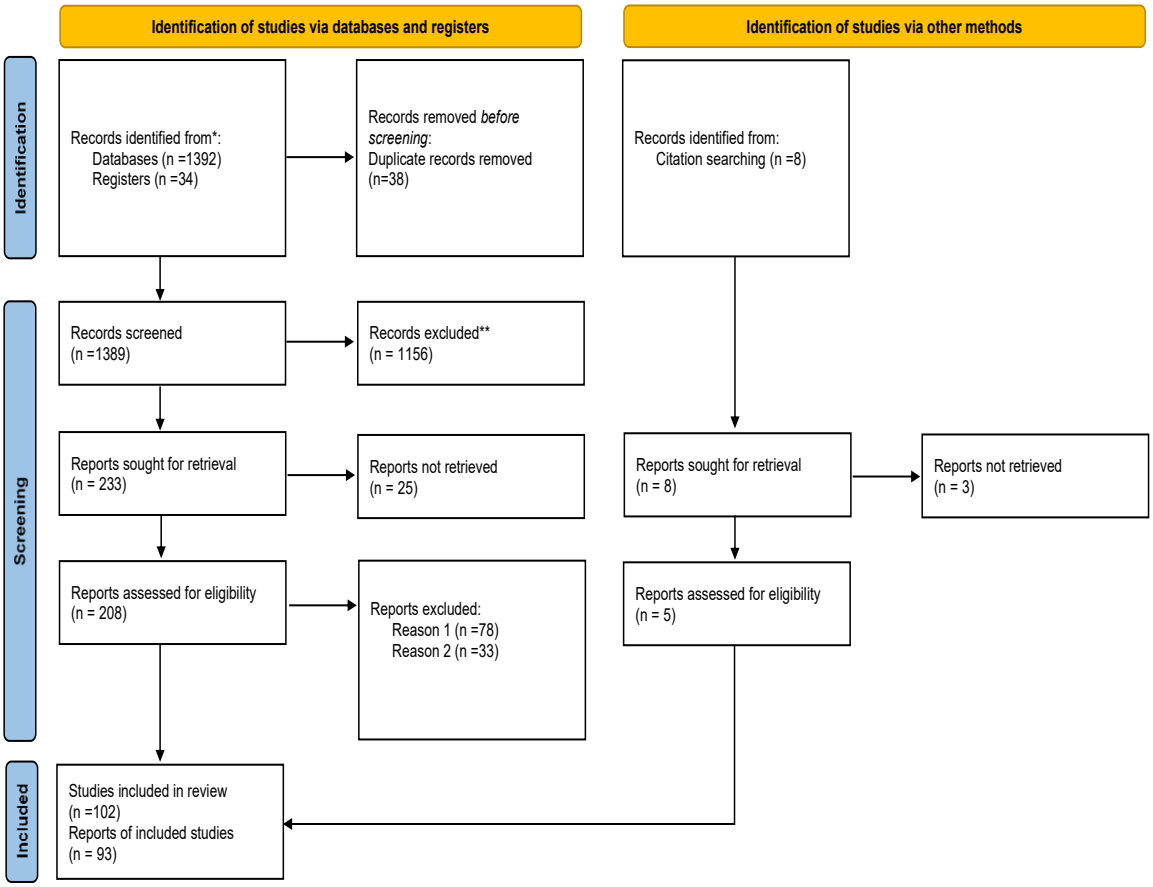



*Figure 0-1: PRISMA 2020 flow diagram for new systematic reviews which included searches of databases, registers and other sources*

### *2.4.2. Characteristics of the records*

The chosen articles exhibit diversity across multiple dimensions, including the journals they were published in, the fields of study they belong to, their publication years, and the countries where the case studies were conducted, as shown in Table 2.1. The predominant journals, accounting for more than 75% of the archived papers, included the Journal of Accident Analysis & Prevention, IEEE Transactions on Intelligent Transportation Systems, Journal of Safety Research, Transportation Research Board, and Transportation Research Part F: Traffic Psychology and Behavior. Additionally, recurrent publications were found in the Journal of Transportation Research Procedia, Institute of Transportation Engineers, International Journal of Scientific & Technology Research (IJSTR), Journal of Road Safety, Results in Engineering, Society of Automotive Engineers (SAE International), Safety Science Traffic Injury Prevention, and Proceedings of Transport Infrastructure and Systems in a changing world (TIS ROMA 2017 & 2019), each with one to two publications. In general, the diverse range of outlets utilized highlights that the journals address a broad spectrum of audiences.

*Table 0-1: Characteristics of the records*

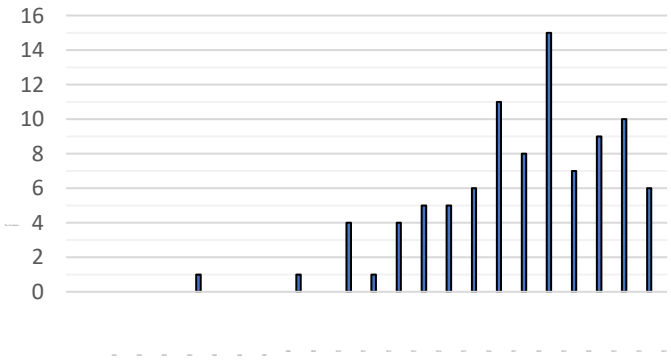
**Field of study**

| Category | No | % |
|---|---|---|
| Civil Engineering | 35 | 38 |
| Technology based research on Transport planning and road safety | 28 | 27 |
| Transportation Engineering | 25 | 30 |
| Mechanical and Industrial Engineering | 5 | 5 |

**Country of case study by region and country**

| | | |
|---|---|---|
| **North America:** | | **60%** |
| | Canada | 2 |
| | USA | 54 |
| **Europe:** | | **19%** |
| | Austria | 2 |
| | France | 1 |
| | Germany | 1 |
| | Italy | 4 |
| | Sweden | 8 |
| | Turkey | 2 |
| **Asia:** | | **14%** |
| | China | 8 |
| | Iran | 2 |
| | Taiwan | 3 |
| **Australia:** | | **6%** |
| | Australia | 6 |

**Publication year**




The oldest article was published in 2005, and the number of articles has steadily increased since 2011. Ninety-five per cent of the articles have been published within the last 12 years, indicating that utilizing NDS studies and using the study for understanding the consequences of overlooking accident risks are relatively recent. Research methodologies vary, as no specific approaches are excluded if the articles meet other criteria. Results indicate that most studies have been conducted in engineering fields and research institutions related to transportation. Regarding the geographical distribution of case studies, a total of 12 countries were identified among the 93 articles that conducted either a single case study or a comparative case study.

These countries were categorized by region, with North America being the most frequently mentioned, accounting for 60% of the total, predominantly the United States (n = 54) and Canada (n=2). Europe

and Asian countries together account for 33% of the total, including Austria (n=2), France (n=1), Germany (n=1), Italy (n=4), Sweden (8), Turkey (2), China (8), Iran (2), and Taiwan (3). Additionally, there were 6 articles from Australia. Most selected articles and case studies are primarily from North America, with the United States contributing to over 58% of the total articles covered. This dominance is attributed to the robust automotive and transportation research sector in the United States, as well as the well-developed infrastructure for data collection and analysis, exemplified by initiatives like the Second Strategic Highway Research Program (SHRP 2). The availability of such data facilitates large-scale studies like naturalistic driving studies (NDS), which focus on understanding real-world driving behavior so that it becomes such a strong ground overlooked risks in the road network. While naturalistic driving studies towards addressing missed accident risks have gained prominence in the U.S., other countries such as Sweden and China also recognize the value of this research approach and are beginning to conduct their own studies. However, the U.S. remains a leader in this field due to its history of investment, large-scale projects, and collaborative efforts in transportation research.

### 2.5.Current knowledge base

This systematic review focuses on the concept of missed crash risk, which refers to the likelihood of traffic incidents that go undocumented, often due to minimal damage or lack of police involvement. By analyzing these events, the review aims to highlight the importance of NDS that traditional crash data might overlook. Considering this, most studies indicate that the previous approach to investigating traffic safety and operations relies heavily on traditional data sources such as spot sensors, loop detectors, and historical crash data. Despite the advancements in traffic surveillance technology and increasing concerns about traffic safety, few studies have established links between real-time traffic flow parameters and crash occurrences, as noted by Abdel-Aty & Pande (2005) and Yu et al. (2024). These traditional data sources may not accurately reflect the actual driving environment at a granular level and could introduce biases due to experimental controls. Addressing these limitations, Naturalistic Driving Studies (NDSs) offer a solution by providing comprehensive real-time driving data to more accurately examine safety issues. For instance, Yu et al. (2021) aim to uncover patterns in freeway loop detector data that may predict traffic crashes, offering fresh insights into road safety. Additionally, the extensive data set from the Second Strategic Highway Research Program's naturalistic driving study has significantly contributed to the use of NDS data in this field, despite being limited to North America.

In a while, with the emergence of NDSs, accident risks started to be investigated more comprehensively by leveraging the detailed and real-time data they provide. Studies can be performed using the various categories identified from NDS data and NDSs, enhancing our understanding of traffic safety across different domains. For instance, this chapter clusters the studies that emerged from the literature review into different categories. The clusters are categorised based on their main focus of analysis, and they are organised in a manner that they directly answer the research questions in the introduction, so that there is minimal overlap and a functional separation between methodological development, risk modelling, contextual analysis, and behavioural assessment. Cluster 1 focuses on naturalistic driving data and methodological advancements, enabling researchers to develop predictive models and optimize driving simulations to preemptively identify and mitigate crash risks. The main distinction of this cluster is that it categorises the studies based on the typologies of models (e.g., econometric models, machine learning, deep learning, and hybrid models), hence answering the research question on the methodological frameworks of proactive risk assessment. Cluster 2, which deals with crash propensity and safety measures, can utilize NDS data to pinpoint specific traffic patterns and behaviors that increase crash likelihood, facilitating the design of targeted safety interventions.

In contrast to Cluster 4, this category focuses on the use of outcome-oriented crash risk estimation and measurement of the risk of a crash or its severity using statistical or predictive models. This cluster in particular isolates the contextual and infrastructural determinants of risk, which can be differentiated as opposed to the driver intrinsic behavioral mechanisms. In Cluster 3, the impact of environmental and external factors on driving safety can be more accurately assessed by analyzing real-world conditions captured by NDS, such as weather impacts and road infrastructure influences. Cluster 4, focusing on driving behavior patterns, can leverage the detailed behavioral data provided by NDS to identify risky driving behaviors and develop corrective measures. As opposed to Cluster 2, which focuses directly on the modelling of crash propensity as an outcome, Cluster 4 focuses on the behavioral mechanisms (distraction, glance behavior, aggression, volatility) leading up to the risk of crash, but not the crash probability, itself. This approach allows for the evaluation of driver responses to various environmental conditions, ultimately enhancing the effectiveness of safety interventions and Advanced Driver Assistance Systems (ADAS). Below are the detailed descriptions of each category.

### *2.5.1. Naturalistic Driving Studies (NDS) and methodological advancements*

Recognizing the importance of NDS, the initial challenge was to develop robust methods and approaches due to the sector's novelty. Over the years, NDS have primarily relied on video and kinematic data to analyze crash, near-crash, and critical-incident events, emphasizing the need for robust statistical methods for effective analysis. The initial studies of NDS were mostly statistically and econometrically based to provide crash and near crash modelizations. To take an example, Jovanis et al. (2011) constructed hierarchical and multilevel models of naturalistic driving event data, which strictly deal with omitted-variable biases and enhance the model robustness. In a comparable way, Arbabzadeh & Jafari (2017) used regularised multinomial logistic regression to create tailored real-time models of safety risks prediction. These methods were focused on interpretability and causal inference, which became the methodological basis of the early NDS-based crash risk model.

The crash and near-crash detection were also largely done with parametric statistical methods. Based on SHRP2 NDS data, Ali et al. (2019) compared logistic regression models with non-parametric ones, which emphasises the advantages and drawbacks of conventional statistical frameworks. These econometric techniques were very important in helping to come up with empirical relationships of driver behaviour, roadway conditions and the crash outcomes. With the growth in size and complexity of NDS datasets, machine learning solutions grew in popularity. Examples of such studies include Das & Ahmed (2021), who used ensemble learning algorithms, such as XGBoost, Random Forest, AdaBoost, and Support Vector Machines, on lane-change prediction with SHRP2 trajectory-level information, with high predictive accuracy. On the same note, Bugusa & Patil (2019) used ensemble prediction algorithms to improve the estimation of crash risks. These machine learning techniques outperformed conventional econometric models in predictive performance, especially in the nonlinear relationships and high-dimensional data generated on the basis of the kinematics of vehicles, roadway and driver characteristics.

Recently, deep learning models have been proposed to learn temporal and sequential driving behaviour dependencies. As shown by Hasan & Naji (2022), Long Short-Term Memory networks excelled in the classification of near-crash risk levels compared to the traditional machine learning models. Huang et al. (2020) also demonstrated that deep neural networks performed better than shallow models in crash detection when real-time radar sensor data were used. The proposed deep learning methods allow automatic feature mining and better modeling of dynamic driving patterns,

which is a substantial methodological breakthrough in proactive risk detection. The driving behavior analysis has also been enlarged by the technological improvements in driver's behavior analysis, increasing the methodological armoury of Naturalistic Driving Studies. Arvin et al. (2019) used the driving volatility concept to measure changes in vehicle dynamics and driver behaviour to create predictive crash risk models using real-time monitoring data. Equally, Cetinkaya & Acarman (2023) developed a driver impairment detection system, which combines both vehicle sensor data and eye-tracking measurements, and show how multi-source data fusion is efficient at-risk detection. These papers explain how traditional statistical models evolve into sophisticated predictive and detection systems that can be used to support real-time safety systems.

In addition to the use of predictive modelling, other methodological developments are signal processing and algorithmic detection systems. Ahlstrom et al. (2011) created more advanced signal processing algorithms to boost the quality of eye- and head-tracking data under naturalistic conditions, and the distraction detection accuracy was enhanced by about 10 percent. Subsequently, Ahlström et al. (2021) proposed AttenD2.0, a context-aware distraction detector algorithm that makes use of various attention buffers that are dependent on driving requirements. These algorithm systems are compatible with real-time monitoring applications and are the basis of integration into Advanced Driver Assistance Systems.

*Table 0-3: Studies on Naturalistic Driving Studies (NDS) and methodological advancements*

| Authors (Year) | Location | Type of Study | Data Sources | Sample Size | Period | Analytical Tool | Relevant Insights |
|---|---|---|---|---|---|---|---|
| **Ahlstrom et al. (2011)** | Gothenburg, Sweden | Signal Processing (Quantitative) | SeMiFOT (Sweden-Michigan Field Operational Test) | 44 drivers, 10,000+ trips | 6+ months | Signal Quality Filtering & Enhancement Algorithms | Improved eye-tracking data quality for naturalistic conditions by 10% in detection accuracy. |
| **Ahlström et al. (2021)** | Linköping, Sweden | Algorithm Development (Quantitative) | n/a | 16 bus drivers | n/a | AttenD2.0 (Multi-Buffer Algorithm) | Context-dependent distraction detection using multiple attention buffers based on driving demands. |
| **Ahmed et al. (2022)** | Wyoming, USA | Systematic Review of NDS (Quantitative) | 117 Global Naturalistic Driving Studies | 117 studies (global review) | 2005-2020 | PRISMA Review Protocol, Meta-analysis | NDS data offers trajectory-level insights into human driving behavior, supporting CAV integration. |
| **Ali et al. (2021)** | Brisbane, Australia | Crash Risk Analysis in Connected Environment (Quantitative) | Driving Simulator " CARRS-Q (Connected Driving) | 78 participants, 3 driving environments | 1 year | Extreme Value Theory (Peak Over Threshold) | Connected environments significantly reduce crash risk during mandatory lane changes, especially with perfect communication. |
| **Ali et al. (2022)** | Brisbane, Australia | Crash Risk Estimation (Quantitative) | Driving Simulator - CARRS-Q | 78 participants | 1 year | Extreme Value Theory (Generalized Pareto) | Connected environments reduce crash risk during mandatory lane changes. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Arbabzadeh & Jafari (2017)** | New Jersey, USA | Predictive Modeling (Quantitative) | SHRP2 Naturalistic Driving Data | n/a | n/a | Elastic Net Regularized Multinomial Logistic Regression | Customized real-time risk models using driver-specific behavior for proactive safety measures. |
| **Bugusa & Patil (2019)** | Arba Minch, Ethiopia & Pune, India | Crash Risk Prediction (Quantitative) | n/a | n/a | n/a | Ensemble Prediction Algorithms | Ensemble models improved crash risk prediction using outcome-based features. |
| **Das & Ahmed (2021)** | Wyoming, USA | Lane-Change Prediction (Quantitative) | SHRP2 NDS + RID (Trajectory-level) | 1,200 lane changes | 3 years | XGBoost, Random Forest, AdaBoost, SVM, etc. | XGBoost achieved 97% accuracy in predicting lane changes using trajectory and sensor data. |
| **Das & Ahmed (2021)** | Wyoming, USA | Lane-Change Prediction (Quantitative) | SHRP2 NDS + RID | 3,400+ drivers | 3 years | XGBoost, RF, AdaBoost, SVM, etc. | XGBoost achieved 97% accuracy in predicting lane changes. |
| **Huang et al. (2020)** | Iowa, USA | Crash Detection & Risk Estimation (Quantitative) | Radar Sensor Data (Interstate 235, Iowa) | n/a | n/a | Deep Learning Models | Deep learning outperformed shallow models for crash detection and prediction. |
| **Hussain et al. (2022)** | Brisbane, Australia | Hybrid ML + EVT Crash Estimation (Quantitative) | n/a | n/a | 4 days (6 AM to 6 PM) | Unsupervised ML (iForest, LOF, SVM) + EVT | Hybrid model using ML improved accuracy in estimating crash risk from conflicts. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Jovanis et al. (2011)** | Virginia, USA | Quantitative Modeling | 100-Car Naturalistic Driving Study | 22 crashes, 36 near-crashes, 170 incidents | 12-13 months | Binary Logit, Multilevel Modeling | Showed omitted-variable bias in context variables; multilevel modeling improved analysis. |
| **Naji et al. (2022)** | Wuhan, China | Risk Classification of Near-Crashes (Quantitative) | n/a | 1690 near-crash events from 41 drivers | n/a | Hierarchical Clustering, Adaptive Lasso Regression, SVM, RF, MLP, Ordinal Probit, Multinomial Logit, LSTM, GRU | Deep learning models (LSTM, GRU) outperformed others in classifying near-crash risk levels. Five distinct risk categories identified: minimal to severe. Key features include speed, braking pressure, road type, and driver experience. |
| **Sharmin et al. (2023)** | Connecticut, USA | Behavioral Psychology & Crash Prediction (Quantitative) | SHRP2 Naturalistic Driving Data | Multiple SHRP2 participants | 1 year | Latent Class Clustering, Mixed-Effects Logistic Regression | Psychological driver profiles significantly impact crash risk under certain road conditions. |
| **Zheng et al. (2021)** | Nanjing, China | Crash Prediction by Segment Type (Quantitative) | n/a | n/a | n/a | Bayesian Logistic Regression (BLR) | Downstream average speed is a strong crash precursor across segment types. |

### *2.5.2. Crash Risk Estimation and Targeted Safety Applications Using Naturalistic Driving Studies*

NDSs have emerged as a pivotal approach for enhancing targeted safety in traffic systems. By capturing real-world driving behavior through continuous monitoring of drivers, vehicles, and environmental conditions, NDSs provide an unparalleled depth of data. This approach allows researchers to pinpoint risk factors and create targeted safety interventions for specific driving situations and behaviors, offering insights beyond what historical data alone can reveal. Consequently, NDSs play a crucial role in advancing our understanding of crash dynamics and improving traffic safety strategies. Building on the theme of improving safety analysis, Dozza and González (2013) explored the automatic processing of driver video information to improve the identification of safety-critical events in NDS data. By analyzing approximately 400 video sequences, they concluded that drivers' individual reactions are key to recognizing safety-critical events. This study highlights the potential of automated video processing to enhance the efficiency of analyzing large-scale NDS data. Expanding on modeling approaches, Shaon et al. (2019) considered both crash frequency and severity to design effective safety improvement programs.

Turning to work zone safety, Bharadwaj et al. (2019) investigated crash risk factors using NDS data. They developed a binary logistic regression model to estimate crash risk, identifying driver inattention and non-driving-related secondary tasks as significant factors increasing the risk of safety-critical events. This study emphasizes the need for targeted interventions to enhance safety in work zones. Additionally, Zurlinden et al. (2020) examined the relationship between complex traffic flow phenomena and crash risk on urban motorways. They identified high traffic density as a major factor increasing crash risk and suggested that modern measurement devices can provide valuable data for predicting crashes and identifying high-risk locations. On the topic of high-resolution traffic sensing data, Yu et al. (2021) emphasized its importance in crash risk analysis. Using the HighD Dataset from German highways, they demonstrated that disturbed traffic flows and close following distances significantly impact high-risk event occurrences. Their study highlighted the need for proactive safety measures to reduce crash likelihood. Changjian Zhang et al. (2021) investigated real-time tyre force data to identify high-risk crash locations on freeway segments with horizontal curvature. Their analysis revealed that longitudinal and lateral stability indicators, obtained from vehicle kinetic experiments, effectively predicted crash risk on high-risk road segments. Additionally, Azizi and Hadi (2021) explored the use of connected vehicle and infrastructure sensor data to enhance traffic

performance estimation and prediction. They proposed using disturbance metrics, such as oscillations and disturbance durations, to improve traffic state and safety predictions, recommending these metrics for decision tools in traffic management. Focusing on improving crash risk prediction efficiency, Qikang Zheng et al. (2021) identified optimal crash precursors for different freeway section types. Their Bayesian logistic regression models demonstrated that downstream average speed is a key crash precursor for all segment types, providing valuable insights for real-time crash risk modeling. Moreover, Das and Ahmed (2021) proposed reliable lane-change prediction models using SHRP2 NDS data. Their study found that the XGBoost model, incorporating vehicle kinematics, machine vision, driver, and roadway features, achieved the highest prediction accuracy and F1-score, suggesting its potential for developing advanced driver assistance systems (ADAS). Hasan & Naji (2022) compared machine and deep learning methods for classifying the risk levels of near crashes. They found that deep learning models, particularly the LSTM model, outperformed traditional machine learning models in classification performance, highlighting the potential of deep learning for improving transportation safety.

Lastly, Ahmed et al. (2022) conducted a systematic review of NDS studies over the past fifteen years, identifying key research topics such as driver behavior, crash causation, and driver distraction. Their review emphasized the importance of NDS data for integrating human behavior into connected and autonomous vehicle (CAV) systems, facilitating the safe implementation of CAV technology. In a related study, Ali et al. (2022) used Extreme Value Theory to examine mandatory lane-changing crash risk in connected and traditional environments. They found that the connected environment significantly reduces crash risk, particularly under perfect communication conditions, demonstrating the benefits of connected vehicle technology for traffic safety. Furthermore, Hussain F. et al. (2022) proposed a hybrid modeling framework combining machine learning and extreme value theory to estimate crash risk from traffic conflicts. They found that the iforest algorithm outperformed other sampling techniques in identifying extremes, suggesting its potential for improving crash risk estimation models. Finally, Sharmin et al. (2023) evaluated driver attributes and roadway environmental factors for predicting safety-critical events using NDS data. Their latent class clustering approach identified different driver types, revealing that models incorporating driver factors more accurately predicted crash risk. This study suggests new avenues for educational and behavioral interventions to improve road safety. In conclusion, this comprehensive literature review synthesizes key findings from various studies, illustrating the significant contributions of Naturalistic Driving Studies in enhancing targeted traffic safety measures.

*Table 0-4: Studies on Naturalistic Driving Studies in enhancing targeted traffic safety measures.*

| Authors (Year) | Location | Type of Study (all Quantitative) | Data Sources | Sample Size | Period | Analytical Tool | Relevant Insights |
|---|---|---|---|---|---|---|---|
| **Zhang & Khattak (2011)** | Virginia, USA | Spatiotemporal Analysis | n/a | n/a | 1 year | Statistical Analysis of Time & Distance Gaps | Analyzed time-distance links in primary-secondary incidents for improved incident management. |
| **Zheng et al. (2014)** | Beijing, China | Cluster Analysis | Chinese Naturalistic Driving Data | Near-crash cases dataset | n/a | K-means Clustering | Clustered near-crashes by braking features; braking speed and trigger factors most correlated with risk. |
| **Abdelrahman et al. (2018)** | Ontario, Canada | Statistical Risk Analysis | 100-Car Naturalistic Driving Study | 829 crash/near-crash, 19,616 baseline events | 2003-2004 | Odds Ratio & Relative Risk | Icy and wet roads under heavy traffic have highest crash risk; confirmed with statistical significance. |
| **Ali et al. (2019)** | Wyoming, USA | Crash Detection | SHRP2 Naturalistic Driving Data | 213 near-crashes (30 rain, 60 clear) | 2018-2019 | Binary Logistic, Decision Tree, ANN | Non-parametric models like ANN and Decision Trees outperformed logistic regression in near-crash detection. |
| **Dozza & Gonzalez (2013)** | Sweden, Göteborg | Video Processing Algorithm | EuroFOT Project (100 Volvo cars) | 400 video sequences | n/a | Automatic Video Classification Algorithms | Video-based recognition can improve automated identification of safety-critical events in NDS. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Rahman Shaon et al. (2019)** | Wisconsin, USA | Multivariate Regression | Crash data (Wisconsin rural highways) | Crash data across severity levels | n/a | Multivariate Multiple Risk Source Regression | Multivariate model better captures distinct engineering and behavioral crash risk sources. |
| **Bharadwaj et al. (2019)** | Missouri, USA | Logistic Regression Modeling | SHRP2 Naturalistic Driving Study | 3,000 drivers, 50M vehicle miles | SHRP2 data collection period | Binary Logistic Regression, ROC | Inattention and secondary tasks over 6s significantly raise crash/near-crash risk in work zones. |
| **Zurlinden et al. (2020)** | Melbourne, Australia | Complexity-Risk Relationship | Traffic & Infrastructure Sensor Data of the city | Urban motorway segments | n/a | Traffic Flow & Event Mapping | Traffic complexity increases cognitive load, linked to higher crash rates in dense traffic. |
| **Yu et al. (2021)** | Shanghai, China | Trajectory-Based Prediction | HighD Dataset (German highways) | High-frequency vehicle trajectories | n/a | Logistic Regression (RPLR, RELR, Standard) | Trajectory-level features (e.g., time to collision) predict high-risk events with up to 97% accuracy. |
| **Zhang et al. (2021)** | Nanjing, China | Real-Time Risk Identification | n/a | Segment-based tire force measurements | n/a | Negative Binomial & Random Effects Models | Tire force stability indicators help identify risk on curved freeway segments with 71.8% accuracy. |
| **Das & Ahmed (2021)** | Wyoming, USA | Lane-Change Prediction | SHRP2 NDS + RID | 3,400+ drivers, 3 years | 3 years | XGBoost, RF, AdaBoost, SVM, etc. | XGBoost achieved 97% accuracy in predicting lane changes. |
| **Ali et al. (2022)** | Brisbane, Australia | Crash Risk Estimation | Queensland Driving Simulator - CARRS-Q | 78 participants | 1year | Extreme Value Theory (Generalized Pareto) | Connected environments reduce crash risk during mandatory lane changes. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Sharmin et al. (2023)** | Connecticut, USA | Behavioral Psychology & Crash Prediction | SHRP2 Naturalistic Driving Data | Multiple SHRP2 participants | 1year | Latent Class Clustering, Mixed-Effects Logistic Regression | Psychological driver profiles significantly impact crash risk under certain road conditions. |
| **Hussain et al. (2022)** | Brisbane, Australia | Hybrid ML + EVT Crash Estimation | n/a | Traffic conflicts at 3 intersections (4 weekdays) | 4 days (6 AM to 6 PM) | Unsupervised ML (iForest, LOF, SVM) + EVT | Hybrid model using ML improved accuracy in estimating crash risk from conflicts. |
| **Qikang Zheng et al. (2021)** | Nanjing, China | Crash Prediction by Segment Type | n/a | Various freeway segment types | n/a | Bayesian Logistic Regression (BLR) | Downstream average speed is a strong crash precursor across segment types. |
| **Ahmed et al. (2022)** | Laramie, Wyoming, USA | Systematic Review of NDS | 117 Global Naturalistic Driving Studies | 117 studies (global review) | 2005-2020 | PRISMA Review Protocol, Meta-analysis | NDS data offers trajectory-level insights into human driving behavior, supporting CAV integration. |
| **Ali et al. (2021)** | Brisbane, Australia | Crash Risk Analysis in Connected Environment | Queensland Driving Simulator - CARRS-Q | 78 participants, 3 driving environments | 1 year | Extreme Value Theory (Peak Over Threshold) | Connected environments significantly reduce crash risk during mandatory lane changes, especially with perfect communication. |

### *2.5.3. Assessing Driving Safety: The Impact of Driver Behavior and Environmental Factors*

The Naturalistic Driving Studies provide detailed real-world information that enable the safety determinants to be investigated in different yet interacting domains of factors. According to the literature reviewed, the influencing variables can be grouped into four main categories, including: (1) roadway and traffic conditions, (2) distraction and glance behaviour, (3) contextual driving environment, and (4) driver emotional and secondary-task factors. To begin with, Abdelrahman A.et al. (2018) leveraged the 100-Car Naturalistic Driving Study (NDS) dataset to analyze traffic and road surface conditions in relation to crash and SCE. They quantified risks using odds ratios and relative risk measures, finding that icy roads with level-of-service (LOS) A, wet roads with LOS D, and dry roads with LOS D have the highest risk for crashes and near-crashes. This study underscores the importance of understanding environmental conditions to improve traffic safety. In a similar Ali E.M. et al. (2019) focused on detecting near-crashes on freeways by comparing environmental conditions and vehicle kinematics signatures of near-crash events to normal driving. Ahlstrom et al. (2011) highlight the significant role of distraction and inattention as primary causes of crashes, emphasizing the need for advanced data processing to improve the detection of distractions in real-world settings. Similarly, Tian et al. (2013) demonstrate a linear relationship between off-road glance duration and crash probability, with traffic density significantly moderating this relationship. This underscores the necessity of considering environmental factors in safety assessments.

Moreover, Tivesten & Dozza (2014) delve into how visual-manual tasks and driving context affect driver glance behavior. They reveal that drivers adapt their glance behavior based on the driving environment, such as during turns or in the presence of other vehicles. This context-dependent behavior is crucial for designing effective driver support systems and evaluating the safety of in-vehicle tasks. In line with this, Bärgman et al. (2015) discuss how naturalistic driving research has reshaped our understanding of driver behavior and crash risk. The insights gained from recording drivers in real-world conditions challenge traditional empirical approaches and highlight the importance of context-specific safety measures. Additionally, the emotional state of drivers also plays a significant role in driving safety. Precht L. et al. (2017) examine the impact of anger on driving behavior using naturalistic driving data, finding that anger leads to more aggressive driving rather

than cognitive overload. This finding is particularly relevant when considering how external provocations can influence driver behavior. Furthermore, Bakhit et al. (2018) analyze the increased crash risk associated with various secondary tasks, identifying high-risk activities like texting and manipulating objects. These findings provide a basis for legislative measures and resource allocation aimed at reducing distraction-related crashes.

*Table 0-5: Key Safety Factors Identified in Naturalistic Driving Studies*

| Factor Cluster | Specific Variables | Reported Effect on Safety | Authors (Year) |
|---|---|---|---|
| **Roadway & Traffic Conditions** | Road surface (icy, wet), level of service, traffic density | Higher crash risk under adverse surface & high-density conditions | Abdelrahman et al. (2018); Ali et al. (2019) |
| **Visual Distraction & Glance Behaviour** | Off-road glance duration, attention shifts | Increased crash probability; moderation by traffic density | Tian et al. (2013); Ahlstrom et al. (2011) |
| **Contextual Driving Environment** | Turning, lead vehicle presence | Adaptive glance behaviour; altered crash risk patterns | Tivesten & Dozza (2014); Bärgman et al. (2015) |
| **Emotional State** | Driving anger | Increased aggressive behaviour and crash likelihood | Precht et al. (2017) |
| **Secondary Tasks** | Texting, object manipulation | Substantially elevated crash/near-crash risk | Bakhit et al. (2018) |

*Table 0-6: Studies on Assessing Driving Safety*

| Authors (Year) | Location | Type of Study (all Quantitative) | Data Sources | Sample Size | Period Covered | Analytical Tool | Relevant Insights |
|---|---|---|---|---|---|---|---|
| **Ahlstrom et al. (2011)** | Gothenburg, Sweden | Signal Processing | SeMiFOT (Sweden-Michigan Field Operational Test) | 44 drivers, 10,000+ trips | 6+ months | Signal Quality Filtering & Enhancement Algorithms | Improved eye-tracking data quality for naturalistic conditions by 10% in detection accuracy. |
| **Tian et al. (2013)** | Indiana, USA | Crash Risk Analysis | VTTI 100-Car Naturalistic Driving Study | n/a | n/a | Linear and Nonlinear | Cumulative off-road glance time increases |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | Probability Modeling | crash risk; traffic density is a moderating factor. |
| **Tivesten & Dozza (2014)** | Gothenburg, Sweden | Glance Behavior Analysis | EuroFOT Swedish Naturalistic Driving Data | n/a | 2013-2014 | Statistical Comparison of Glance Behavior by Context | Driving context (e.g., turns, lead vehicles) affects glance behavior during phone use. |
| **Bärgman et al. (2015)** | Michigan, USA | Counterfactual Simulation | SHRP2 Naturalistic Driving Study | 37 crashes, 186 near-crashes | Data collected before 2015 | "What-if" Counterfactual Simulation | Simulated changes in glance behavior show significant influence on crash/injury risk. |
| **Precht et al. (2017)** | Munich, Germany | Emotion-Based Driving Risk Analysis | SHRP2 Naturalistic Driving Study | n/a | n/a | GLMM and LMM Models | Anger causes deliberate aggressive driving more than errors, increasing crash likelihood. |
| **Bakhit et al. (2018)** | Louisiana, USA | Crash/Near-Crash Risk Assessment | SHRP2 Naturalistic Driving Study | 3,000+ drivers across 6 states | 2010-2013 | Bivariate Probit, Baseline-Category Logit, Rule Mining | Tasks like texting and object manipulation drastically raise crash risk odds. |
| **Cetinkaya & Acarman (2023)** | Istanbul, Turkiye | Driver Impairment Detection | 100-Car Naturalistic Driving Study | 828 trips, 68 crashes, 760 near-crashes | 2003-2004 | Boosted Decision Trees, Sensor Fusion | Impairment can be detected using sensor and glance data with high accuracy. |
| **Arvin et al. (2019)** | Tennessee, USA | Crash Intensity Modeling | SHRP2 Naturalistic Driving Study | 617 crash events, 0.18 million trajectory observations | 2010-2013 | Path Analysis, Tobit Model, Ordered Probit (Fixed & Random Parameters) | Pre-crash driving volatility (speed and deceleration instability) significantly increases crash intensity. Distracted and aggressive driving indirectly raise severity risk. Environmental factors like congestion also contribute. |

### *2.5.4. NDSs for Enhanced Incident Management and Traffic Safety*

Traffic operations research increasingly recognizes incident management and operational safety monitoring as essential components for reducing crash occurrence, minimizing response times, and improving overall network efficiency. Within this context, Naturalistic Driving Studies (NDS) have emerged as an important data source not only for analyzing behavioral or environmental determinants of crash risk but also for supporting real-time incident detection and traffic management systems. By continuously capturing vehicle dynamics and driver behavior in real-world conditions, NDS enables researchers and traffic authorities to identify precursors of hazardous situations and develop operational strategies aimed at improving roadway safety and system performance.

Several studies demonstrate the contribution of NDS to operational safety analysis and incident detection. Habibovic et al. (2013) used naturalistic driving data to investigate driver behavior in car–to–pedestrian conflict situations, contributing to improved incident causation models and the design of context-sensitive safety technologies. Similarly, Zheng et al. (2014) applied clustering techniques to naturalistic driving data to classify traffic incidents more effectively and identify high-risk driving situations within operational traffic environments. From a methodological perspective, Wu & Jovanis (2013) proposed systematic screening procedures for detecting Safety-Critical Events (SCEs) in naturalistic datasets, enabling more reliable identification of crash surrogates that can be used within traffic monitoring and safety management systems. Complementing this work, Kluger et al. (2016) demonstrated that combining kinematic vehicle data with signal-processing techniques improves the detection of safety-critical scenarios, thereby supporting proactive traffic control and incident response strategies.

Further research highlights the role of NDS in the development of advanced safety applications and intelligent transportation systems. For example, Wang et al. (2017) developed a collision-avoidance system using naturalistic driving data to estimate risk levels during lane-change manoeuvres, illustrating the potential of NDS for real-time warning systems. Using the SHRP2 naturalistic driving dataset, Xu (2021) examined ramp-related crash mechanisms and showed that detailed behavioural fault characterization can support the design of targeted operational countermeasures. In addition, Antin et al. (2019) emphasized that large-scale naturalistic datasets provide valuable insights for traffic operations research and can inform system-level safety interventions and operational policies. Collectively, these studies demonstrate that Naturalistic Driving Studies play a crucial role not only

in crash risk modelling but also in operational traffic safety management, including incident detection, traffic monitoring, and the development of intelligent transportation systems. In contrast to Cluster 3, which primarily focuses on the influence of environmental and behavioral determinants of crash risk, this cluster emphasizes system-level applications of NDS, highlighting their potential to enhance traffic management strategies, improve incident response mechanisms, and support the development of proactive safety technologies.

Overall, the reviewed studies demonstrate that the methodological approaches applied in Naturalistic Driving Studies (NDS) and telematics-based road safety research can generally be classified into two major categories: traditional statistical models and Artificial Neural Network (ANN)/machine learning-based approaches. Each methodological group presents distinct advantages, limitations, and areas of applicability depending on the objectives of the study, the nature of the available data, and the intended analytical outcomes. Traditional statistical approaches, including regression-based, probabilistic, and econometric models, are widely utilized due to their interpretability, theoretical transparency, and ability to explicitly quantify the relationships between explanatory variables and safety outcomes. These methods are particularly valuable for identifying significant contributing factors, conducting hypothesis testing, and supporting policy-oriented interpretations through measurable parameter estimates and effect sizes. Furthermore, statistical models are generally less computationally demanding and can perform effectively even with relatively structured and moderate-sized datasets. However, their underlying assumptions and relatively rigid functional structures may limit their ability to fully capture the complex, nonlinear, and dynamic interactions commonly observed in driver behaviour, traffic exposure, roadway characteristics, and environmental conditions within NDS data.

In contrast, ANN-based and other machine learning approaches have increasingly gained attention because of their strong predictive capability and flexibility in handling large-scale, high-dimensional, and nonlinear datasets generated from telematics and sensor technologies. These approaches can better capture hidden behavioural patterns and complex interactions among variables without requiring strict statistical assumptions. Consequently, several studies have reported improved predictive performance when applying ANN-based methods for crash prediction, driver behaviour analysis, and safety-critical event detection. Nevertheless, these approaches also present important limitations, including reduced interpretability, higher computational complexity, sensitivity to

hyperparameter selection and data quality, as well as the potential risk of overfitting. Moreover, the "black-box" nature of many ANN models may limit their direct applicability for policy interpretation and practical decision-making where explainability is essential. Therefore, the reviewed literature suggests that neither methodological approach can be considered universally superior; rather, their suitability depends on the specific research context and objectives. Statistical models remain highly valuable for explanatory analysis and policy interpretation, whereas ANN-based approaches may provide enhanced predictive performance for complex behavioural and telematics-driven datasets. This methodological diversity highlights the growing need for integrated and balanced modelling frameworks capable of combining interpretability with predictive capability in future road safety research.

.

*Table 0-7: Studies on NDSs for Enhanced Incident Management and Traffic Safety*

| Authors (Year) | Location | Type of Study (all Quantitative) | Data Sources | Sample Size | Period Covered | Analytical Tool | Relevant Insights |
|---|---|---|---|---|---|---|---|
| **Habibovic et al. (2013)** | Gothenburg, Sweden & Tsukuba, Japan | Driver-Pedestrian Incident Causation | Naturalistic Driving Study video recordings from Japan | 90 car-to-pedestrian incident videos | n/a | Modified DREAM (Driving Reliability and Error Analysis Method) | Identified driver attention failures and unexpected pedestrian behaviors as key causes. Recommended interactive and autonomous ADAS based on incident context. |
| **Wu & Jovanis (2013)** | Pennsylvania, USA | Event Detection Methodology | 100-Car Naturalistic Driving Study | 63 road departure events | Early 2000s | Statistical screening using kinematic triggers & diagnostics | Screening with kinematic variables (e.g., lateral/longitudinal acceleration) enhances the identification of well-defined surrogate safety events. |
| **Zheng et al. (2011)** | Beijing, China | Driving Risk Classification | Naturalistic Driving Data from China | n/a | n/a | K-means clustering using braking features | Velocity during breaking and event triggers strongly influence driving risk classification in near-crash cases. |
| **Wu & Thor (2015)** | Virginia, USA | Crash Sequence Analysis | Integrated Vehicle-Based Safety System program (IVBSS) | 2 freeway rear-end events (1 crash, 1 near-crash) | n/a | Comparative crash sequence modeling | Highlighted the importance of temporal headway, attention, and driver misjudgment in rear-end crash progression. |
| **Wang et al. (2017)** | USA | Collision Avoidance System Design | 100-Car Naturalistic Driving Study | 2,000,000 vehicle miles; 241 participants | 12-13 months | Fuzzy inference system, OC-SVM, rule-based classifier | Fuzzy system effectively detects danger level during lane changes; combined with fault classifier to predict responsibility. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Xu (2021)** | Nevada, USA | Ramp-Related Crash Analysis | SHRP2 Naturalistic Driving Study | 428 crashes/near-crashes, 100 baseline events | n/a | Behavioral fault classification & statistical analysis | Driver distraction is the primary fault in ramp-related crashes; female drivers and congestion amplify unsafe behavior likelihood. |
| **Antin et al. (2019)** | Virginia, USA | Methodological Framework for NDS | SHRP2 Naturalistic Driving Study | 3,500+ drivers, 32M miles, 5.5M trips | 2010-2013 | Comprehensive instrumentation & observational methodology | Provided detailed and continuous pre-crash behavior data; valuable for traffic operations and risk analysis research. |
| **Arvin & Khattak (2020)** | Tennessee, USA | Crash Risk & Distraction Duration Analysis | SHRP2 Naturalistic Driving Study | 7394 baseline, 1228 near-crashes, 617 crashes | 2010-2014 | Fixed and Random Parameter Logistic Regression | Distractions (texting, dialing) and impairments (alcohol/drugs) significantly increase crash risk. Distraction duration is a key crush factor. |
| **Wu & Wang (2021)** | Taiwan | Rear-End Crash Risk Modeling | SHRP2 Naturalistic Driving Study | 17 crashes, 433 near-crashes | 2010-2015 | Structural Equation Modeling (SEM) | Sequential interaction of traffic flow, driver behavior, perception time, and precipitating events linked to rear-end crash outcomes. |
| **Arvinet et al. (2019)** | Tennessee, USA | Crash Intensity Modeling (Quantitative) | SHRP2 Naturalistic Driving Study | 617 crash events, ~180,000 trajectory observations | 2010-2016 | Path Analysis, Tobit Model, Ordered Probit (Fixed & Random Parameters) | Pre-crash driving volatility (speed and deceleration instability) significantly increases crash intensity. Distracted and aggressive driving correlates with volatility, which indirectly affects crash severity. Environmental factors like congestion and junction types also play a key role. |

## 2.6.Research Trends, Gaps and Future research orientations

### *2.6.1. Research Trends*

Over the past two decades, research in the road safety field has witnessed a major change in the method of data-driven study, with Naturalistic Driving Studies (NDS) becoming the leading method of reflecting real-world driver behavior and analyzing the safety of traffic with a higher level of accuracy. Compared to standard sources of crash data, i.e. police reports, aggregate data on historical crashes, etc., NDS allows real-time, high-resolution tracking of vehicles, drivers, and the environment. The shift is a methodological transformation of the way the outcomes of crashes are analyzed in the past to the proactive, behavior-oriented strategy, which can detect near-crashes, safety-critical events, risky maneuvers, and pre-crash signs in real time. By so doing, NDS directly overcomes the limitations of the traditional crash datasets, by capturing surrogate safety indicators, including sudden braking, lane deviations, glance behavior, and volatility, which would otherwise remain unnoticed in the official crash reporting (RQ1).

The latest trends show that quantitative methods are predominantly used nowadays, with most studies combining traditional econometric models with the novel machine learning and deep learning tools to forecast safety-critical events. The initial studies have used hierarchical modelling, logistic regression, and multilevel modelling, whereas recent studies are progressively using ensemble learning (e.g., XGBoost, Random Forest) and deep neural networks (e.g., LSTM, GRU). The evolution is seen as diversification of the typology of methods in which statistical interpretability, predictive performance, and real-time applicability are evaluated together (RQ2). Such methodology improvements have been used to improve the ability to model nonlinear relations and time dependencies in large-scale NDS data sets. Regarding the explanatory factors, the interaction between the driver behavior and the environmental factors is consistently presented in the literature as being the most critical factor of crash and near-crash risk. Distraction, length of glance, involvement in secondary task, emotional state, traffic density, road surface conditions, and configuration of roadways are the common contributory factors to safety-caused events. A combination of eye-tracking, video data, vehicle kinematics and real-time telemetry also allows a multidimensional evaluation of the combined impact of behavioral and contextual factors on missed or undetected crash risk (RQ3). Such a multifactorial view is an enormous improvement on the traditional crash-only analyses.

The United States has been the most active geographically in NDS research, which is in part because of the seminal studies, including the 100-Car Study and SHRP2, which have provided a basis on which many publications have been built. Nevertheless, even with increasing input by China, Sweden, Germany and Australia, the geographic focus of research in North America restricts the externalizable of results worldwide. Moreover, methodological sophistication has advanced, but comparatively low numbers of studies explicitly combine crash severity results or reveal the large-scale functioning of NDS-based forecast systems. These trends demonstrate that literature contains structural constraints such as geographic inequality, lack of severity-oriented analysis, and feasibility of scale (RQ4). In addition, NDS studies are now coming into close contact with new technologies in connected and autopilot vehicles, sophisticated driver assistance systems, and smart transportation systems. NDS is increasingly used in the field of designing real-time alert systems and risk prevention efforts by providing predictive analytics on the likelihood of a crash and driver behaviour under different circumstances. This growing status highlights a larger shift in the study of transportation safety: a shift between a reactive crash reporting system to a behavioral and system-based, proactive safety evaluation system.

### *2.6.2. Gaps in the current state of the art*

Despite the growing impact of Naturalistic Driving Studies (NDS) in advancing traffic safety research, several critical gaps limit the full translation of their findings into global and practical contexts. A predominant limitation lies in the geographical concentration of NDS research in North America, particularly within the United States. While these studies benefit from robust infrastructure and institutional support, their findings often reflect local traffic norms, infrastructure quality, and regulatory conditions. This presents a challenge when attempting to generalize results for low- and middle-income countries or regions with vastly different driving cultures and road systems. Expanding the geographic scope of NDS research is essential to ensure broader applicability and to address traffic safety challenges in underrepresented areas where road safety risks are often more severe.

Another key limitation is the underrepresentation of crash severity in many NDS-based studies. While significant efforts have been made to identify and analyze precursors to crashes and near crashes, few studies have directly linked these behavioral patterns with the actual severity of resulting incidents such as fatalities, injuries, or property damage. This limits the ability of current models to inform safety interventions that not only prevent crashes but also mitigate their consequences. Integrating severity-focused data, such as medical or emergency response outcomes, into NDS analysis would

enhance the predictive and practical value of these studies. Furthermore, high costs and logistical complexity associated with large-scale NDS deployments continue to hinder their scalability, particularly in resource-limited contexts. Most studies remain confined to specific vehicle types, road segments, or driver demographics, limiting the generalizability of insights across broader transportation networks.

In addition to these challenges, the integration of NDS findings into real-time traffic systems and emerging technologies remains an underexplored frontier. Although some progress has been made in using NDS data to support the development of Advanced Driver Assistance Systems (ADAS) and connected vehicle frameworks, few studies have demonstrated the seamless, closed-loop application of NDS-derived models in real-world operational environments. This gap prevents the full deployment of predictive safety systems that could dynamically respond to evolving road and driver conditions. To bridge this divide, interdisciplinary collaboration is urgently needed bringing together behavioral scientists, data engineers, software developers, and transportation planners to translate observational insights into deployable, real-time safety interventions. Such integration would not only enhance road safety outcomes but also support the broader goals of intelligent transportation systems and autonomous vehicle development.

### *2.6.3. Future research orientations*

Building upon the current state of Naturalistic Driving Studies (NDS), future research should prioritize expanding geographic diversity to ensure the global applicability of findings. While much of the existing knowledge originates from North America, driving behaviors, infrastructure conditions, and policy frameworks vary significantly worldwide. Future studies should focus on underrepresented regions particularly in low- and middle-income countries, to capture diverse driving environments and cultural contexts. This would not only improve the generalizability of results but also support evidence-based interventions tailored to local realities. Collaborative international projects and open-access data sharing initiatives could play a key role in achieving this inclusiveness. Another critical direction for future research lies in the integration of crash severity analysis into predictive modelling frameworks. Existing NDS-based studies tend to focus primarily on crash occurrence and frequency, often overlooking the severity of outcomes. Future work should therefore incorporate variables related to injury severity, economic impact, and post-crash consequences by linking Naturalistic Driving Data with complementary sources such as health records, insurance databases, and emergency response datasets. Integrating these dimensions enables the development of more comprehensive safety models

that capture not only the likelihood of crash occurrence but also the potential magnitude of their consequences.

Such an approach is essential for advancing integrated crash risk assessment, where risk is conceptualized as the combined interaction of crash frequency, crash severity, and exposure. By incorporating severity-related information into predictive frameworks, researchers can move beyond isolated probability models and construct composite risk indicators that reflect both the likelihood and the impact of crashes. This integrated perspective supports more effective safety prioritization, allowing policymakers and infrastructure planners to identify locations where crashes are not only frequent but also potentially severe. Moreover, integrated risk models can inform the development of intelligent transportation systems and proactive safety management strategies capable of both preventing crashes and mitigating their consequences when they occur, thereby contributing to a more comprehensive and system-based approach to road safety.

Equally important is the translation of NDS insights into real-time applications. To date, few studies have successfully embedded NDS-derived behavioral models into operational systems, such as Advanced Driver Assistance Systems (ADAS), traffic control centers, or connected vehicle environments. Future research should aim to close this gap by developing scalable, real-time analytics pipelines capable of ingesting continuous driving data and delivering actionable insights. This requires interdisciplinary collaboration among behavioral scientists, data scientists, software developers, and transportation engineers. Emphasis should also be placed on cost-effective technologies such as edge computing, low-cost sensors, and smartphone-based monitoring that can support real-time deployment, especially in resource-constrained settings. Ultimately, the next wave of research should not only advance theoretical understanding but also contribute to practical, deployable solutions that enhance road safety on a global scale.

### *2.6.4 Critical Synthesis and Research Motivation*

In summary, the reviewed literature demonstrates that Naturalistic Driving Studies (NDS) have significantly advanced road safety research by enabling the observation of real-world driving behaviour and the identification of Safety-Critical Events (SCEs) that are often overlooked in traditional crash databases. From a methodological perspective, statistical and econometric approaches have provided valuable insights through their interpretability, theoretical transparency, and ability to quantify the influence of explanatory variables using explicit parameter estimation and hypothesis testing. However, these approaches often struggle to capture the complex nonlinear interactions, behavioural dynamics, and high-dimensional characteristics inherent in telematics and trajectory-based datasets. Conversely, Artificial Neural Networks (ANNs), machine learning, and deep learning techniques have demonstrated superior predictive capabilities and greater flexibility in modelling complex behavioural and contextual relationships. Nevertheless, these methods are frequently criticized for their limited interpretability, higher computational requirements, sensitivity to data quality and model calibration, and potential risk of overfitting.

The literature further reveals that most existing studies focus on specific aspects of road safety, such as driver behaviour analysis, crash prediction, crash severity estimation, or localized hotspot identification, often treating these elements independently. Furthermore, relatively few studies integrate behavioural indicators, roadway infrastructure characteristics, environmental conditions, traffic exposure, crash frequency, and crash severity within a unified analytical framework. Limited attention has also been given to network-wide risk assessment, multi-scale analysis, and the application of explainable artificial intelligence techniques to improve transparency and practical decision-making. Consequently, there remains a need for an integrated framework capable of combining the interpretability of statistical models with the predictive power of artificial intelligence approaches while simultaneously addressing crash occurrence, crash severity, exposure, and behavioural precursors of risk. These research gaps provide the principal motivation for the framework developed in this dissertation.

# 3. CHAPTER 3

# METHODOLOGY

## 3.1.Methodological Framework Overview

This study proposes an integrated framework for assessing road safety risk at the network level by combining telematics data, crash records, roadway infrastructure characteristics, traffic exposure indicators, and environmental information within a unified analytical structure. The framework adopts a proactive and system-based perspective in which crash frequency and crash severity are modeled as complementary dimensions of road safety risk, while behavioral indicators derived from telematics observations are integrated with traditional crash-based evidence to capture potential risk precursors before crash occurrence. By jointly incorporating behavioral, infrastructural, environmental, and exposure-related factors, the proposed approach supports a more comprehensive and data-driven assessment of safety conditions across the road network and provides a basis for proactive road safety management.

Road safety risk is conceptualized through two interrelated dimensions: crash occurrence and crash consequences (Kashkanova, 2025). As per, Skaug et al. (2025) deicted that crash frequency represents the likelihood of crash events within the transport system, whereas crash severity reflects the magnitude of resulting outcomes. When considered alongside exposure conditions, these dimensions enable comprehensive risk estimation. Building upon this dual-dimensional conceptualization of risk, the framework adopts a system-oriented perspective in which safety outcomes emerge from interactions among human behavior, roadway infrastructure, traffic dynamics, environmental context, and exposure levels .These components generate operational risk indicators and safety-critical events that precede crashes and shape long-term safety patterns. Two theoretical foundations guide this approach: behavioral safety theory and risk-based road safety assessment.

### *3.1.1 Behavioral safety theory*

Behavioral safety theory posits that severe crashes are preceded by a large number of unsafe behaviors, near-miss events, and operational deviations (Heinrich, 1941; Bird, 1984). This relationship is represented by the Safety Pyramid (Fig. 3.1), illustrating the progression from routine unsafe acts to severe outcomes. Behavioral safety theory is based on the idea that serious crashs are

not isolated or random events, but rather the outcome of a chain of more frequent and less severe unsafe conditions (Heinrich 1941), and later adapted to road safety (Bird, 1984), observed that for every major injury crashs, there are many more minor injuries, property-damage incidents, near-misses, and unsafe acts occurring beforehand. In other words, severe crashes represent only the visible "tip" of a much larger base of unsafe behaviors and operational deviations. This relationship is commonly illustrated through the **Safety Pyramid** (Fig. 3.1). At the base of the pyramid are routine unsafe acts and behavioral deviations (e.g., speeding, harsh braking, distraction). Above this level are near-miss events and minor incidents, which occur less frequently but indicate increasing risk. At the top are severe crashes, which are rare but have the most serious consequences. The pyramid structure emphasizes that reducing frequent unsafe behaviors at the base can significantly decrease the probability of severe outcomes at the top. In the context of road safety, this theory supports the use of proactive indicators such as Safety-Critical Events (SCEs) because monitoring and mitigating unsafe behaviors before crashes occur can prevent the escalation from minor operational risk to severe crash consequences.

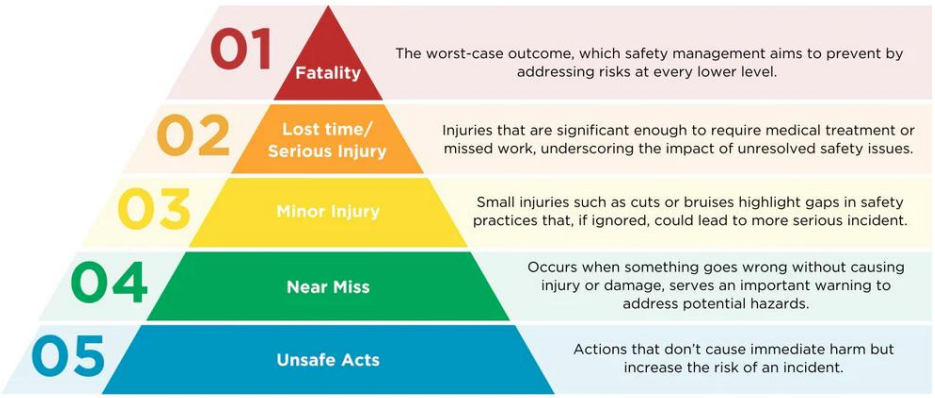



*Figure 0-1: Safety Pyramid (Heinrich (1941); Bird (1984) adapted to road safety)*

### *3.1.2 Risk-based road safety assessment*

In parallel, risk-based road safety frameworks define safety outcomes as the combined effect of crash likelihood, exposure, and consequences (Hauer Ezra, 1997; Elvik et al., 2009). In this structure, crash frequency represents probability, measuring how often crashes are expected to occur under given roadway and traffic conditions. Crash severity represents impact, reflecting the magnitude of harm once a crash occurs. Thus, two locations with similar crash counts may differ substantially in overall risk if the consequences are more severe in one case. Exposure represents the level of interaction

between road users and potentially hazardous conditions. Crashes can only occur when users are present and interacting within the system; therefore, higher traffic volumes or longer travel distances increase the opportunity for crashes.

By jointly considering probability, impact, and exposure, this framework enables a more balanced and meaningful assessment of road safety across the network. This perspective aligns with the European road safety target hierarchy (Fig. 3.2), which links high-level strategic objectives such as the reduction of fatalities and serious injuries to measurable intermediate outcomes and concrete operational interventions. Within this hierarchy, improvements in infrastructure design, speed management, and user behavior are treated as mechanisms through which final safety outcomes are achieved. The present PhD research contributes to this structure by developing quantitative models that measure and explain both crash occurrence and crash consequences, while incorporating proactive behavioral indicators derived from naturalistic driving data. In doing so, the research provides measurable intermediate risk indicators that can support strategic safety targets at the network level.

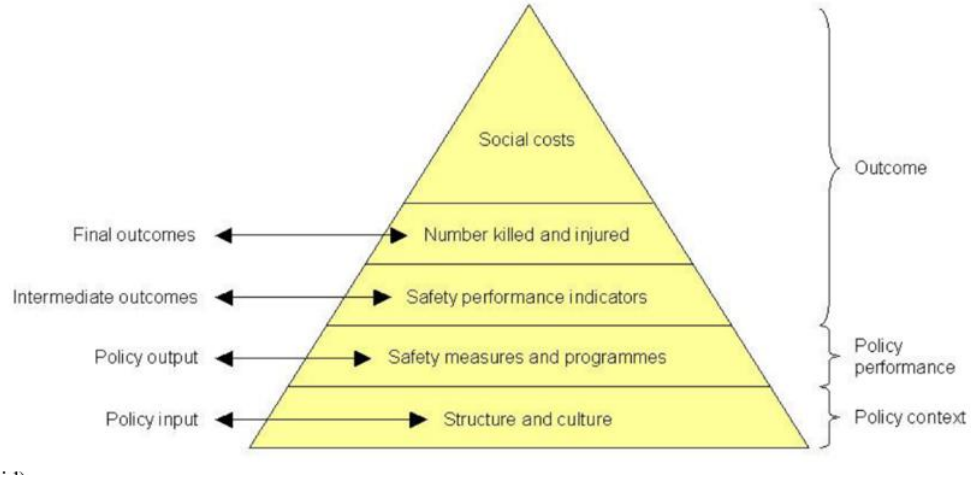



*Figure 0-2: European road safety target hierarchy (SUNflower pyramid)*

This conceptualization is further consistent with Directive (EU) 2019/1936 on road infrastructure safety management, which promotes proactive, risk-based, and network-wide safety assessment supported by evidence-based decision-making. Rather than relying solely on historical crash concentrations, the Directive encourages systematic screening of the entire road network using predictive and preventive approaches. Building on these principles, the PhD operationalizes the hierarchy through a quantitative modeling framework in which crash frequency models estimate crash likelihood, crash severity models quantify consequence magnitude, and exposure variables account for interaction intensity. By integrating telematics-based safety-critical events with traditional crash and infrastructure data, the research translates European policy principles into a measurable analytical

architecture that supports proactive network screening, risk mapping, and prioritized intervention planning. Building on these principles, the research operationalizes the hierarchy through a quantitative framework in which:

- crash frequency models estimate crash occurrence probability
- crash severity models quantify consequence magnitude
- exposure variables represent interaction intensity within the transport system

Road safety risk is therefore formulated as:

$$\textbf{Risk} = f(\textbf{Frequency} \times \textbf{Severity} \times \textbf{Exposure})$$

The proposed methodology operationalizes this formulation by integrating telematics-based behavioural indicators, roadway characteristics, exposure measures, and crash outcomes into a unified analytical structure. The conceptual analytical pathway is illustrated in Fig. 3.3. Within this framework, multi-source datasets including behavioural indicators from naturalistic driving data, crash records, geometric and infrastructural attributes, traffic conditions, environmental factors, and exposure metrics are harmonized into consistent modelling inputs. This integrated pipeline links data preparation, behavioural risk indicator extraction, and predictive modelling outputs in a traceable sequence. By jointly representing behavioural, infrastructural, traffic, environmental, and exposure dimensions within a single modelling architecture, the framework supports simultaneous estimation of crash occurrence mechanisms and severity outcomes. This unified structure enhances explanatory and predictive performance and enables comprehensive network-level risk estimation, spatial prioritization of high-risk locations, and evidence-based planning of safety interventions.

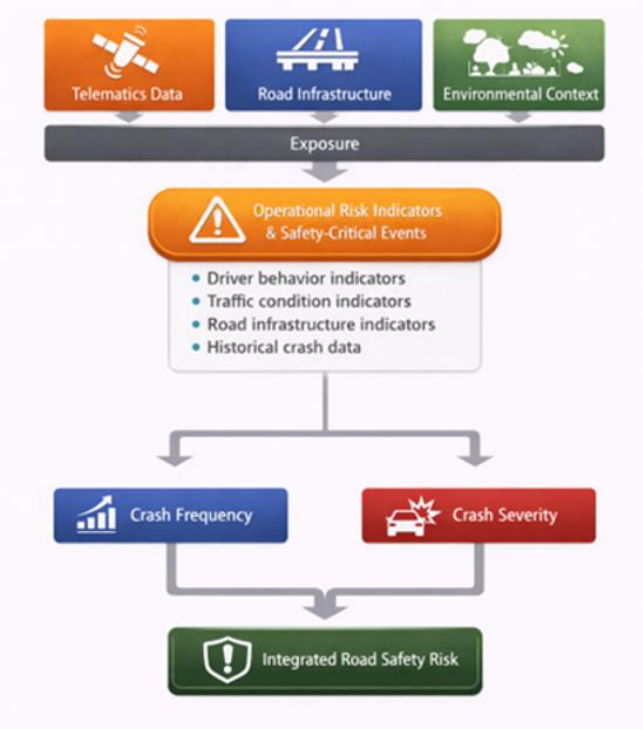



*Figure 0-3: Conceptual analytical pathway of methodology*

### 3.2.Structure of the Methodology Chapter

The subsequent sections present the components of the methodological framework in a structured sequence, beginning with data sources and preprocessing procedures, followed by the identification of operational risk indicators, crash frequency and severity modelling, model validation, composite risk formulation, and spatial analysis. The overall workflow is illustrated in Fig. 3.4 (Full Methodological Flowchart).

The remainder of this chapter is organized as follows:

• **Section 3.3:** Phase 1 – Telematics Data Engineering and Behavioral Event Construction (Sections 3.3.1–3.3.4)

• **Section 3.4:** Phase 2 – Parallel Modelling of Crash Frequency and Crash Severity (Sections 3.4.1–3.4.3)

• **Section 3.5:** Model validation and interpretation

• **Section 3.6:** Composite risk index formulation

• **Section 3.7:** Spatial risk mapping and decision-support application

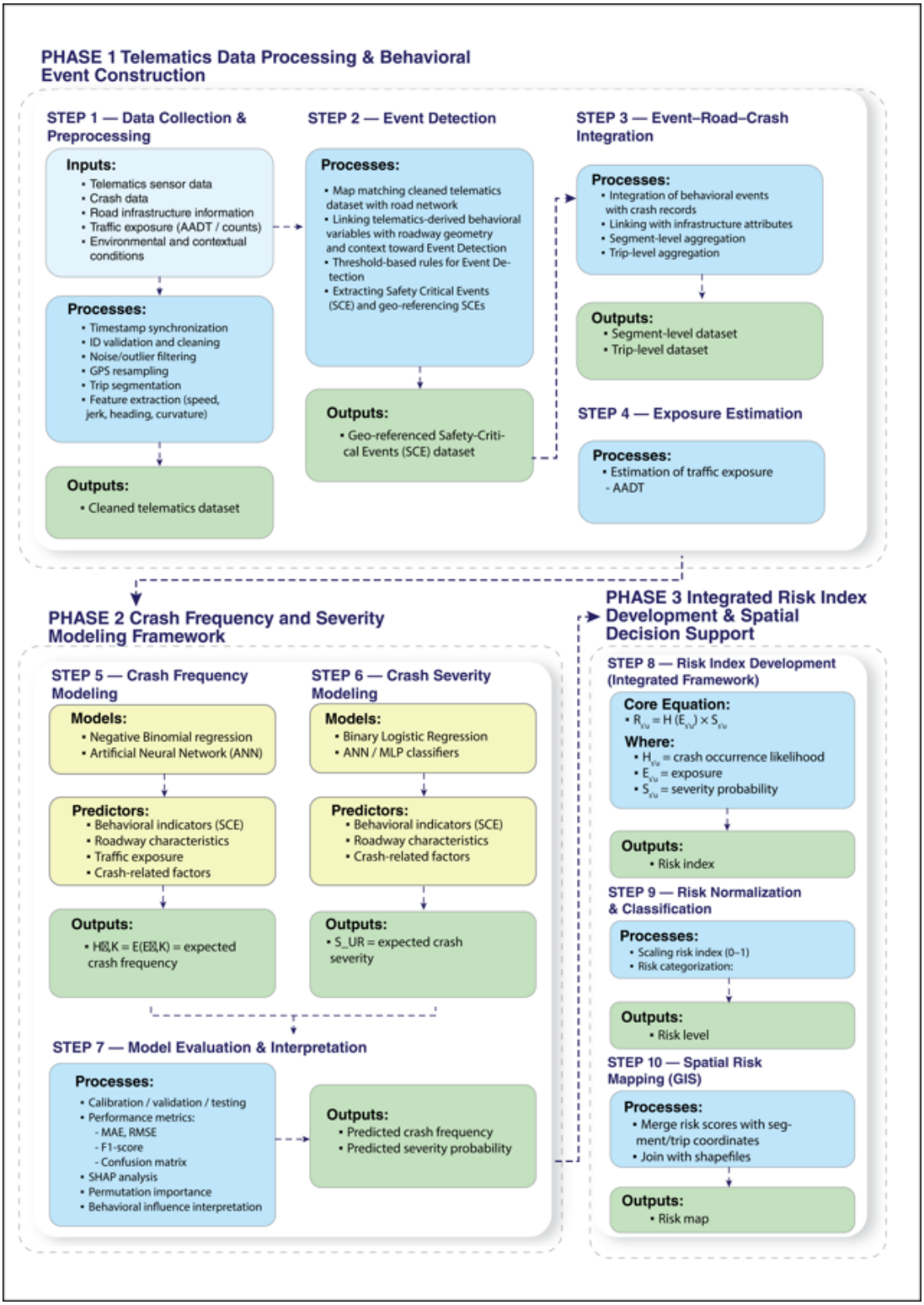



*Figure 0-4: Full Methodological Flowchart*

## 3.3.Phase 1: Telematics Data Engineering and Behavioral Event Construction

Before implementing the methodological framework, it is essential to delineate the study area and identify the data sources required to support a comprehensive risk assessment process. The presentation of data sources follows established practices in road safety research, ensuring transparency regarding spatial coverage, temporal scope, and analytical suitability, while the methodological design remains tailored to the objectives of this study. Clearly defining data provenance and structure is particularly important in multi-source safety analyses, where inconsistencies in scale or quality may affect model validity (Pei et al., 2012). The study focuses on a road network environment in which telematics observations, roadway characteristics, crash records, and exposure indicators are integrated to support predictive safety modelling and risk estimation. The growing availability of telematics and naturalistic driving data has transformed road safety analysis by enabling direct observation of driver behaviour in real-world conditions, including speed variability, harsh acceleration or braking, and trajectory patterns (Shichrur & Ratzon, 2023). These behavioural indicators provide valuable insight into near-crash situations and unsafe manoeuvres that often precede actual crashes but are not captured in conventional databases.

Given the objective of developing a data-driven, behaviour-based safety framework, data sources are selected based on availability, spatial representativeness, temporal consistency, and relevance for modelling both crash occurrence and severity. The data collection process incorporates heterogeneous sources grouped into four main categories: telematics trajectory data, roadway and environmental attributes, traffic exposure information, and historical crash data. Integrating these complementary datasets enables the construction of a multi-dimensional safety database capable of supporting advanced predictive modelling and proactive identification of hazardous conditions across the network (Moradi, 2022).

### *3.3.1. Step 1 - Data Collection and Pre-processing*

The first step involves the acquisition and preparation of multi-source datasets required for safety analysis. The primary data source consists of telematics-based vehicle trajectory data, including timestamped GPS coordinates and derived motion information. These data are complemented by crash databases, roadway inventory information, and traffic exposure indicators such as Average Annual Daily Traffic (AADT), which are used in subsequent modelling stages.

**Telematics sensor data**:

Telematics data consist of in-vehicle measurements capturing real-world driving activity through GPS trajectories and accelerometer recordings. These data enable reconstruction of vehicle movement and extraction of motion-based indicators such as speed variation, jerk, heading change, and curvature (Winlaw et al., 2019; Paefgen et al., 2013). Such indicators support the identification of operational driving patterns and safety-critical events used in later modelling stages. To operationalize these behavioral indicators, telematics signals are transformed into quantitative kinematic variables derived from trajectory and acceleration data.

To ensure analytical reliability and internal consistency, the raw telematics trajectories were subjected to a systematic preprocessing workflow prior to the derivation of motion-based indicators. The preprocessing stage aimed to correct temporal inconsistencies, remove erroneous records, and ensure the spatial and kinematic validity of the observations. Cleaning procedures included timestamp validation and ordering, removal of duplicated observations, filtering of unrealistic motion values (e.g., implausible speed or acceleration values), elimination of incomplete geographic coordinates, and verification of vehicle identifiers. These steps were necessary to mitigate measurement noise and inconsistencies commonly present in large-scale telematics datasets.In addition, the telematics trajectories examined to evaluate the temporal sampling behavior and the statistical characteristics of the variables. Mostly the sampling interval between consecutive observations ($\Delta t$) shows considerable variability across the dataset.

Again, to ensure these kinematic variables are analytically reliable and comparable across the dataset, the raw telematics trajectories undergo a structured preprocessing workflow before behavioral indicators are computed. This preprocessing stage resolves temporal irregularities, removes erroneous or incomplete records, and standardizes trajectories into coherent driving episodes suitable for subsequent modelling. The resulting cleaned data then serve as the basis for the trip segmentation, temporal resampling, and sliding-window decomposition procedures described next.

*3.3.1.1 Trip Segmentation, Resampling, and Windowing Framework*

To enable consistent behavioural analysis from Naturalistic Driving Data (NDD), a structured preprocessing workflow is applied, consisting of trip segmentation, temporal resampling, and sliding-window decomposition (Dingus et al., 2006; Klauer et al., 2006). This workflow ensures temporal

coherence, comparability across trajectories, and reliable detection of short-duration behavioural patterns.

***Trip Segmentation***

Raw telematics data typically consist of sequential vehicle measurements recorded at irregular time intervals and may include interruptions caused by parking, signal loss, congestion, or temporary device inactivity. To preserve behavioral continuity and ensure analytical consistency, a trip segmentation procedure is implemented prior to further processing. Observations separated by a predefined temporal gap are treated as belonging to different driving episodes, and each continuous movement sequence is assigned a unique trip identifier. This procedure prevents extended inactivity periods from being incorrectly interpreted as part of the same trajectory and establishes a coherent analytical unit for subsequent resampling, windowing, and behavioral analysis.

The temporal threshold is defined to distinguish short operational pauses (e.g., traffic signals, congestion delays, brief stops) from true trip termination events (e.g., parking or destination arrival). In urban and peri-urban driving contexts, temporary stops are typically short-lived, whereas destination-related inactivity tends to persist for longer durations. Exploratory analysis of the inter-observation time-gap distribution can further support the identification of a natural separation between brief interruptions and prolonged inactivity periods. The adoption of a time-gap–based segmentation rule is consistent with established practices in GPS trajectory and telematics research, where trip identification commonly relies on temporal continuity criteria (Axhausen et al., 2002; Stopher et al., 2007; Gong et al., 2012; Zheng, 2011).

Formally, let $t_i$and $t_{i+1}$denote two consecutive timestamps associated with vehicle $v$. A new trip is defined whenever the temporal gap exceeds a predefined threshold $\tau$:

$$t_{i+1} - t_i > \tau \tag{3.1}$$

This condition operationalizes the temporal continuity rule by ensuring that only temporally connected observations are grouped within the same trip, thereby maintaining behavioral integrity within each driving episode and supporting robust downstream safety analysis.

***Temporal Resampling***

Naturalistic driving datasets often contain non-uniform sampling intervals, which can bias the estimation of kinematic derivatives and reduce comparability across trips. To ensure temporal consistency, trajectories are resampled to a constant frequency. Let raw observations be denoted as $P(v, r, t)$, where $v$ represents the vehicle, $r$ the trip, and $t$ the timestamp. After interpolation to a fixed frequency $f$(e.g., $f = 2$Hz), the time interval becomes

$$\Delta t = \frac{1}{f} = 0.5 \text{ seconds} \tag{3.2}$$

and the standardized trajectory is expressed as:

$$P^*(v, r, t_k), t_k = t_0 + k\Delta t \tag{3.3}$$

This temporal standardization produces uniformly spaced observations from which speed, acceleration, jerk, heading change, and curvature can be consistently computed. By enforcing equal time intervals, the procedure improves numerical stability, reduces sensitivity to irregular sampling noise, and ensures comparability of behavioral indicators across trips.

***Sliding Window Decomposition***

To capture short-term driving dynamics and avoid reliance on single-point observations, trajectories are decomposed into fixed-duration temporal windows using a sliding-window approach. This representation allows behavior to be evaluated over consistent time horizons, improving robustness against sensor noise and enhancing the detection of abrupt maneuvers. Following resampling, each trip trajectory $P^*(v, r, t_k)$is divided into windows of length $\omega$:

$$w_j = \{P^*(v, r, t_k) \mid t_k \in [t_j, t_j + \omega]\}, t_j = t_0 + j\delta \tag{3.4}$$

where $\delta$ denotes the step size (windows overlap if $\delta < \omega$). With sampling frequency $f$, the number of observations per window is:

$$n_\omega = \omega \cdot f \tag{3.5}$$

Behavioral indicators are computed at the window level, for example:

$$A_y(w_j) = \max_{t_k \in w_j} \mid a_y(t_k) \mid \tag{3.6}$$

The window-based framework evaluates behavior over short, consistent time horizons by aggregating consecutive observations within each temporal segment. This approach preserves short-term driving dynamics while reducing sensitivity to isolated sensor noise and single-point anomalies, ensuring that detected patterns represent sustained and meaningful behavioral deviations. The resulting representation provides a stable basis for identifying Safety-Critical Events using rule-based thresholds, statistical methods, or machine-learning techniques. The sliding-window decomposition is applied independently within each segmented trip, and windows are strictly confined to trip boundaries. By preventing windows from spanning temporally disconnected driving episodes, the procedure preserves behavioral integrity and avoids artificial continuity across separate trips.

Overall, the integrated segmentation–resampling–windowing procedure establishes a robust analytical framework in which safety-relevant behaviors such as sudden braking, rapid acceleration, vehicle instability, and excessive speed are identified systematically without relying on isolated observations or overly aggregated trip-level summaries. By preserving behavioral continuity within trips and ensuring temporal consistency across trajectories, the workflow enhances the reliability and interpretability of dynamic driving patterns. Following the establishment of a temporally consistent trajectory dataset, kinematic indicators describing vehicle motion are computed from consecutive observations. *which is calculated as follows:*

**Speed estimation**

Vehicle speed is computed from the spatial displacement between consecutive GPS observations and the corresponding time difference:

$$v_i = \frac{d_i}{\Delta t_i} \tag{3.7}$$

where:

- $v_i$= speed at observation $i$(m/s),
- $d_i$= distance between consecutive GPS points,
- $\Delta t_i = t_i - t_{i-1}$= time difference between observations.

**Jerk calculation**

Jerk, describing abrupt variations in driving behavior, is calculated as the temporal derivative of acceleration measured by accelerometers:

$$jerk_i = \frac{a_i - a_{i-1}}{\Delta t_i} \tag{3.8}$$

Where:

- jerk at observation *i* (m/s³),
- $a_i$= acceleration at observation *i* (m/s²),
- $a_{i-1}$= acceleration at observation *(i−1)* (m/s²),
- $\Delta t_i = t_i - t_{i-1}$= time difference between consecutive observations (s).

This indicator captures sudden changes in driver control and is commonly associated with safety-critical maneuvers (Bagdadi & Várhelyi, 2011; Eren, H. et al., 2012).

**Heading change estimation**

Vehicle directional change is derived from the angular difference between consecutive trajectory segments available from the raw dataset :

$$\Delta\theta_i = \theta_i - \theta_{i-1} \tag{3.9}$$

where $\theta_i$represents the movement bearing at observation $i$.

Heading change describes turning behavior and directional instability.

**Curvature estimation (three-point rule)**

Road trajectory curvature is estimated using three consecutive GPS points $P_{i-1}, P_i, P_{i+1}$. The radius of the circumcircle passing through these points is calculated as:

$$R_i = \frac{abc}{4A} \tag{3.10}$$

where:

- $a, b, c$ are the distances between the three points,
- $A$is the area of the triangle formed by the points.

Curvature is then derived as:

$$\kappa_i = \frac{1}{R_i} \tag{3.11}$$

Higher curvature values indicate sharper vehicle turning trajectories (Lipar et al., 2011).

The outcome of this process is a clean, structured telematics dataset enriched with derived kinematic and geometric indicators, including speed variation, longitudinal and lateral acceleration, jerk, heading change, and curvature. These variables form the core behavioral inputs for subsequent Safety-Critical Event detection and serve as fundamental explanatory components in the development of crash frequency and crash severity models within the broader methodological framework.

*3.3.1.2 Additional Data Sources*

**Crash data:**

In addition to the telematics dataset, official crash records were incorporated to provide the empirical outcome layer of the analysis. Historical crash datasets contain structured information on crash occurrence, spatial location, injury severity, involved road users, environmental conditions, and temporal attributes. Data availability varies across countries depending on reporting systems. For example, in the Italian context, official crash records are obtained from the Istituto Nazionale di Statistica (ISTAT) Road Traffic Crash Database, which represents the national crash reporting system and provides standardized, nationwide crash information.

The ISTAT database includes detailed attributes on crash location (municipality, road reference, and geographic coordinates where available), time of occurrence (year, month, day, hour), injury severity classification (fatal, Sever injury), vehicle and road user types involved, as well as environmental and lighting conditions. Within the proposed methodological framework, crash data serves two principal functions. First, they provide the dependent variables required for estimating crash frequency and crash severity models in Phase 2. Second, they act as ground-truth observations against which telematics-derived behavioral indicators and Safety-Critical Events are evaluated.

To ensure analytical coherence, crash records were filtered according to the study period and spatial extent and subsequently linked to the predefined road network segmentation through a segment-based spatial aggregation procedure. This spatial harmonization enables the direct integration of crash outcomes with telematics-derived behavioral measures, roadway infrastructure characteristics, and exposure indicators within a unified segment-level dataset. Through this structured integration, crash data operationalizes the consequence dimension of the safety framework and supports the calibration and validation of behavior-based predictive models.

**Roadway Inventory and Infrastructure Data:**

Road inventory and contextual characteristics were compiled from multiple geospatial sources to ensure a comprehensive representation of infrastructure-related and environmental risk factors within the analytical framework. These sources include official regional Geographic Information System (GIS) datasets, OpenStreetMap (OSM), and structured spatial queries extracted through Overpass-based mapping tools. The integration of these datasets enabled the construction of a harmonized digital road inventory containing geometric, functional, and contextual attributes relevant to safety analysis. To ensure data reliability and attribute consistency, a manual verification process was also conducted, during which key road characteristics were cross-checked and corrected where necessary, acknowledging that digital geospatial datasets may contain inconsistencies, omissions, or classification errors.

The compiled inventory includes variables such as road functional classification (e.g., urban, rural, arterial, local), lane configuration and carriageway structure, horizontal curvature and geometric properties, intersection density, posted speed limits, built-environment characteristics, lighting conditions, and surrounding land-use context. These attributes describe the physical and operational characteristics of the road environment and constitute the structural component of the safety system. Within the methodological framework, roadway inventory variables are incorporated as explanatory factors in the modelling stages, enabling infrastructure characteristics to be explicitly represented alongside behavioral and exposure indicators in the analytical dataset. In addition, selected infrastructure attributes are used during the event-extraction stage. For example, segment-level speed limit information is integrated with telematics-derived speed profiles to identify speeding-related Safety-Critical Events, ensuring that behavioral thresholds are evaluated relative to the regulatory context of each road segment.

**Environmental and Contextual Conditions:**

Environmental and contextual factors are incorporated to account for external conditions that influence driving behavior and crash probability. These variables capture the spatial and temporal variability of the operational environment and complement the behavioral, infrastructure, and exposure-related components of the safety framework. The environmental dataset includes attributes such as weather conditions, lighting (e.g., daylight, nighttime, illuminated or non-illuminated segments), land-use classification, and surrounding built-environment characteristics. Weather and

lighting information are integrated using temporally referenced crash records and contextual datasets, while land-use and built-environment indicators are derived from geospatial inventory sources and spatial overlays with the road network.

Within the modelling framework, these contextual variables are treated as explanatory factors that may simultaneously influence behavioral responses and crash outcomes. For example, drivers may adapt speed or maneuvering patterns under adverse weather conditions, while reduced visibility during nighttime or in non-illuminated segments may increase crash severity. By explicitly incorporating environmental and contextual conditions, the methodology captures exogenous influences on safety performance and strengthens the system-based interpretation of crash risk across different spatial and temporal contexts.

### *3.3.2. Step 2 – Extracting and Geo-referencing Safety Critical Events (SCEs)*

*3.3.2.1 Integration of Road Infrastructure Context*

Since Step 1 already established the harmonized roadway inventory and contextual database, Step 2 focuses specifically on the **map-matching procedure**, which operationally links each telematics observation to the correct infrastructure segment within that inventory. In the methodological framework, map matching is performed after data cleaning and temporal standardization, and its purpose is to spatially anchor every GPS observation to the digital road network so that behavioral indicators can be interpreted relative to infrastructure and environmental attributes.

Methodologically, the map-matching process begins with a clean set of geo-referenced telematics observations, where the spatial position of vehicle $v$ during trip $r$ at time $t$ is formally denoted as $x(v, r, t) = (lat, lon)$, representing its geographic coordinates within the adopted reference system. For each observation, a candidate set of nearby road segments is identified within a predefined search radius using spatial indexing techniques (e.g., R-tree structures). Candidate selection is based on multiple criteria, including proximity to the segment geometry, directional consistency between vehicle heading and segment bearing, and network topology constraints. A composite scoring function is then applied to evaluate the likelihood that a given segment corresponds to the true vehicle position. This score typically integrates (i) the perpendicular distance from the GPS point to the segment, (ii) the angular difference between vehicle trajectory and segment orientation, and (iii) path continuity constraints to preserve temporal and spatial consistency across consecutive observations within the

same trip. The segment with the highest composite likelihood is ultimately selected as the matched segment, thereby transforming raw geographic coordinates into network-referenced observations suitable for infrastructure-integrated safety analysis.

To preserve behavioral continuity, map matching is performed trip-by-trip, meaning that resampling and windowing (already conducted in Step 1) are maintained within each trip before spatial assignment. This ensures that consecutive observations within a trip are matched coherently along plausible network paths rather than independently assigned, reducing spatial jitter and unrealistic segment switching. Network topology is used to verify that consecutive matched segments are connected within the digital graph, thereby enforcing route feasibility.

Once map matching is completed, each telematics record inherits the full set of roadway and contextual attributes from the matched segment. This includes functional classification, geometric configuration, speed limits, lane structure, curvature, intersection presence, lighting conditions, land use, and other environmental characteristics previously compiled. This spatial integration enables behavioral indicators such as speed variation, acceleration, jerk, heading change, and curvature to be evaluated relative to regulatory and infrastructural context (e.g., detecting speeding relative to segment-level posted speed limits).

From a methodological standpoint, map matching therefore serves as the critical integration bridge between behavioral signals and structural roadway conditions. It transforms raw trajectory data into infrastructure-aware behavioral observations, enabling the subsequent extraction of Safety-Critical Events (SCEs) and supporting the development of crash frequency and severity models within a spatially consistent and network-referenced analytical framework.

*3.3.2.2 Safety-Critical Event (SCE) Definition and Detection*

Following the completion of map matching and infrastructure contextualization, the next methodological step consists of defining and extracting **Safety-Critical Events (SCEs)** from the network-referenced telematics dataset. Safety-Critical Events are defined as observable driving situations characterized by abrupt or extreme maneuvers that indicate elevated crash risk or operational instability, consistent with behavioral safety theory (Klauer et al., 2006; Bagdadi, 2013). Typical examples include sudden braking, rapid acceleration, sharp turning, speeding, and lane deviation. These events represent behavioral precursors to crashes and capture near-miss situations

that are typically absent from official crash databases. As such, SCEs function as proactive risk proxies, allowing the identification of latent safety conditions that precede documented crash outcomes.

After spatial integration, the cleaned and map-matched telematics dataset is transformed into a structured, geo-referenced SCE database. This transformation converts continuous trajectory and accelerometer measurements into discrete, behaviorally meaningful events through a systematic workflow that includes indicator construction, rule-based detection, temporal window evaluation, multi-layer validation, and final geo-referencing. Event detection is implemented using calibrated threshold criteria derived from scientific literature, empirical observation of the dataset, and engineering judgement. Each SCE is defined as a sustained deviation from normal driving behavior exceeding predefined magnitude and duration limits. For example, sudden braking is identified through excessive longitudinal deceleration, sharp turning through elevated lateral acceleration or curvature, and speeding through sustained exceedance of the segment-specific posted speed limit inherited during map matching. Evaluating signals over temporal windows reduces sensitivity to measuring noise and ensures behavioral persistence rather than isolated point anomalies.

To enhance robustness and minimize false detections, the threshold-based framework incorporates a multi-layer validation structure. This includes signal integrity screening to remove unrealistic or corrupted sensor values, temporal persistence requirements to confirm sustained maneuvers, contextual plausibility checks based on matched roadway characteristics (e.g., curvature consistency with sharp turning), and cross-indicator confirmation using multiple telematics signals (e.g., acceleration and jerk alignment). This layered validation approach balances sensitivity and specificity while maintaining methodological transparency. The resulting SCE database constitutes the behavioral core of the analytical framework and serves as the primary input for subsequent aggregation, crash frequency modelling, and severity analysis stages.

#### *3.3.2.3 Formal Definition of the Mined Observational Records and Event-Based Analytical Dataset*

From the integrated telematics and roadway inventory data, a structured dataset of network-referenced observations is constructed. Let $T$ denote the monitoring period and $t \in T$ a generic timestamp. Let $V$ be the set of observed vehicles during $T$, with $v \in V$ representing a specific vehicle. For each vehicle $v$, let $R(v)$ denote the set of trips performed, and $r \in R(v)$ a generic trip. A telematics observation is defined as $P(v, r, t)$, representing the recorded data for vehicle $v$, trip $r$, at time $t$, including at

minimum its spatial position and motion signals. The spatial position is denoted as $x(v,r,t) = (lat, lon)$, while the tri-axial acceleration vector is expressed as $a(v,r,t) = (a_x, a_y, a_z)$, and the vehicle heading (bearing) as $\theta(v,r,t)$.

Let $S$ denote the set of road segments in the digital road network, with $s \in S$representing a generic segment. Each segment is associated with a vector of roadway attributes $A(s)$, including functional classification, geometric configuration, lane structure, speed regime, surface condition, lighting, directionality, and pedestrian facilities. A map-matching function $\Phi(\cdot)$assigns each spatial observation to a road segment such that $s(v,r,t) = \Phi(x(v,r,t))$, where $s(v,r,t) \in S$is the segment associated with the observation. This step transforms raw geographic coordinates into network-referenced observations enriched with infrastructure attributes.

To support behavioral inference and event construction, observations within each trip are organized into temporally standardized windows. Let $W(v,r)$denote the ordered set of windows associated with trip $r$ of vehicle $v$, and $w \in W(v,r)$a generic window representing a short time interval containing a sequence of resampled observations $\{P(v,r,t)\}$. Within each window, a vector of behavioural indicators $F(w)$is computed, including speed metrics, longitudinal and lateral acceleration extrema, jerk summaries, heading-change measures, and curvature indicators derived from the resampled signals. A representative road segment $\bar{s}(w)$is then assigned to each window by applying the map-matching function to all observations within the window and selecting the most consistent segment (e.g., via majority assignment or highest confidence). The contextual attribute vector associated with the window is defined as $C(\bar{s}(w)) = A(\bar{s}(w))$, optionally augmented with exposure or environmental descriptors when available.

Consequently, the detection of Safety-Critical Events (SCEs) in this research is based on a threshold-based behavioral identification framework designed to ensure both robustness and interpretability. The approach draws on empirical safety thresholds reported in established naturalistic driving studies, including the 100-Car Study and relevant AASHTO guidelines, which provide reference values for abnormal driving dynamics such as sudden deceleration, excessive acceleration, sharp turning, or instability patterns. Rather than relying on a single indicator, the methodology applies a multi-layered threshold logic, where several behavioral and contextual conditions must be satisfied simultaneously within a defined temporal window. In practice, events are only flagged when multiple threshold criteria are exceeded within the same observation window, such as combinations of acceleration

magnitude, jerk intensity, speed variation, or heading change. This layered detection structure substantially improves event reliability by filtering out noise, sensor artefacts, or isolated fluctuations that could otherwise generate false positives. Consequently, the detected events represent meaningful safety-relevant behavioral episodes rather than random measurement anomalies.

Formally, Safety-Critical Events are identified through a generalized rule-based detection function $\Psi(\cdot)$, which maps behavioural indicators and contextual attributes into a specific event classification:

$$\Psi(F(w), C(\bar{s}(w))) \rightarrow e_{\text{type}} \text{ or } \emptyset \tag{3.12}$$

where $e_{\text{type}}$ denotes a specific SCE class (e.g., braking, acceleration, turning, speeding, lane instability), and $\emptyset$ indicates that no event is detected when the combined thresholds are not satisfied. This formulation explicitly reflects the multi-layered decision mechanism, where the behavioral feature vector $F(w)$and contextual attributes $C(\bar{s}(w))$jointly determine whether the predefined threshold structure is met.

***The pseudo-code of the multi-layer Data Screening algorithm (Sharp Turn used as illustrative SCE example) is reported here after:***

*M1(u) ← ∅*
*For each window w ∈ M0(u) do // first screening (core dynamics)*
*if |acc_y(w)| ≥ θacc and curvature(w) ≥ θcurve then*
*M1(u) ← M1(u) ∪ {w}*
*end if*
*end for*

*M2(u) ← ∅*
*For each window w ∈ M1(u) do // second screening (maneuver consistency)*
*if heading_change(w) ≥ θhead then*
*M2(u) ← M2(u) ∪ {w}*
*end if*
*end for*

*M3(u) ← ∅*
*For each window w ∈ M2(u) do // third screening (feasibility constraint)*
*if speed(w) ≥ θspd then*

*M3(u) ← M3(u) ∪ {w}*
*end if*
*end for*

*// event consolidation*
*Group consecutive windows in M3(u) whose temporal gap ≤ Δt*
*For each group g do*
*select w* = argmax(|acc_y(w)|) within g // peak maneuver intensity*
*Ψ(F(w*), C(s̄(w*))) → e_type = Sharp Turn*
*E ← E ∪ {w*}*
*end for*

*Return E*
***Notes:***
***Ψ(F(w), C(s̄(w))) → e_type is operationalized here as layered threshold screening:***
***M0 → M1 → M2 → M3, where only windows satisfying all thresholds are retained as events.***

Ψ(F(w), C(s̄(w))) → e_type is operationalized through a **multi-layer threshold screening procedure.** The candidate window set is progressively refined through sequential filtering stages ($M_0 \rightarrow M_1 \rightarrow M_2 \rightarrow M_3$, ) where only windows satisfying all behavioral and feasibility constraints are retained. Finally, temporally adjacent candidate windows are grouped and consolidated to avoid multiple detections of the same maneuver, and the peak-intensity window within each group is labeled as a **Sharp Turn event as an example.**

Based on these definitions, the final analytical output of this phase is the geo-referenced Safety-Critical Event database, defined as:

$$E=\{e|e=(v,r,w,s^{-}(w),e_{type},F(w),C(s^{-}(w)),t_{ref}(w)),\ \Psi(F(w),C(s^{-}(w)))\neq\emptyset\} \quad (3.13)$$

Each element $e \in E$represents a validated behavioral episode associated with a specific vehicle, trip, temporal window, and road segment, while also incorporating behavioral descriptors and infrastructure context. The resulting dataset therefore forms the core behavioral layer of the methodological framework, capturing high-resolution indicators of driving instability. These threshold-derived behavioral signals subsequently serve as the primary inputs for the crash frequency models, crash severity models, and the integrated network risk assessment developed in the following phases of the research.

In summary, Step 2 converts validated driving maneuvers into geo-referenced Safety-Critical Events by linking each episode to its representative road segment and contextual attributes. The resulting dataset integrates behavioral, infrastructure, exposure, and environmental information within a consistent spatial framework. This structured SCE database forms the behavioral foundation for subsequent crash frequency, severity, and integrated risk modelling stages.

#### *3.3.2.4 Justification of Threshold Values for Safety-Critical Event Detection*

The detection of Safety-Critical Events (SCEs) in this research was based on threshold values selected through a structured calibration process rather than arbitrary judgement. The purpose of these thresholds was to identify driving manoeuvres and vehicle responses that plausibly reflect elevated crash risk while minimizing false detections caused by normal operational variability or sensor noise. Threshold selection was guided by three complementary principles. First, values reported in previous Naturalistic Driving Studies and surrogate safety literature were reviewed to identify commonly accepted ranges for abnormal driving behaviour, including sudden braking, harsh acceleration, excessive lateral manoeuvres, speeding, and vertical instability events. Second, engineering relevance was considered so that the adopted thresholds reflected behaviourally meaningful deviations from normal driving conditions, such as manoeuvres associated with aggressive behaviour, increased workload, or potential conflict situations. Third, empirical inspection of the Brescia telematics dataset was undertaken to ensure that the resulting event frequencies were realistic and operationally interpretable.

Accordingly, sudden braking events were identified using abnormal longitudinal deceleration combined with rapid jerk response, rapid acceleration events captured unusually aggressive acceleration behaviour, speeding events were defined relative to segment-specific posted speed limits, and sharp-turn events required combined evidence from lateral dynamics, heading variation, and roadway curvature. Additional temporal persistence and contextual plausibility rules were introduced to reduce isolated noise-induced detections. Where appropriate, nearby alternative threshold values were also examined through sensitivity checks to confirm that the principal behavioural patterns remained stable under minor parameter variation. This calibration strategy improves the interpretability, robustness, and practical usefulness of the final SCE indicators adopted in the subsequent modelling stages.

### *3.3.3. Step 3 Event–Road–Crash Integration*

Following the construction of the geo-referenced Safety-Critical Event dataset, this stage integrates behavioural events with roadway infrastructure attributes and historical crash information to produce a unified analytical dataset for safety modelling. This step marks the transition from behavioural signal extraction to explanatory and predictive analysis, where telematics-derived behaviour, infrastructure conditions, crash occurrence, and crash severity are examined within a consistent spatial framework. Using the formal definitions introduced in the previous phase, the behavioral layer is represented by the geo-referenced Safety-Critical Event database as defined by equation 3.13.

Each element corresponds to a validated behavioural episode associated with a vehicle, trip, temporal window, and road segment, including behavioural descriptors, contextual infrastructure attributes, and temporal reference.

*3.3.3.1 Integration of behavioral, infrastructure, and crash dimensions*

Following geo-referencing, behavioral events are systematically integrated with roadway characteristics and historical crash records through their shared spatial association with road segments. Because each Safety-Critical Event is already linked to a specific segment, infrastructure integration is performed by joining segment-level attributes such as geometric configuration, lane structure, curvature, traffic control, speed limits, lighting conditions, and intersection characteristics directly to the behavioral dataset. This enables driving patterns to be interpreted within the physical and regulatory environment in which they occur.

Historical crash data are incorporated by spatially allocating crash records to the same segment structure, introducing two complementary safety dimensions: **crash frequency**, representing the spatial distribution and occurrence of crashes across the network, and **crash severity**, reflecting the consequences of crashes in terms of injury outcomes, damage levels, or fatality risk. By combining behavioral precursors with both crash occurrence and crash consequence data, the framework captures the full safety chain from operational instability to realized crash outcomes. Within this structure, Safety-Critical Events are treated as behavioral drivers that influence both crash probability and outcome severity, allowing simultaneous analysis of crash likelihood, expected severity, and the interaction between behavior, infrastructure, and environmental conditions. This integrated perspective positions behavior as the central explanatory bridge between dynamic driving processes and measurable safety performance.

*3.3.3.2 Multi-Scale Aggregation*

To support predictive modelling and system-level safety analysis, the event-level dataset is transformed into modelling-ready analytical units through a structured multi-scale aggregation process. The two aggregation scales segment-level and trip-level represent distinct but complementary analytical perspectives within the safety system. Both scales incorporate behavioral indicators derived from Safety-Critical Events (SCEs), while preserving the spatial and temporal linkages established during earlier methodological stages.

Because map-matching and infrastructure linkage were completed during the SCE extraction phase, each behavioral event was already associated with a specific roadway segment and its infrastructural attributes prior to aggregation. Historical crash records were subsequently spatially allocated to the same segment structure. This ensured that behavioral precursors, infrastructure characteristics, crash occurrence, and crash severity outcomes were aligned within a consistent spatial framework before constructing higher-level analytical datasets.

*3.3.3.3 Formal Definition of the Aggregated Analytical Datasets*

Building on the geo-referenced event-level dataset:

$$E = \{e \mid e = (v, r, w, \bar{s}(w), e_{type}, F(w), C(\bar{s}(w)), t_{ref}(w))\} \tag{3.15}$$

Two modelling-ready datasets are constructed.

1. *Segment-Level Dataset (Macro-Level Perspective)*

Let $S$ denote the set of road segments and $s \in S$ a generic segment.

The segment-level analytical dataset is defined as:

$$D_{seg} = \{d_s \mid d_s = (s,\ B_s,\ A(s),\ H_s,\ S_s)\} \tag{3.16}$$

*where:*

- $B_s$represents aggregated behavioural indicators (e.g., SCE counts, relative frequencies, intensity measures such as mean or maximum magnitudes) associated with segment $s$;
- $A(s)$denotes infrastructure and contextual attributes of segment $s$;
- $H_s$represents crash occurrence measures (frequency dimension);

- $S_s$represents crash severity outcomes linked to segment $s$.

At this macro scale, behavioral variables are aggregated across all trips traversing the same roadway segment. Operationally, aggregation is implemented by grouping all event records by segment identifier and applying count-based and descriptive statistical functions to each behavioral variable. Because infrastructure attributes and crash counts are already spatially harmonized with event data, the resulting indicators jointly reflect behavioral concentration patterns and structural roadway conditions.

The segment-level dataset therefore captures spatial heterogeneity in behavioral instability and crash performance across the road network. It allows behavioral precursors to be interpreted within their broader infrastructural and exposure context, supporting network screening and high-risk segment identification.

*2. Trip-Level Dataset (Micro-Level Perspective)*

Let $R(v)$denote the set of trips and $r \in R(v)$a generic trip.

The trip-level analytical dataset is defined as:

$$D_{trip} = \{d_r \mid d_r = (v, r, B_r, E_r, S_r)\} \tag{3.17}$$

*where:*

- $B_r$represents behavioural indicators aggregated along trip $r$;
- $E_r$denotes exposure-related descriptors associated with the trip (e.g., duration, distance, mean traffic intensity);
- $S_r$represents safety outcomes or severity-related proxies linked to the trip context.

At this micro scale, aggregation is performed by grouping events by unique trip identifiers and computing within-trip totals, proportional distributions, and dispersion measures for each SCE category. Infrastructure attributes are not redefined at this stage; instead, the analytical focus remains on preserving behavioral continuity within temporally bounded journeys. Safety outcome indicators are attached after behavioral aggregation, maintaining a clear temporal relationship between behavioral patterns and potential consequences.

The trip-level dataset captures behavioral dynamics within individual driving episodes, reflecting driver-specific operational responses and temporal variability rather than spatial clustering effects.

*3.3.3.4 Conceptual Importance of the Dual Perspective*

The simultaneous consideration of macro- and micro-level structures is fundamental for comprehensive safety analysis. Segment-level aggregation reveals how behavioral instability concentrates spatially within specific roadway environments where infrastructure design, exposure, and crash outcomes interact. Trip-level aggregation, by contrast, reveals how behavioral variability manifests within individual journeys independent of fixed spatial boundaries.

Together, these complementary perspectives allow the framework to distinguish between:

- Infrastructure-influenced behavioral clustering (network effect), and
- Driver-level behavioral instability processes (operational effect).

This multi-scale structure strengthens the explanatory depth, internal consistency, and policy relevance of the modelling framework developed in subsequent phases. It ensures that behavioral signals, infrastructure characteristics, crash frequency, and crash severity are jointly represented across both spatial and temporal dimensions, enabling a system-level understanding of crash risk mechanisms.

### *3.3.4. Step 4 Exposure Estimation*

**Exposure Estimation and Integration**

Exposure measures represent the intensity of road use and the level of interaction between vehicles and infrastructure. Indicators such as Average Annual Daily Traffic (AADT), traffic counts, and congestion proxies quantify demand and allow crash occurrence to be interpreted relative to traffic volume rather than raw counts (Hauer, 2001; Lord & Mannering, 2010). Because both Safety-Critical Events (SCEs) and crashes increase mechanically with traffic flow, exposure provides the necessary denominator for distinguishing high-demand locations from genuinely high-risk conditions.

The final step of Phase 1 estimates and integrates traffic exposure to normalize behavioral and crash outcomes across the network. Without exposure adjustment, segments with higher traffic volumes may appear riskier simply due to greater demand. By incorporating exposure, behavioral event rates and crash outcomes can be interpreted relative to traffic activity, enabling valid comparisons across segments and operating environments.

**Exposure sources and construction**

Exposure indicators are derived primarily from traffic volume measures, particularly AADT where available. When direct counts are incomplete, proxy indicators inferred from telematics data such as vehicle presence, pass counts, or relative flow intensity are used. These proxy measures are calibrated to align with the road network structure to ensure consistency with observed traffic patterns. This process produces exposure descriptors at both segment and trip levels. These descriptors may subsequently be incorporated into modelling as offsets, weights, or normalization factors, depending on the analytical objective.

**Integration with the Multi-Scale Safety Dataset**

Exposure is integrated after Step 3, once the segment- and trip-level analytical units are constructed. At this stage, exposure becomes operationally meaningful because behavioral, infrastructural, and crash variables are already harmonized within consistent analytical structures.

1. *Segment-Level Exposure Integration*

Let $E_s$denote exposure associated with road segment $s$(e.g., AADT or telematics-derived flow proxy). The exposure-extended dataset is defined as:

$$D_{seg}^{exp} = \{d_s^{exp} \mid d_s^{exp} = (s, B_s, A(s), H_s, S_s, E_s)\} \quad (3.18)$$

where:

- $B_s$: aggregated behavioural indicators,
- $A(s)$: infrastructure/context attributes,
- $H_s$: crash frequency,
- $S_s$: crash severity,
- $E_s$: segment exposure descriptor.

This structure supports exposure-adjusted crash frequency modelling (e.g., via offsets in count models) and enables normalization of behavioral rates (e.g., SCEs per unit exposure), improving comparability across segments with heterogeneous traffic demand.

2. *Trip-Level Exposure Integration*

Let $E_r$denote exposure associated with trip $r$(e.g., travelled distance, duration, or route-level intensity proxy). The trip-level dataset becomes:

$$D_{trip}^{exp} = \{d_r^{exp} \mid d_r^{exp} = (v, r, B_r, E_r, S_r)\} \quad (3.19)$$

where:

- $B_r$: trip-aggregated behavioural indicators,
- $E_r$: trip exposure descriptor,
- $S_r$: severity-related outcome or proxy.

This formulation enables behavioral risk to be interpreted relative to the amount of driving performed, supporting analysis of behavioral risk accumulation and its relationship to severity outcomes under varying operating conditions.

Within the modelling framework, exposure plays a central role by normalizing crash frequency so that occurrence is interpreted relative to traffic demand rather than raw counts, standardizing behavioral indicators to enable meaningful comparison of event rates across units with different traffic volumes or trip durations, and contributing to integrated risk estimation where exposure interacts with crash frequency and severity to produce comprehensive safety indicators. The output of this step is therefore an exposure-adjusted, multi-scale safety dataset in which behavioral signals, infrastructure characteristics, crash frequency, crash severity, and traffic demand are jointly represented, forming the analytical foundation for subsequent crash frequency and crash severity modelling and ensuring that safety performance is evaluated relative to demand rather than absolute event totals.

### 3.4.Phase 2 – Parallel Modelling of Crash Frequency and Crash Severity

Following Phase 1, which focused on telematics data engineering, behavioural event construction, and integration of infrastructure, crash, and exposure information, the research proceeds to the predictive modelling stage. The outcome of the previous phase is a multi-scale, exposure-adjusted analytical dataset in which Safety-Critical Events, roadway characteristics, crash records, and contextual variables are jointly represented at both segment and trip levels. The modelling framework developed in this phase is designed to be generalizable and transferable across different road networks and datasets. It provides a methodological approach that can be applied wherever behavioural telematics data, infrastructure attributes, crash records, and exposure measures are available. This generalized structure supports replication and adaptation in diverse safety contexts. Hence,Phase 2 aims to quantitatively model the mechanisms governing crash occurrence and crash consequences. While Phase 1 established the behavioural and infrastructural foundations of safety risk, this phase translates those components into statistical and machine-learning models that estimate crash frequency and crash

severity as complementary dimensions of road safety. A key feature of the framework is the adoption of a parallel modelling strategy in which crash occurrence and crash consequences are treated as distinct but interrelated processes. This distinction reflects the understanding that factors influencing whether a crash occurs may differ from those determining the severity of its outcomes. Accordingly, modelling proceeds along two analytical streams:

**Crash frequency modelling**

- Negative Binomial Regression
- Artificial Neural Networks (FNN / MLP)

**Crash severity modelling**

- Logistic Regression
- Artificial Neural Networks (FNN / MLP classifiers)

These models are estimated at both the segment and trip levels, allowing safety to be examined at different analytical scales. The segment-level specification reflects infrastructure-related safety performance across the road network, while the trip-level specification captures behavioural variability and driver–environment interactions along individual trajectories. By combining these two perspectives, safety is interpreted simultaneously as a spatial characteristic of the network and as a behavioural process emerging from vehicle movements. In this phase, the risk-based framework is translated into quantitative form by connecting behavioural indicators, infrastructure attributes, contextual variables, and exposure measures with observed crash outcomes. Crash frequency modelling estimates the probability of crash occurrence under specific system conditions, whereas crash severity modelling assesses the likelihood of more serious consequences once a crash has occurred. Modelling these components separately allows road safety to be decomposed into two fundamental mechanisms: occurrence and consequence.

Crash risk is treated as a latent construct that cannot be observed directly but can be inferred through probabilistic relationships. The models describe how behavioral instability, roadway characteristics, environmental context, and exposure jointly influence safety outcomes. The methodological strategy integrates econometric and machine-learning techniques to balance interpretability and predictive strength: statistical models offer transparent, explainable associations between predictors and crashes, while neural-network approaches capture nonlinear patterns embedded in telematics-derived data.

The outputs of Phase 2 consist of two probabilistic indicators: the expected crash frequency (conceived as a driver of crash likelihood) and the predicted probability of severe outcomes (representing vulnerability). These measures do not yet define overall crash risk; rather, they provide the foundational components that are subsequently combined with exposure metrics to construct a composite risk index in Phase 3. Through this structured modelling framework, Phase 2 establishes the critical analytical bridge between integrated behavioral–infrastructural data and network-level risk evaluation, supporting proactive and evidence-based safety assessment across the transport system.

Categorical binarization and collinearity control. Before Phase 2 modelling, categorical predictors across behavioral, infrastructural, exposure related, and contextual domains were binarized (one hot encoding with a held out reference; rare level consolidation; ordinal handled as binary only if non monotone/sparse). To manage redundancy after binarization, we used a collinearity matrix tailored to data type—Cramér's V for categorical–categorical pairs, point biserial for binary–continuous, and Pearson/Spearman for continuous–continuous—to screen pairwise dependence (e.g., near duplicates, complements, over granular dummies). We then computed Variance Inflation Factors (VIFs) from provisional GLM design matrices to assess system level multicollinearity. When clusters were flagged, we dropped or merged redundant indicators, retaining the variable with stronger bivariate association with the outcome and clearer substantive meaning. This collinearity matrix + VIF workflow yields a parsimonious, identifiable, and policy interpretable predictor set for crash frequency and crash severity models.

### *3.4.1. Econometric Modelling of Crash Frequency (Negative Binomial Model)*

#### *3.4.1.1 Conceptual Basis and Negative Binomial Specification*

Within the parallel modelling framework, the first analytical component focuses on estimating crash occurrence. While crash severity modelling addresses the consequences once crashes occur, quantifying crash frequency is essential for evaluating overall safety performance. Crash frequency models estimate the expected number of crash events under varying behavioural, infrastructural, contextual, and exposure conditions. In this research, crash occurrence is interpreted as a probabilistic outcome influenced by driver behaviour, roadway characteristics, environmental conditions, and traffic demand.

Raw crash counts alone cannot adequately represent safety performance because observed crashes reflect both intrinsic crash risk and differences in traffic exposure. Locations with higher traffic volumes may experience more crashes even if their inherent safety performance is superior to lower-volume sites. Therefore, crash occurrence must be modelled as a stochastic process that explicitly accounts for exposure and system heterogeneity rather than relying on simple count comparisons.

Crash data are discrete, non-negative integers that typically exhibit overdispersion, meaning that the variance exceeds the mean. This phenomenon arises due to unobserved heterogeneity across modelling units, such as differences in driver populations, traffic composition, environmental conditions, roadway design features, enforcement intensity, spatial clustering, and reporting practices (Cameron & Trivedi, 2013). For this reason, the Negative Binomial (NB) model provides an appropriate statistical framework, as it accommodates overdispersion while allowing exposure to be incorporated explicitly (Lord & Mannering, 2010; Miaou, 1994).

Crash occurrence is modelled as a stochastic count process within a Generalised Linear Model (GLM) framework using a Negative Binomial (NB) error distribution. Let $Y_i$denote the observed crash count for modelling unit $i$(e.g., road segment or trip). The response variable is discrete and non-negative, representing the total number of crashes observed within a defined spatial and temporal unit. Accordingly, crash frequency is interpreted as the expected number of crash events occurring at a site over a specified period (e.g., one year).

Formally, crash counts are assumed to follow a Negative Binomial distribution:

$$Y_i \sim \mathrm{NB}(\mu_i, k) \tag{3.20}$$

where $\mu_i$represents the expected crash frequency and $k$(or equivalently $\alpha$) is the dispersion parameter. The NB model extends the Poisson model by relaxing the restrictive equidispersion assumption. While the Poisson distribution imposes $\mathrm{Var}(Y_i) = \mu_i$, crash data in real-world road networks typically exhibit overdispersion, meaning that the variance exceeds the mean. The NB variance structure is given by:

$$\mathrm{Var}(Y_i) = \mu_i + \alpha\mu_i^2 \tag{3.21}$$

This quadratic variance formulation allows variability to increase with the expected crash frequency, reflecting empirical evidence that higher-risk conditions are often associated with greater uncertainty and instability. The dispersion parameter captures unobserved heterogeneity arising from latent behavioural differences, traffic composition, enforcement intensity, environmental variability, spatial clustering, and reporting inconsistencies. By accommodating overdispersion, the NB model provides

unbiased standard errors and more reliable statistical inference compared with simpler Poisson formulations.

By accommodating overdispersion, the Negative Binomial model produces unbiased standard errors, improves goodness-of-fit, and provides more reliable statistical inference compared with the simpler Poisson formulation. Consequently, it represents a robust and theoretically grounded approach for modelling crash frequency within the integrated behavioral–infrastructure–exposure safety framework developed in this dissertation.

#### *3.4.1.2 Verification of Distributional Assumptions*

Before applying a GLM with NB error distribution, it is necessary to verify whether crash data are overdispersed. This is achieved by computing the dispersion parameter $D$, defined as the ratio between the variance and the mean of crash counts:

$$D = \frac{\sigma^2}{\mu} \tag{3.22}$$

where $\sigma^2$denotes the variance of the crash count distribution and $\mu$denotes its mean.

Interpretation of the dispersion ratio is straightforward:

- $D = 1$: crash data are consistent with a Poisson distribution (equidispersion).
- $D > 1$: crash data are overdispersed, supporting the adoption of a Negative Binomial specification.
- $D < 1$: crash data are underdispersed, suggesting a binomial-type structure.

In road safety research, overdispersion is commonly observed due to spatial heterogeneity and behavioural variability across the network. Therefore, the NB specification is generally appropriate and consistent with established modelling practice in crash frequency analysis.

### *3.4.2.* ***Model Structure, Functional Form, and Treatment of Exposure***

Crash frequency for modelling unit $i$(segment or trip) is linked to explanatory variables through a logarithmic link function within the Generalized Linear Model framework. The log-link ensures strictly positive predicted values and enables a multiplicative interpretation of model coefficients. Let $E_i$denote the exposure measure (e.g., Annual Average Daily Traffic for segments or travelled distance/time for trips), and let $X_i$denote a vector of behavioural, infrastructural, and contextual predictors.

**Exposure Treated as an Offset**

When exposure is treated as an offset, the model is specified as:

$$\log(\mu_i) = \log(E_i) + \beta_0 + \beta' X_i \tag{3.23}$$

or equivalently,

$$\mu_i = E_i \exp(\beta_0 + \beta' X_i) \tag{3.24}$$

In this formulation, exposure represents crash opportunity and is assumed to scale proportionally with crash occurrence. Treating exposure as an offset separates traffic demand from intrinsic safety characteristics. Consequently, the estimated coefficients $\beta$describe the effect of behavioural, infrastructural, and contextual variables on crash rates rather than raw crash counts. This specification maintains a clear conceptual distinction between exposure (opportunity for crashes) and structural risk factors (determinants of crash propensity).

In this research, exposure is primarily incorporated as an offset to preserve consistency within the risk-based modelling framework and to ensure that parameter estimates reflect intrinsic safety performance.

**Exposure Treated as a Predictor (Alternative Specification)**

Alternatively, exposure can be included directly as an explanatory variable:

$$\log(\mu_i) = \beta_0 + \beta_1 \log(E_i) + \beta' X_i \tag{3.25}$$

In this case, the coefficient $\beta_1$estimates the elasticity between crash frequency and exposure.

- $\beta_1 = 1$: proportional scaling between crashes and exposure
- $\beta_1 < 1$: sublinear growth (e.g., congestion effects or behavioural adaptation)
- $\beta_1 > 1$: superlinear growth (e.g., increased interaction density)

This specification provides flexibility in modelling nonlinear exposure effects and allows the exposure–crash relationship to be empirically estimated. While both formulations are applicable at the segment and trip levels, the offset specification is adopted as the primary modelling strategy in

this dissertation to clearly distinguish crash opportunity from intrinsic risk determinants, while the predictor-based formulation is recognized as a complementary approach when scaling effects are of analytical interest.

**Model Adequacy and Predictive Performance**

Model adequacy and predictive performance are evaluated using several complementary diagnostics.First, the deviance ratio (regression deviance divided by degrees of freedom) is used to assess overall model fit. Values close to unity indicate appropriate specification. The overall statistical significance of the model is assessed using the chi-square ($\chi^2$) statistic. Second, predictive accuracy is evaluated using the Root Mean Squared Error (RMSE) and the percentage ratio of predicted to observed crashes. Let $Y_i$denote observed crash frequency and $\hat{F}_i$the predicted frequency. The RMSE is computed as:

$$\text{RMSE} = \sqrt{\frac{1}{|\ I\ |}\sum_{i\in I}(\ Y_i - \hat{F}_i)^2} \tag{3.26}$$

where $|\ I\ |$represents the total number of modelling units. Lower RMSE values indicate better predictive performance. The comparison between predicted and observed crash totals further assesses potential systematic bias in estimation.

Finally, the magnitude, direction, and statistical significance of estimated coefficients are examined to evaluate the influence of each explanatory variable on crash frequency. Exponentiated coefficients may be interpreted as incidence rate ratios, quantifying multiplicative changes in expected crash counts associated with one-unit variations in predictors. Within the broader risk-based methodological framework of this research, the Negative Binomial model provides a probabilistic estimate of expected crash occurrence under varying behavioral, infrastructural, contextual, and exposure conditions. Crash risk is not directly observable; rather, it is inferred through such probabilistic models. The estimated expected crash frequency represents the *occurrence component* of risk and constitutes a fundamental input for subsequent integration with crash severity modelling and exposure measures in the composite risk index developed in the following phase.

### *3.4.3. Econometric Modelling of Crash Severity (Binomial Logit Model)*

**Conceptual Basis and Logistic Specification**

Following the estimation of crash frequency, the modelling framework proceeds to the analysis of crash consequences to capture the second fundamental dimension of road safety risk, namely severity. While frequency modelling quantifies the expected number of crash events occurring within a defined spatial and temporal unit, severity modelling evaluates the conditional probability that a crash results in serious outcomes. This distinction is methodologically essential, as the mechanisms governing crash occurrence differ from those influencing crash consequences. Whereas crash frequency is primarily associated with exposure, interaction density, and behavioral conflict generation, crash severity is more directly influenced by impact energy, roadside protection, geometric configuration, environmental visibility, and behavioral extremity. Modelling severity as a separate probabilistic process therefore enables a clearer representation of consequence vulnerability within the overall risk structure.

In this dissertation, crash severity is modelled using a Binomial Logit regression framework. The choice of binomial logistic specification is motivated by the nature of the dependent variable, which is defined as binary. For each modelling unit $i$(segment or trip), the severity outcome is represented by a binary variable $Y_i^S \in \{0,1\}$, where $Y_i^S = 1$denotes a severe crash outcome and $Y_i^S = 0$denotes a non-severe outcome. The operational definition of "severe" depends on the available crash database and may correspond to fatal or serious injury, or to a high-damage classification, thereby ensuring transferability across different empirical contexts. Given this binary structure, the objective is to estimate the conditional probability $\pi_i = P(Y_i^S = 1 \mid X_i)$, where $X_i$represents the vector of behavioural, infrastructural, and contextual explanatory variables.

Logistic regression is widely adopted in crash severity research because it is specifically designed to model binary outcome variables while constraining predicted values within the interval $[0, 1]$, thereby ensuring consistency with probability theory. In the context of injury severity analysis, the binomial logit framework enables the conditional probability of severe outcomes to be expressed as a function of behavioural characteristics, infrastructure design attributes, traffic conditions, and contextual vulnerability factors (Savolainen et al., 2011; Mannering & Bhat, 2014). These studies emphasize that severity modelling requires a probabilistic structure capable of capturing the complex interaction between driver behavior, roadway environment, and crash circumstances. Furthermore, crash severity

risk often exhibits nonlinear escalation effects with respect to variables such as speed, roadside protection, geometric configuration, and adverse environmental conditions. The log-odds transformation embedded within the logistic model provides a statistically stable and interpretable framework for estimating such relationships, allowing coefficients to be interpreted through odds ratios.

The binomial logit model links the severity probability to the explanatory variables through the log-odds transformation:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \gamma_0 + \gamma' X_i \tag{3.27}$$

where $\gamma_0$denotes the intercept term and $\gamma$is the vector of parameters associated with the predictors. The transformation of probability into log-odds removes the natural bounds of probability (0 and 1), thereby allowing a linear additive specification in the transformed space. This structure ensures that predicted values remain strictly within the admissible probability range after back-transformation.

The predicted probability of severe outcomes is obtained by applying the logistic function:

$$\pi_i = \frac{e^{\gamma_0+\gamma' X_i}}{1+e^{\gamma_0+\gamma' X_i}} \tag{3.28}$$

This functional form maps the linear combination of predictors into a probability bounded between 0 and 1. Conceptually, the numerator represents the combined influence of vulnerability factors that elevate the likelihood of severe consequences, while the denominator normalizes the expression to ensure theoretical consistency with probability theory. The coefficients $\gamma$can be interpreted through odds ratios $e^{\gamma_j}$, which represent the multiplicative change in the odds of severe outcomes associated with a one-unit increase in predictor $X_j$, holding other variables constant. This interpretation provides meaningful behavioural and infrastructural insights into severity escalation mechanisms.

The binomial logit specification is particularly suitable for crash severity modelling because injury outcomes often exhibit nonlinear escalation effects. For example, incremental increases in speed environment, reduced roadside protection, or diminished visibility may lead to disproportionately

higher probabilities of severe injury outcomes. The logistic transformation captures this nonlinear relationship between explanatory variables and outcome probabilities while ensuring that predicted probabilities remain bounded between 0 and 1. Model parameters are estimated using Maximum Likelihood Estimation (MLE), which provides consistent and efficient estimates under standard assumptions. Moreover, modelling crash severity separately from crash frequency allows vulnerability-related factors influencing injury consequences to be identified independently of exposure-driven mechanisms governing crash occurrence.

**Model Adequacy and Predictive Performance**

Model adequacy and predictive performance are evaluated through complementary goodness-of-fit and diagnostic measures. Estimation is conducted using Maximum Likelihood techniques, and overall model fit is assessed through the log-likelihood value at convergence and the Likelihood Ratio (LR) test comparing the full model with the intercept-only specification. Explanatory improvement is quantified using McFadden's pseudo-$R^2$, defined as:

$$R^2_{McFadden} = 1 - \frac{\ln L_{full}}{\ln L_{null}} \quad (3.29)$$

where $L_{full}$ and $L_{null}$ denote the likelihoods of the full and null models, respectively. Although pseudo-$R^2$ values are not directly comparable to ordinary least squares $R^2$, they provide an indication of relative improvement over a baseline model.

Predictive performance is further evaluated through classification-based diagnostics, including the confusion matrix, overall accuracy, sensitivity (true positive rate), specificity (true negative rate), and precision metrics. Given that severe crashes may represent a relatively smaller proportion of total crashes, particular attention is paid to sensitivity and balanced accuracy to ensure that the model effectively identifies high-risk severe outcomes. Additionally, the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) are used to assess discriminatory power in a threshold-independent manner. Higher AUC values indicate stronger capability of the model to distinguish between severe and non-severe outcomes across varying probability thresholds. Calibration assessment is performed by comparing predicted probabilities with observed severity proportions across probability intervals, ensuring that predicted risk levels correspond to empirical frequencies.

Within the integrated modelling framework, the predicted severity probabilities $\hat{\pi}_i$represent the vulnerability component of crash risk. When combined with the expected crash frequency derived from the Negative Binomial count model and with exposure measures, they contribute to the construction of composite crash risk indicators. This parallel probabilistic structure modelling occurrence and severity separately ensures methodological coherence, interpretability, and transferability of the proposed risk-based safety assessment framework across segment-level and trip-level analytical scales.

### *3.4.4. Neural Network-Based Modelling of Crash Frequency and Severity*

In addition to conventional econometric models, this research employed Feedforward Neural Networks (FNNs) as the principal artificial intelligence approach for crash frequency and crash severity modelling. The selection of FNN was motivated by both the characteristics of the available dataset and the analytical objectives of the study.

First, the assembled datasets primarily consisted of structured tabular variables combining telematics-derived behavioural indicators, roadway infrastructure attributes, environmental conditions, crash characteristics, and traffic exposure measures. Such heterogeneous predictor structures are well suited to Feedforward Neural Networks, which are capable of learning nonlinear relationships and interaction effects among multiple input variables without requiring predefined functional assumptions. This characteristic is particularly relevant in road safety analysis, where crash risk often emerges through complex interactions between behaviour, roadway context, and exposure conditions. Second, compared with traditional regression-based approaches, FNN models provide greater flexibility in capturing threshold effects, non-monotonic responses, and higher-order dependencies that may remain hidden under linear or log-linear formulations. This was particularly important for the present study, where behaviours such as speeding, sudden braking, and instability may influence crash risk differently depending on traffic conditions, road type, and infrastructure complexity.

Third, the choice of FNN was also influenced by the structure of the telematics inputs. Although the raw data originated from trajectories, the final modelling datasets were transformed into aggregated behavioural indicators and segment-level or trip-level explanatory variables rather than long continuous sequential streams. For this reason, sequence-based deep learning architectures such as Long Short-Term Memory (LSTM) networks or Transformer models were considered less essential

for the present modelling stage, as these methods are generally most advantageous when long temporal dependencies in raw time-series data are explicitly retained.

Fourth, compared with tree-based ensemble methods such as Random Forest, XGBoost, or LightGBM, Feedforward Neural Networks offered a consistent modelling architecture applicable to both regression-type frequency tasks and classification-based severity tasks within a unified analytical framework. This facilitated methodological consistency across modelling phases while maintaining strong predictive capability.

Importantly, the selection of FNN should not be interpreted as implying universal superiority over all alternative artificial intelligence models. Rather, it represents a context-dependent methodological choice that balances predictive performance, computational feasibility, scalability, and compatibility with the structured datasets developed in this dissertation. To strengthen interpretability, Explainable Artificial Intelligence techniques, including SHAP values and permutation importance analysis, were additionally employed to identify the most influential predictors and mitigate common concerns associated with black-box models.

Therefore, the adoption of Feedforward Neural Networks was considered methodologically appropriate for the objectives of this research, namely, to complement econometric inference with nonlinear predictive modelling while preserving sufficient transparency for road safety decision-making. Future research could further extend this framework by benchmarking FNN performance against gradient boosting methods, recurrent neural networks, graph neural networks, and other emerging architectures under identical validation conditions.

This capability is particularly valuable in transportation safety analysis, where crash risk emerges from intricate interactions among behavioral, infrastructural, and environmental factors. Similar comparative studies between econometric and artificial intelligence approaches have shown that neural network models can achieve improved predictive performance by capturing these nonlinear relationships, while econometric models remain useful for interpretability and policy insights (Barabino & Ventura, 2025).

Let the input vector for modelling unit $i$(segment or trip) be defined as

$$X_i = (x_{i1}, x_{i2}, \ldots, x_{ip}) \tag{3.30}$$

where $p$ denotes the number of explanatory variables representing behavioural indicators, roadway infrastructure characteristics, contextual factors, and exposure measures.

Within the neural network, the input variables are processed through a sequence of hidden layers that apply nonlinear transformations to the data. The transformation in the first hidden layer can be expressed as

$$h^{(1)} = f(W^{(1)}X_i + b^{(1)}) \tag{3.31}$$

and for deeper layers the transformation becomes

$$h^{(l)} = f(W^{(l)}h^{(l-1)} + b^{(l)}) \tag{3.32}$$

where $W^{(l)}$denotes the weight matrix associated with layer $l$, $b^{(l)}$represents the bias vector, and $f(\cdot)$is the activation function. Common activation functions include the **Rectified Linear Unit (ReLU), sigmoid, and hyperbolic tangent functions.** These nonlinear transformations allow the network to capture complex interaction patterns among explanatory variables that may not be adequately represented through traditional linear modelling approaches.

Within the modelling framework of this dissertation, the **FNN architecture is applied to both crash frequency prediction and crash severity prediction.** For crash frequency estimation, the neural network produces a predicted crash occurrence level for each modelling unit $i$:

$$\hat{\mu}_i = g(W^{(L)}h^{(L-1)} + b^{(L)}) \tag{3.33}$$

where $\hat{\mu}_i$denotes the predicted crash frequency, $L$represents the output layer, and $g(\cdot)$is the output activation function ensuring predictions are expressed on an appropriate scale. This formulation enables the network to capture nonlinear escalation mechanisms governing crash occurrence, such as compound interactions between behavioural instability, traffic exposure, roadway curvature, and infrastructure characteristics.

For crash severity modelling, the neural network produces a probability representing the likelihood of severe crash outcomes. The output layer applies to a **logistic activation function:**

$$\hat{\pi}_i = \sigma(W^{(L)}h^{(L-1)} + b^{(L)}) \tag{3.34}$$

*where*

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

The resulting value represents the predicted probability that a crash associated with modelling unit $i$results in severe consequences, satisfying the probabilistic constraint

$$0 \leq \hat{\pi}_i \leq 1$$

(3.36)

This probabilistic formulation allows the neural network to model severity vulnerability mechanisms that may arise from nonlinear interactions between behavioral patterns, infrastructure characteristics, and environmental conditions.

Thus in this study, behavioral information derived from NDD is represented through the extraction of **Safety-Critical Events (SCEs)** during the data engineering phase. These events include maneuvers such as sudden braking, rapid acceleration, sharp turning, and speeding episodes that represent behavioral precursors to crash risk. Because the safety implications of these events depend on combinations of driver behavior, roadway design, traffic conditions, and environmental context, the relationship between SCEs and crash outcomes is inherently multidimensional. The hierarchical structure of the neural network allows the model to learn these behavioral–environmental interaction patterns and translate them into predictions of crash frequency and crash severity.

**Model Adequacy and Predictive Performance**

To ensure methodological robustness, the predictive capability of the Feedforward Neural Network was evaluated through a structured validation strategy. Initially, the modelling dataset was partitioned into training, validation, and testing subsets. The training dataset was used to estimate the neural network parameters through forward and backpropagation. The validation dataset supported hyperparameter tuning and enabled the implementation of early stopping procedures to prevent overfitting. The independent testing dataset was then used to evaluate the predictive capability of the model using performance metrics appropriate to the modelling task.

After confirming that the neural network exhibited satisfactory predictive capability, the model was re-estimated using the complete dataset. This final estimation step allows the model to fully exploit the informational content of the available data and provides more stable prediction outputs for interpretation and variable-importance analysis. The results reported in this research therefore correspond to the neural network estimated on the full dataset.

For crash frequency prediction, which represents a regression problem, predictive performance was evaluated using the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE). These indicators measure the deviation between predicted and observed crash frequencies and are defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \tag{3.37}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3.38}$$

where $y_i$ represents the observed crash frequency and $\hat{y}_i$ the predicted crash frequency.

For crash severity prediction, model performance was evaluated using the **F1-score**, which balances precision and recall and is particularly suitable for crash severity datasets where severe crashes represent a minority class. Precision and recall are defined as

$$Precision = \frac{TP}{TP + FP} \tag{3.39}$$

$$Recall = \frac{TP}{TP + FN} \tag{3.40}$$

and the F1-score is computed as

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{3.41}$$

Beyond predictive accuracy, model interpretability was examined using **SHAP (SHapley Additive explanations)** values, which quantify the contribution of each predictor variable to the neural network predictions. The SHAP value for feature $j$is defined as

$$\phi_j = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!}\left[f_{S\cup\{j\}}(x) - f_S(x)\right] \tag{3.42}$$

where $F$ represents the full set of input features and $S$denotes subsets of features excluding feature $j$. SHAP values allocate the difference between the model prediction and the baseline expectation across input variables, enabling the relative influence of behavioural indicators, infrastructure characteristics, contextual variables, and exposure measures to be interpreted. Through this combined evaluation framework, the Feedforward Neural Network provides both **strong predictive capability and interpretable insights into the nonlinear mechanisms linking driver behavior, infrastructure conditions, environmental factors, and traffic exposure to crash frequency and crash severity outcomes.**

Within this framework, the outputs of Phase 2 are interpreted as probabilistic indicators of safety system performance rather than final measures of crash risk: crash frequency models estimate the

likelihood of crash occurrence, while crash severity models estimate the probability of severe outcomes. Because exposure has already been incorporated within the modelling structure, these outputs provide exposure-adjusted estimates that form the basis for the subsequent derivation of composite crash risk indicators for network-level safety assessment.

### *3.4.5 Model Assumptions and Methodological Conditions*

To enhance methodological transparency and ensure reliable inference, the principal assumptions underlying the modelling approaches adopted in this research are explicitly stated.

**Negative Binomial Regression Assumptions**

The Negative Binomial Regression model was employed for crash frequency analysis because crash counts are discrete, non-negative, and commonly exhibit overdispersion. Its principal assumptions are as follows: (i) the dependent variable represents count outcomes measured over consistent analytical units; (ii) observations are conditionally independent given the explanatory variables; (iii) the conditional mean is linked to predictors through the logarithmic link function; (iv) the variance exceeds the mean according to the Negative Binomial variance structure; (v) explanatory variables do not exhibit severe multicollinearity; and (vi) exposure is appropriately represented through an offset term or explanatory variable. These conditions were examined through dispersion statistics, correlation screening, and Variance Inflation Factor checks.

**Logistic Regression Assumptions**

The Logistic Regression model was used for crash severity analysis where the dependent variable is binary (severe versus non-severe outcome). Its principal assumptions are: (i) the dependent variable is dichotomous and correctly classified; (ii) observations are independent; (iii) the log-odds of the outcome are approximately linearly related to continuous predictors; (iv) no severe multicollinearity exists among explanatory variables; and (v) sufficient observations are available in each outcome category to permit stable parameter estimation. These conditions were assessed through class-balance inspection, correlation analysis, and coefficient stability checks.

**Feedforward Neural Network Conditions**

The Feedforward Neural Network (FNN) model was adopted to capture nonlinear relationships and higher-order interactions among behavioural, infrastructural, contextual, and exposure-related variables. Unlike econometric models, neural networks rely less on strict distributional assumptions; however, several methodological conditions remain important. These include: (i) training and

evaluation data originating from comparable population structures; (ii) sufficient sample size for stable weight estimation; (iii) appropriate encoding and scaling of predictors; (iv) the presence of nonlinear relationships that justify flexible modelling structures; and (v) control of overfitting through validation procedures, early stopping, and hyperparameter tuning. These conditions were addressed through systematic preprocessing and repeated performance assessment.

### 3.5. Phase 3: Integrated Risk Index Development, Normalization, and Spatial Risk Mapping

Phase 3 represents the operational stage of the methodological framework in which the predictive outputs generated in Phase 2 are transformed into a unified, interpretable, and spatially actionable crash risk structure. While the preceding phases focused on behavioral event construction and probabilistic modelling of crash occurrence and severity, this phase integrates those components into a comprehensive risk assessment tool capable of supporting practical safety decision-making. The objective is to move from statistical prediction to an operational representation of safety performance across the road network. Such integration is consistent with contemporary risk-based safety management approaches that emphasize proactive assessment using predictive indicators rather than relying solely on historical crash records (Nguyen et al., 2024).

Phase 2 produces two key probabilistic outputs: the expected crash occurrence (frequency component) and the probability of severe outcomes (severity component). These components correspond to the likelihood and consequence dimensions of safety risk. Phase 3 integrates them into a single composite metric consistent with risk-based safety theory, thereby operationalizing the concept of crash risk as the interaction between occurrence and consequence processes. This integration enables the transformation of model-based predictions into spatial safety intelligence capable of supporting network screening and evidence-based intervention planning.

Recent telematics-based safety research highlights the importance of incorporating behavioral indicators derived from Naturalistic Driving Data into safety assessment frameworks. Behavioral precursors such as volatility in speed, abrupt maneuvers, and instability measures are strongly associated with both crash probability and injury severity, while near-miss events captured through connected vehicle systems occur far more frequently than crashes and therefore provide valuable early indicators of hazardous conditions. Integrating these behavioral signals with predictive crash models enables proactive identification of high-risk locations and strengthens the capacity of data-driven safety management systems (Shen & Yang, 2020; Wali & Khattak, 2020; Hashmi et al., 2024).

### *3.5.1. Exposure Treatment and Risk Formulation*

A central methodological issue in integrated road safety assessment is the treatment of exposure, which represents the level of interaction between road users and hazardous conditions and thus determines the opportunity for crashes to occur. In transportation safety research, exposure is commonly operationalized through measures such as traffic volume, vehicle-kilometers travelled, or travel time, all of which quantify the amount of traffic subject to potential risk (Lord & Mannering, 2010). Without appropriate exposure adjustment, locations carrying higher traffic demand may appear disproportionately unsafe simply because they experience more interactions, rather than because of inherently hazardous conditions.

Two principal modelling strategies are typically employed. When exposure is incorporated as an offset in crash frequency models, predicted crash occurrence is normalized by traffic demand, producing estimates of crashes per unit of exposure. This approach is widely used in Safety Performance Functions and network screening studies because it enables fair comparison across segments with different usage levels while isolating structural safety characteristics (AASHTO, 2020). Under this formulation, exposure influences risk indirectly through the occurrence term.

Alternatively, exposure may be treated as an independent contributor to risk. In this case, predicted occurrence reflects intrinsic crash likelihood conditional on roadway, behavioral, and environmental factors, while exposure represents the magnitude of traffic subjected to that risk. This formulation aligns with classical risk theory, where overall risk arises from the combined effects of hazard probability, exposure, and consequence (Ronan et al., 2020).

In this framework, exposure can be incorporated in two alternative ways. When it is included as an offset in the crash-frequency model, predicted crash occurrence is already normalized by traffic demand, and therefore represents exposure-adjusted crash intensity. Under this specification, the integrated risk measure reflects intrinsic safety performance rather than total societal burden. This is the formulation adopted in the present research, as it aligns with the objective of identifying road elements with elevated inherent risk independent of traffic volume.

(i) **Exposure incorporated as an offset**

With exposure embedded in the occurrence model, expected crashes are expressed per unit of exposure, enabling comparable evaluation across segments with different demand levels. The resulting risk index is defined as:

$$R_{s,u} = H_{s,u}(E_{s,u}) \times S_{s,u} \tag{3.43}$$

where $H_{s,u}(E_{s,u})$is the exposure-adjusted occurrence component and $S_{s,u}$is the predicted severity probability.

(ii) **Exposure treated as an independent factor**

Alternatively, exposure may be included explicitly in the final risk expression, in which case crash occurrence reflects intrinsic likelihood, and exposure contributes separately to total risk. This leads to the classical multiplicative form:

$$R_{s,u} = H_{s,u} \times E_{s,u} \times S_{s,u} \tag{3.44}$$

This second specification is conceptually valid but is not adopted here. Using separate models for the risk components provides clearer control over the contribution of each element, avoids the complexity of joint modelling structures, and supports targeted safety interpretation. Given these advantages and the methodological aims of this study, the offset-based formulation is retained for all subsequent analyses.

### *3.5.2. Composite Risk Index Construction*

The composite risk index integrates the probabilistic outputs of crash frequency and crash severity models into a single operational measure of safety risk. Frequency models estimate how often crashes are expected to occur under specified traffic, infrastructural, and behavioral conditions, whereas severity models quantify the likelihood that those crashes will result in serious or fatal outcomes. Individually, these components provide partial information; their integration yields a comprehensive indicator that captures both the probability of occurrence and the magnitude of consequences, consistent with contemporary risk-based road safety assessment frameworks (Lord & Mannering, 2010). When exposure is incorporated as an offset in the frequency model, predicted crash occurrence

already reflects traffic demand. In this case, the composite risk index for each analytical unit is defined as:

$$R_i = \lambda_i \times P_i \tag{3.45}$$

where:

• $R_i$ = composite risk score for unit i

• $\lambda_i$ = predicted crash occurrence (exposure-adjusted)

• $P_i$ = probability of severe outcomes

This formulation represents the expected severity-weighted crash occurrence under given behavioral, infrastructural, and environmental conditions. It captures both how frequent crashes are likely to occur and how severe their outcomes are expected to be, thereby providing a unified representation of safety performance. The multiplicative structure reflects the conceptual understanding that safety risk emerges from the interaction between crash likelihood and outcome severity rather than from either component alone (Ronan et al., 2020; Winlaw et al., 2019).

Unlike traditional crash counts, the index incorporates predictive information derived from behavioral indicators, infrastructure characteristics, contextual variables, and exposure conditions. Such integration is particularly important in telematics-based safety analysis, where near-miss events and operational instability can reveal latent risk before severe crashes occur (Mannering et al., 2016). Consequently, the composite index supports proactive identification of hazardous locations even in the absence of large numbers of historical crashes, facilitating preventive interventions and evidence-based safety management.

The Composite Crash Risk Index (CCRI) is estimated using predictions obtained from the Negative Binomial Regression crash frequency models and the Logistic Regression crash severity models. Predicted crash frequency values combine with the corresponding predicted probabilities of severe outcomes to compute the integrated risk measure for each analytical unit. This approach ensures that traffic exposure is incorporated through the offset term in the econometric models, allowing the estimation of exposure-adjusted crash risk. Results from the Feed-Forward Neural Network models

are not used for the composite index because neural networks do not support offset specifications, meaning exposure cannot be incorporated in the same exposure-normalized manner.

### *3.5.3. Risk Normalization and Road Network Ranking*

The raw composite risk index $R_{s,u}$, although theoretically consistent, is not directly suitable for operational network screening. Raw risk scores may exhibit substantial variability across the network due to differences in traffic demand, environmental conditions, and infrastructure characteristics. Direct comparison of unscaled values may therefore be misleading. To enhance interpretability and support visualization, the composite risk index is normalized.

Min–max normalization is applied to rescale risk values to the interval $[0, 1]$:

$$R^*_{s,u} = \frac{R_{s,u} - \min(R)}{\max(R) - \min(R)} \tag{3.46}$$

where $R^*_{s,u}$denotes the normalized risk score for segment $s$in unit $u$, and $\min(R)$and $\max(R)$represent the minimum and maximum risk values observed in the dataset.

This transformation preserves relative differences while enabling meaningful comparison across segments and facilitating integration with spatial analysis tools. The normalized score represents the position of each unit within the overall risk distribution rather than an absolute measure of danger. Once the crash risk values have been computed for all paths, a ranking procedure is required to identify High Crash Risk Sites. A variety of methods exist for constructing risk-ranking scales; however, in this study a five-level scale is adopted based on the empirical distribution of the computed risk scores. Quantile-based scaling is used because it adapts naturally to the intrinsic distribution of the network, avoids rigid or arbitrary thresholds, and ensures a balanced representation of safety conditions across the entire system. Rather than imposing fixed cut-offs, this approach reflects the *relative* risk hierarchy of the specific network under analysis, which is particularly important for proactive screening and prioritization.

The ranking procedure begins by ordering all paths from the lowest to the highest risk value. Thresholds are then derived using key distributional indicators, including the minimum and maximum values, the lower, middle, and upper quartiles ($Q_1$, $Q_2$, and $Q_3$), and the interquartile range (IQR),

defined as $IQR = Q_3 - Q_1$. The inclusion of the IQR supports the identification of segments that exhibit anomalously high crash risk, which is essential for pinpointing the most critical paths. Traditionally, the IQR is used to detect statistical outliers by extending the first and third quartiles by $1.5 \cdot IQR$. Values lying beyond these extended thresholds are typically removed to prevent distortion of results. In contrast, this study intentionally uses the upper IQR extension **not to remove extreme values**, but to **highlight** them. Extending the upper quartile to $Q_3 + 1.5 \cdot IQR$creates an additional boundary that isolates the highest-risk paths, enabling the construction of a five-level ranking scale in which the top category represents the paths with the greatest safety concern.

The lower IQR extension ($Q_1 - 1.5 \cdot IQR$) does not yield a meaningful additional class because the crash risk metric is non-negative; thus, $(Q_1 - 1.5 \cdot IQR)$is always greater than or equal to zero and does not contribute to the definition of a distinct lower-risk category.

Building on these distributional indicators, a **quantile-based classification** approach is adopted to ensure a balanced representation of safety conditions across the network. This method avoids rigid, arbitrary thresholds and instead reflects the intrinsic distribution of the computed risk values. As such, it captures the *relative* risk hierarchy of the network, which is particularly valuable for proactive road-safety screening.

The normalized crash-risk values are partitioned into five ordinal classes, each corresponding to an equal-probability interval:

- Very Low Risk: 0–20%
- Low Risk: 20–40%
- Medium Risk: 40–60%
- High Risk: 60–80%
- Very High Risk: 80–100%

$$Class_{s,u} = \begin{cases} 1 & R^*_{s,u} \le Q_{0.20} \\ 2 & Q_{0.20} < R^*_{s,u} \le Q_{0.40} \\ 3 & Q_{0.40} < R^*_{s,u} \le Q_{0.60} \\ 4 & Q_{0.60} < R^*_{s,u} \le Q_{0.80} \\ 5 & R^*_{s,u} > Q_{0.80} \end{cases} \tag{3.47}$$

where $Q_{0.20}$, $Q_{0.40}$, $Q_{0.60}$, and $Q_{0.80}$denote the 20th, 40th, 60th, and 80th percentiles of the normalized risk distribution. Thus, this quantile-based strategy ensures that each class contains a meaningful

portion of the network while preserving relative differences between segments. Paths assigned to the fifth class those above the 80th percentile represent the most critical areas and should be prioritized for road-safety interventions.

### *3.5.4. Spatial Risk Mapping and GIS-Based Decision Support*

After normalization and classification, the derived risk metrics are integrated with spatial network data using Geographic Information Systems (GIS), enabling a geographically explicit representation of safety performance across the road network. Each modelling unit—whether a road segment or a trajectory—is assigned a composite risk score, a normalized value, and a categorical risk level, which are spatially joined to the corresponding elements of the digital road network. Through this GIS-based integration, statistical outputs are transformed into spatially interpretable indicators that reveal how crash risk is distributed across the network. To support operational interpretation, the five-level risk-ranking scheme defined earlier based on quartiles and the IQR extension is summarized in (**Fig 3.5)**, which reports the lower and upper limits of each class together with the corresponding color scale used for visualization. A range of GIS visualization techniques such as graduated color symbology, kernel density surfaces, heatmaps, and thematic overlays are employed to illustrate variations in risk intensity and to highlight spatial clustering. These tools assist in identifying high-risk corridors, concentrated clusters of hazardous segments, and areas where targeted interventions may be required. Spatial analysis functions further support the evaluation of patterns of spatial autocorrelation and network continuity, helping determine whether elevated risk levels occur in isolation or form part of broader systemic conditions.

This GIS-based decision-support framework goes beyond traditional hotspot analyses that rely exclusively on historical crash counts. By incorporating predictive indicators derived from behavioral, infrastructural, and contextual variables, the system enables proactive safety assessment, identifying locations where severe crashes are likely to emerge even in the absence of long crash histories. This capability strengthens the ability of the screening process to detect hazardous locations at an early stage and to anticipate emerging risks. The resulting spatial risk structure provides robust decision-support functionality for transportation planners and road-safety authorities. GIS-based risk maps facilitate the prioritization of infrastructure improvements, the strategic deployment of enforcement and speed-management measures, and continuous monitoring of safety conditions over time. By supporting evidence-based planning and more efficient allocation of resources, the GIS framework serves as a cornerstone for proactive, network-wide road-safety management.

*Table 0-1: Five level crash risk ranking scale based on quartile thresholds and IQR extension*

| Level | Ranges values | |
|---|---|---|
| | Lower limit | Higher limit |
| 1 | (Q1-1,5 IQR) | Q1 |
| 2 | Q1 | Q2 |
| 3 | Q2 | Q3 |
| 4 | Q3 | (Q3 + 1,5 IQR) |
| 5 | (Q3 + 1,5 IQR) | MAX |

### *3.5.5 Model Validation Strategy and Performance Evaluation*

To ensure methodological robustness and predictive reliability, a two-stage validation strategy was adopted in this research. Initially, the analytical datasets were partitioned into training, validation, and testing subsets for the purposes of model development, hyperparameter tuning, and independent performance assessment. This preliminary stage was used to evaluate model behavior, detect potential overfitting, and confirm that the selected modelling structures could generalize beyond the training sample.

After satisfactory predictive performance and model stability had been confirmed, the final models reported in this thesis were re-estimated using the full available dataset. This final estimation step enables the models to exploit the complete informational content of the data, improve parameter stability, and generate the most reliable predictions for interpretation, variable importance analysis, and Composite Crash Risk Index construction. Accordingly, the reported coefficients, predictions, and final performance indicators correspond to models estimated on the full dataset, whereas the train–validation–test procedure served as an internal model verification stage.

Validation of Econometric Models

For the crash frequency and crash severity econometric models, validation focused on statistical adequacy, explanatory consistency, and predictive plausibility. Model quality was examined using Log-Likelihood values, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), deviance statistics, and pseudo-$R^2$ measures where appropriate. Statistical significance of coefficients, confidence intervals, and logical consistency of parameter signs were also evaluated. Preliminary split-sample checks further confirmed stable predictive behavior before final full-data estimation.

### Validation of Neural Network Models

For the Feedforward Neural Network models, the training subset was used to estimate network weights, while the validation subset supported hyperparameter tuning and early stopping procedures to reduce overfitting. The testing subset provided an independent evaluation of predictive capability. Once satisfactory performance had been achieved, the final neural network models were re-trained using the full dataset in order to maximize predictive robustness and produce the final reported outputs.

### Performance Metrics

For crash frequency models, predictive accuracy was assessed using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). For crash severity models, classification performance was evaluated using Area Under the Receiver Operating Characteristic Curve (AUC), balanced accuracy, precision, recall, F1-score, confusion matrices, and Brier score for probability calibration.

This two-stage strategy combines the advantages of independent model validation with the efficiency of final full-data estimation, ensuring that the reported results are both methodologically credible and based on the maximum available evidence.

# 4. CHAPTER 4

# REAL CASE EXPERIMENT

## 4.1.Case study description

The empirical implementation of the proposed behavior-driven and risk-based safety framework is conducted on the road network of the Province of Brescia (Lombardy Region, Northern Italy. The selection of this study area was primarily motivated by the availability of large-scale telematics data integrated with crash, infrastructure, environmental, and traffic exposure information, enabling the operational development of a comprehensive multi-source analytical framework for proactive road safety assessment. Building on this context, the Province of Brescia is one of the most populous provinces in Italy, comprising 205 municipalities and approximately 1.26 million inhabitants, which contributes to the development of a large and highly heterogeneous road network and positions the area among the most complex transport environments in northern Italy. The province is characterized by high mobility demand generated by industrial production, residential settlements, commercial activities, and tourism, resulting in substantial daily traffic flows across multiple functional road classes. The road system includes urban corridors, peri-urban connectors, rural and mountainous routes, industrial access roads, and high-speed arterial infrastructure. Such spatial and functional diversity provides a robust and representative context for analyzing the interaction between infrastructure characteristics, driving behaviour, exposure conditions, and crash outcomes within a unified and integrated modelling framework.

The road network structure further justifies the analytical relevance of the case study. The province comprises more than 340 km of highways, over 1,700 km of provincial roads managed by the Province of Brescia Road Department, approximately 170 km of state roads managed by **Azienda Nazionale Autonoma delle Strade** (ANAS), and more than 200 km of local roads managed by municipalities (Winlaw et al., 2019). This extensive and hierarchically organized system supports both local circulation and long-distance passenger and freight transport, making it particularly suitable for network-wide safety screening and risk estimation. From a geographical perspective, the Province of Brescia is characterized by marked morphological variability. The southern sector is predominantly flat, belonging to the Pianura Padana, while the northern areas progressively transition into rolling terrain and Alpine mountainous environments. This spatial heterogeneity results in substantial variation in roadway geometry, speed regimes, alignment characteristics, and environmental

conditions, all of which influence driving behaviour and crash risk mechanisms. Consequently, the province offers a natural laboratory for investigating behaviour–infrastructure interactions under diverse operational contexts.

With respect to road safety performance, 2023 was selected as the reference year for analysis to ensure temporal consistency with the telematics dataset used in this research. According to the 2023 crash records, approximately 3,720 road crashes were reported within the provincial territory, resulting in more than 5,060 injured persons and around 59 fatalities over the year (Polis Lombardia, 2023). These figures quantify the annual safety burden across the network and provide the empirical baseline for subsequent modelling activities. The spatial distribution of crashes indicates that the majority (approximately 75%) occurred within urban areas, reflecting higher traffic density and increased interaction among vehicles, pedestrians, and cyclists. However, the highest proportion of fatal crashes was recorded on the primary extra-urban network, where higher operating speeds and longer uninterrupted travel segments contribute to increased crash severity. This dual pattern in high frequency in urban contexts and high severity in non-urban corridors reinforces the need for an integrated modelling approach that simultaneously accounts for crash occurrence, exposure, and consequences.

From a methodological standpoint, the Province of Brescia offers a rare opportunity to operate Naturalistic Driving Data (NDD) at the network scale. The availability of continuous telematics observations enables the direct extraction of behavioural indicators and Safety-Critical Events, facilitating the transition from traditional retrospective crash-based evaluation to predictive, behaviour-informed risk modelling. The temporal alignment between telematics observations and crash data ensures analytical coherence, allowing behavioural precursors to be systematically associated with observed crash outcomes. Overall, the combination of network complexity, morphological diversity, high traffic demand, documented safety burden, and extensive telematics data availability makes the Province of Brescia a robust and representative case study for implementing and validating the proposed behaviour-driven, exposure-adjusted, and risk-based road safety assessment framework. The following sections present the empirical results obtained from the application of crash frequency models, crash severity models, and the formulation of the Integrated Road Safety Risk Index within the Brescia road network, demonstrating how behavioural indicators and exposure-adjusted modelling contribute to proactive safety assessment at the network level.

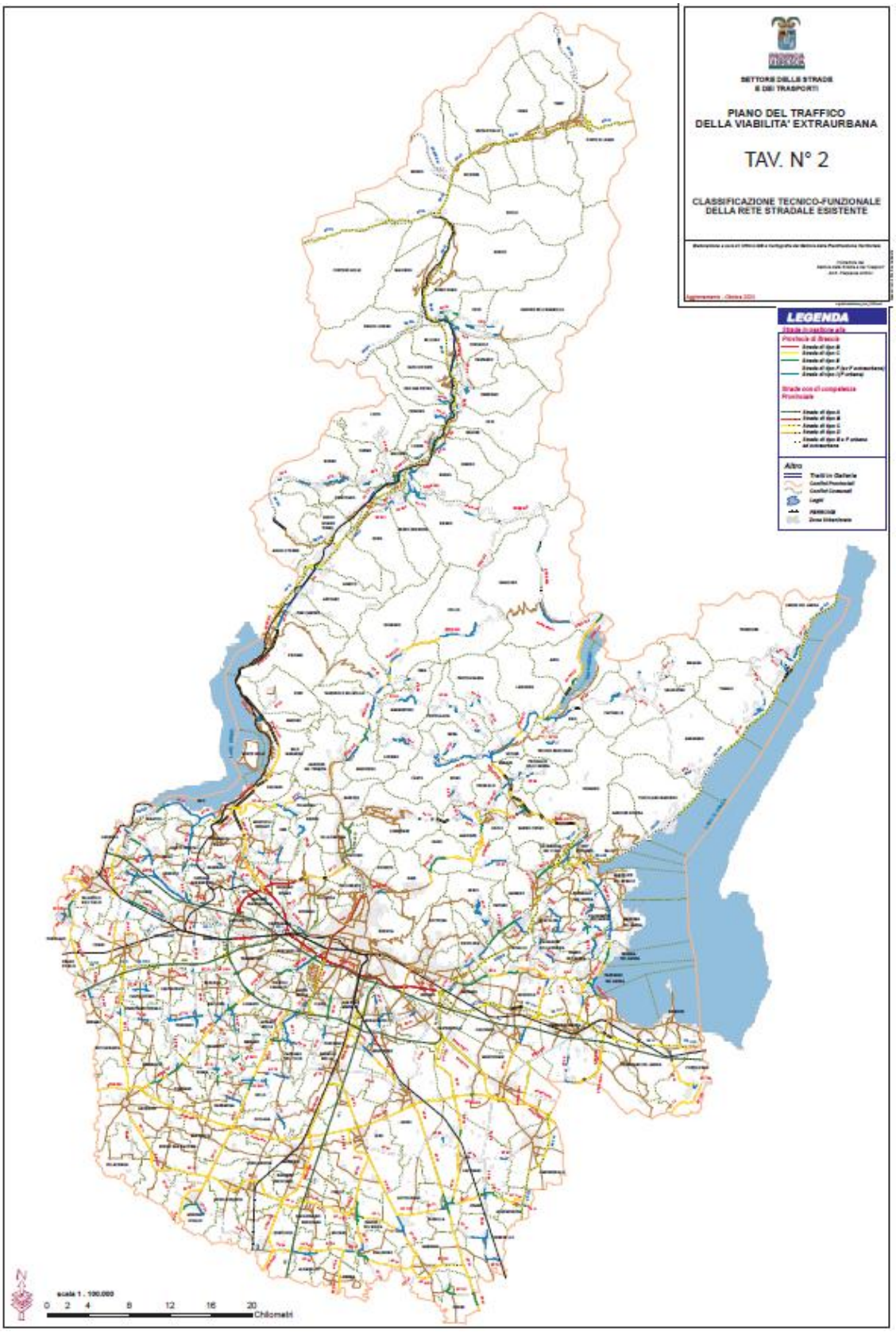



*Figure 0-1: Main Road network of the Province of Brescia considered in the study*

## 4.2.Phase 1 Results – Telematics Data Processing and Safety-Critical Event Extraction

### *4.2.1. Data collection and preparation*

**Telematics Data Preprocessing and Derived Kinematic Indicators**

The primary dataset used in this study consists of telematics-based Naturalistic Driving Data provided by the company named as FasterNet s.r.l., collected from approximately 1,230 instrumented vehicles operating across the road network of the Province of Brescia over 24 hours span. Specifically, the dataset comprises timestamped GPS trajectories and in-vehicle sensor measurements that capture real-world driving behaviour under varying operational and environmental conditions. In total, the telematics database includes approximately 730,000 geo-referenced observations and therefore constitutes the core behavioural data source supporting the analytical framework developed in this research. More precisely, the telematics dataset includes raw trajectory variables and motion-related attributes derived directly from in-vehicle sensing systems. The input data structure is organized around spatial–temporal and motion variables recorded at the vehicle level, including *air_point_id*, *car_id*, *recorded_at*, *latitude*, *longitude*, *heading*, and acceleration components along the three axes (*acc_x*, *acc_y*, *acc_z*). Collectively, these variables enable the reconstruction of vehicle trajectories and the characterization of operational driving conditions with high temporal resolution.

Beyond its technical structure, the dataset provides a substantial empirical basis for behavioural road safety analysis because it covers a heterogeneous transport environment comprising urban, suburban, rural, and mountainous road contexts. Although the sample does not represent the full driving population in a statistically random sense, it captures a broad range of daily mobility patterns, roadway classes, and traffic conditions. Consequently, the dataset is considered sufficiently representative for behavioural risk modelling and exploratory network-level safety screening, while acknowledging that direct extrapolation to all drivers should be interpreted with caution.

From the perspective of fleet composition, the available telematics records primarily refer to passenger cars, reflecting the nature and accessibility of the available data source. Other vehicle categories such as heavy goods vehicles, buses, motorcycles, and commercial fleets are therefore not explicitly represented in the present empirical application. Nevertheless, the principal objective of this research is strongly methodological, namely the development of an integrated framework for behavioural event extraction, crash modelling, and network-level risk assessment. Accordingly, the proposed methodology remains transferable and can be readily updated, recalibrated, and extended whenever

broader multi-vehicle datasets become available in future applications.Potential sources of telematics bias were also considered. Instrumented vehicles may overrepresent insured drivers, technology-adopting users, or frequent commuters, while infrequent drivers and some demographic groups may be underrepresented. Likewise, road sections with higher traffic demand are naturally observed more frequently than lower-volume local roads. To mitigate these limitations, the study integrates telematics-derived indicators with roadway, crash, and exposure variables rather than relying solely on raw event counts.

To address the representativeness, potential bias, and data reliability issues discussed above, the telematics trajectories were further examined before the structured preprocessing stage in order to assess temporal sampling behaviour and overall data quality. Following the above considerations, the telematics trajectories were examined to further assess temporal sampling behaviour and inclusive data quality. The sampling interval between consecutive observations ($\Delta t$) exhibited substantial variability, ranging from 0.0 seconds to a maximum of 300.0 seconds, with a mean of 4.47 seconds, a median of 3.0 seconds, and a standard deviation of 11.16 seconds. Zero-second intervals indicated duplicated timestamps originating from logging or data integration processes, whereas very large intervals reflected signal loss, communication delays, or genuine inactivity when vehicles were stationary. This irregular temporal structure is typical of large-scale naturalistic driving datasets and therefore required a principled segmentation strategy to preserve behavioural continuity. Based on the empirical $\Delta t$ distribution, the maximum observed interval of 300 seconds was adopted as the temporal gap threshold ($\tau = 300$ s) for defining trip boundaries. This threshold prevents derivative-based kinematic variables from being computed across prolonged gaps while avoiding unnecessary fragmentation due to short operational pauses such as congestion or brief stops.

A rigorous preprocessing procedure was then implemented to address missing, duplicated, inconsistent, and noisy records. Invalid coordinates, impossible speeds, duplicate timestamps, incomplete trajectories, and records with major sensor anomalies were removed during cleaning. From the initial 730,151 observations, 455,787 valid geo-referenced records were retained after preprocessing (a reduction of approximately 38%), and vehicles with insufficient coverage were excluded to guarantee stable behavioural assessment across trips, yielding a final analytical sample of 790 instrumented vehicles.

On this cleaned base, trajectories were organized into temporally continuous trips using the gap threshold of $\tau = 300$ s, producing 4,007 coherent driving episodes. Each episode was then resampled to a fixed frequency of $f = 2$ Hz ($\Delta t = 0.5$ s) to stabilize derivative estimates and enable consistent computation of speed, acceleration, jerk, heading change, and curvature. Subsequently, a sliding-window scheme with window length $\omega = 2.5$ s and step size $\delta = 1.25$ s was applied, providing a balanced 50% overlap and generating 1,013,838 points distributed across 67,565 analytical windows. Within each window, indicators such as max |a_y|, max |a_x|, Δspeed, Δθ, and maximum curvature were computed.

The resulting distributions confirm the physical plausibility and behavioural interpretability of the standardized kinematic signals. Across all windows, mean speed was 10.85 m/s (range 0.0–43.1 m/s). Longitudinal acceleration exhibited a mean of –0.36 m/s² (range –17.88 to 20.59 m/s²; SD = 1.81 m/s²), lateral acceleration displayed a mean of –0.12 m/s² (range –17.99 to 20.63 m/s²; SD = 1.66 m/s²), and vertical acceleration had a mean of 0.65 m/s² (range –21.45 to 16.45 m/s²; SD = 2.02 m/s²). Resultant acceleration magnitudes averaged 2.15 m/s² (range 0.012–25.18 m/s²), with a median of 1.45 m/s². Mean resultant jerk was 0.42 m/s³ (range 0.0–24.44 m/s³; SD = 0.62 m/s³), while curvature values averaged 0.0149 $m^{-1}$ with a range of 0.0–0.111 $m^{-1}$. Graphical summaries are provided through a circular heading-rose plot (Fig.4.2( a) and a curvature cumulative distribution function (Fig. 4.2( b)), illustrating the directional structure and geometric variability of the trajectories.

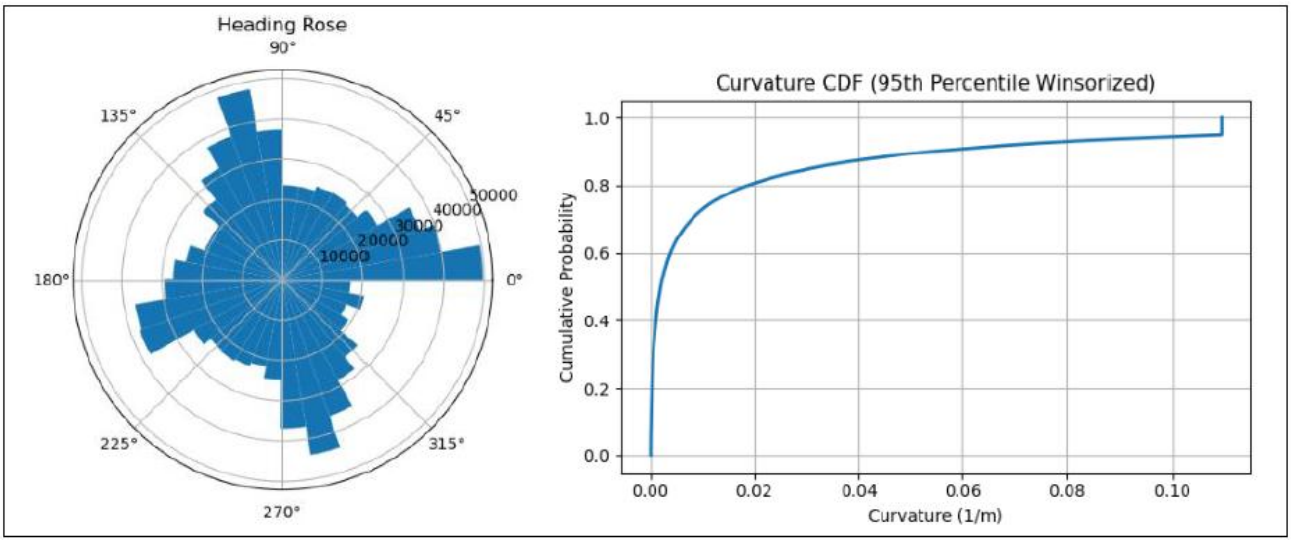



*Figure 0-2: Direction and Curvature Characteristics: (a) Heading Rose and (b) Curvature CDF.*

The heading-rose plot exhibits sharp, well-defined directional lobes with no random scatter, indicating a clean and stable heading signal after resampling and windowing. Likewise, the curvature CDF shows a smooth, monotonic distribution with realistic geometric ranges, confirming that outliers and noise were effectively removed, leaving a physically coherent representation of roadway curvature. Before proceeding to Safety-Critical Event (SCE) extraction, the telematics dataset is complemented with three external information layers official crash records, roadway infrastructure inventory, and environmental/contextual conditions. These datasets operationalize the outcome, structural, and exogenous components of the safety system. Crash data from ISTAT[1] provide the empirical injury-severity and crash-frequency outcomes used for model calibration and validation. Road inventory sources include (*official regional Geographic Information System (GIS) datasets*[2]*, OpenStreetMap (OSM), and structured spatial queries extracted through Overpass-based mapping tools*[3]) supply geometric, functional, and regulatory characteristics at the segment level, while environmental datasets contribute weather, lighting, land-use, and built-environment attributes. We integrate infrastructure and environmental data during with SCE extraction to ensure context-aware event detection and defer the incorporation of crash data to subsequent step where to serve as outcome measures for crash frequency and severity modelling.

*Table 4.1 Summary Characteristics of the Telematics Dataset*

| Category | Indicator | Value |
|---|---|---|
| **Data Source** | Provider | FasterNet s.r.l. |
| | Study area | Brescia |
| | Monitoring period | 24 hours |
| **Raw Dataset** | Initial vehicles | 1,230 |
| | Raw observations | 730,151 |
| | Main vehicle type | Passenger cars |
| **Quality Control** | Valid observations retained | 455,787 |
| | Records removed | 38.0% |
| | Final analytical vehicles | 790 |
| **Temporal Structure** | Mean sampling interval (Δt) | 4.47 s |
| | Median sampling interval | 3.0 s |
| | Standard deviation of Δt | 11.16 s |
| | Maximum observed gap | 300 s |
| | Trip gap threshold (τ) | 300 s |
| **Trajectory Processing** | Trips identified | 4,007 |
| | Resampling frequency | 2 Hz |
| | Time step after resampling | 0.5 s |

[1] https://www.istat.it/en/press-release/road-accidents-2023/
[2] *https://www.geoportale.regione.lombardia.it/* ***and*** *https://sit.provincia.brescia.it/*
[3] https://overpass-turbo.eu/

| | | |
|---|---|---|
| **Windowing Framework** | Window length (ω) | 2.5 s |
| | Step size (δ) | 1.25 s |
| | Window overlap | 50% |
| | Total windows | 67,565 |
| | Total window points | 1,013,838 |
| **Kinematic Summary** | Mean speed | 10.85 m/s |
| | Mean resultant acceleration | 2.15 $m/s^2$ |
| | Mean resultant jerk | 0.42 $m/s^3$ |
| | Mean curvature | 0.0149 $m^{-1}$ |

### *4.2.2. **Safety-Critical Event Extraction***

Following the application of the multi-threshold detection framework described in Phase 1 (Step 2), Safety-Critical Events (SCEs) were successfully extracted from the cleaned and map-matched telematics dataset. Events were flagged only when multiple kinematic and contextual conditions were simultaneously satisfied within a predefined temporal window, thereby ensuring that detected cases represent coherent high-risk maneuvers rather than isolated sensor noise. The analysis focused on six SCE categories (Sudden Braking, Lane Departure, Sharp Turn, Rapid Acceleration, Speeding, and Vertical Disturbance) which collectively capture the principal longitudinal, lateral, vertical, and regulatory dimensions of risky driving behavior. These categories were selected based on their theoretical relevance to crash risk and their measurability using the available telematics variables (acceleration components, jerk, speed, heading variation, curvature, and speed limits). As a result, the extracted SCEs provide a multidimensional behavioral representation of driver–vehicle–infrastructure interaction within the study network. The adopted threshold values were defined to ensure physical plausibility and robustness to measure noise. Specifically, the limits were calibrated to identify dynamic conditions that exceed normal driving variability while remaining consistent with realistic vehicle performance boundaries. The threshold configuration used for event detection is summarized in Table 4.2.

*Table 0-2: Threshold Values for Extraction of SCEs.*

| Event Type | Feature | Threshold | Meaning / Reasoning |
|---|---|---|---|
| **1. Sudden Braking (SBE)** | acc_x_mps2 | < –3.5 $m/s^2$ | Strong deceleration (hard brake) |
| | jerk_x_mps3 | > 5 $m/s^3$ | Sudden change in deceleration |
| | speed_mps | > 8 m/s | Filter out slow-speed noise |
| **2. Lane Departure** | acc_y_mps2 | > 1.2 $m/s^2$ | Hard lateral movement indicates drift or swerve |

| | | | |
|---|---|---|---|
| | jerk_y_mps3 | > 1.5 m/s³ | Sudden steering input |
| | heading_change_deg | > 10° | Significant directional shift |
| | speed_mps | > 5 m/s | Avoid false positives from parking turns |
| **3. Sharp Turn** | acc_y_mps2 | ≥ ±2.5 m/s² | Strong lateral G-force from turning |
| | curvature | ≥ 0.015 | Tight road curve (radius ≈ 66 m) |
| | heading_change_deg | ≥ 20–30° in 2.5s window | Rapid directional change |
| | speed_mps | ≥ 8 m/s | Eliminates low-speed maneuvering |
| **4. Rapid Acceleration** | acc_x_mps2 | ≥ 3.0 m/s² | Strong forward acceleration |
| | jerk_x_mps3 | ≥ 5.0 m/s³ | Sudden thrust |
| | speed_mps | ≤ 25 m/s | Common urban acceleration |
| **5. Speeding** | speed_mps | > (maxspeed × 0.278) | Violation of road speed limit |
| **6. Vertical Disturbance (VDE)** | acc_z_mps2 | > 11.8 or < 7.8 m/s² | Upward or downward jolts beyond ±2 m/s² from gravity baseline |

The completion of Step 2 marks the successful transformation of cleaned and map-matched telematics trajectories into a structured behavioral safety dataset. Through the implementation of the multi-threshold detection framework, a total of **15,311 geo-referenced Safety-Critical Events (SCEs)** were systematically identified across six behavioral categories. The extraction results clearly demonstrate that the behavioral risk profile of the observed vehicle sample is overwhelmingly dominated by **speeding events (77.6%)**, indicating that regulatory non-compliance constitutes the principal operational risk layer within the study network. This marked predominance suggests that excessive speed is not an isolated phenomenon but rather a systematic behavioral pattern across the Province of Brescia road network. From a safety systems perspective, such concentration of events within a single behavioral category implies that speed-related exposure forms the structural backbone of the network's dynamic risk environment.

Secondary behavioral layers are represented by **Vertical Disturbance Events (13.9%)** and **Sudden Braking Events (4.8%)**, each capturing distinct dimensions of driving risk. Vertical Disturbance Events reflect pronounced fluctuations in vertical acceleration and likely emerge from infrastructure-related factors such as pavement irregularities, surface discontinuities, abrupt grade transitions, or

localized geometric inconsistencies. Unlike purely driver-driven behaviors, VDEs embody the interaction between vehicle dynamics and roadway conditions, thereby serving as indicators of potential infrastructure-induced stress. Their relatively high share highlights that risk exposure is not exclusively behavioral but partially embedded in the physical road environment.

Sudden Braking Events, although substantially less frequent than speeding, represent abrupt deceleration patterns that often occur in response to perceived hazards, unexpected traffic interactions, or constrained following distances. Within a proactive safety framework, these events function as dynamic near-miss proxies, signaling localized moments of elevated conflict probability. Their presence provides valuable insight into micro-level traffic instability and reactive driving behavior that may precede crash occurrences.

Lateral instability behaviors, namely **lane departures (2.1%)** and **sharp turns (1.4%)**—appear comparatively infrequent in absolute terms. However, their lower frequency should not be interpreted as lower safety relevance. These maneuvers are commonly associated with high-severity crash mechanisms, including run-off-road events and loss-of-control collisions. Empirical crash literature consistently demonstrates that even rare instability behaviors can contribute disproportionately to severe injury outcomes. Consequently, their inclusion in the behavioral risk structure remains essential for comprehensive crash severity modeling.

This stratified behavioral composition reinforces the need for integrated safety strategies that simultaneously address speed management, infrastructure quality, traffic conflict mitigation, and control stability. The findings therefore not only quantify behavioral exposure but also provide direction for targeted, evidence-based intervention domains within a proactive and risk-based road safety management framework.

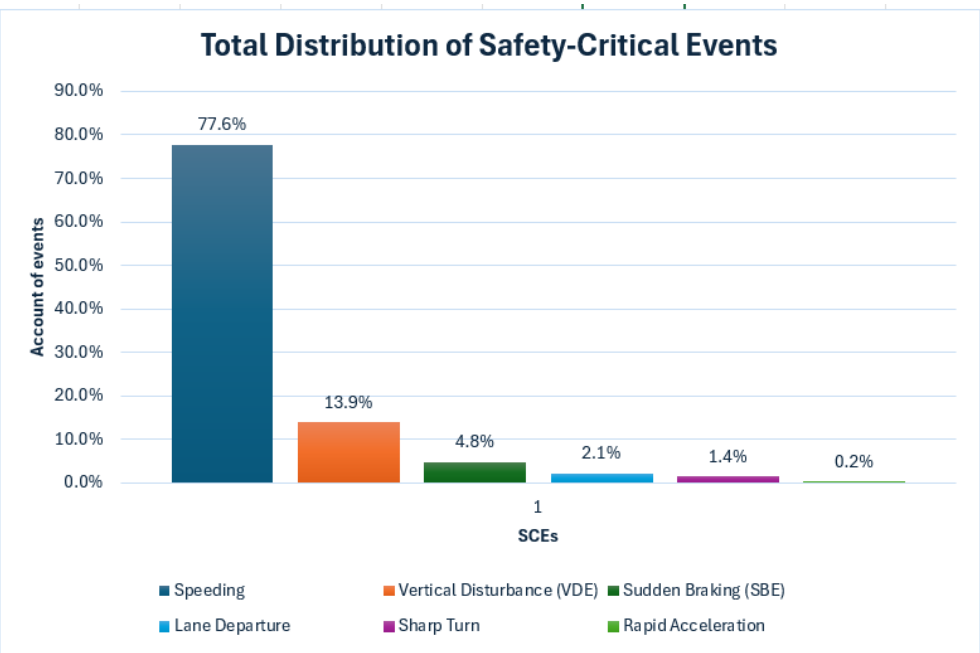



*Figure 0-3: Total Distribution of Safety-Critical Events*

*Figure 4.3 illustrates the overall distribution of detected Safety-Critical Events across the six behavioral categories. The bar chart clearly highlights the dominance of speeding events relative to other dynamic risk indicators.*

Importantly, each detected SCE was time-stamped, spatially geo-referenced, and linked to roadway infrastructure attributes through the map-matching procedure, meaning that the output of Step 2 is not merely a collection of isolated event flags but a fully integrated behavioral-event dataset that preserves the spatial, temporal, and contextual dimensions of risk exposure. This structured SCE dataset forms the core behavioral predictor layer for the subsequent modelling framework and effectively operationalizes the transition from raw telematics signals to quantifiable behavioral risk indicators, providing the empirical foundation for proactive, data-driven crash-risk estimation in the following analytical phases. For methodological transparency, detailed time-series visualizations illustrating the threshold-based detection process for each event category are provided in Appendix A.

### *4.2.3. Event–Road–Crash Integration and Multi-Layer Data Assembly*

Step 3 produces a unified, segment- and trip-aligned analytical dataset in which Safety-Critical Events (SCEs), roadway infrastructure characteristics, and historical crash outcomes are harmonized within a common spatial framework. At this stage, each SCE is fully contextualized by the geometric,

functional, and regulatory attributes of the roadway segment on which it occurred, enabling behavioral instability patterns to be interpreted relative to curvature, lane configuration, intersection density, posted speed limits, lighting conditions, and other infrastructural features. The integration of crash records enriches the dataset with segment-level safety outcomes, yielding **3,884 segments**, including **2,515 segments with recorded crashes**, **3,719 total crashes**, and **5,059 total injuries** mapped to the network. The resulting integrated base contains **6 behavioural event categories** and **52 segment-level predictors** prepared for downstream modelling as depicted in Appendix B.

Because map-matching and infrastructure linkage were completed during SCE extraction, **the segment-level dataset is inherently created at this step**: every behavioural event and every crash record is already aligned to a unique segment. Segment-level aggregation therefore consists of rolling up behavioural indicators across all trips that traverse the same segment computing event counts, relative frequencies, and intensity statistics while retaining the segment's geometric, functional, and contextual attributes. This produces a macro-scale dataset that captures spatial patterns of behavioral instability and their correspondence with roadway characteristics and crash outcomes.

In parallel, **a distinct trip-level aggregation** is performed to characterize behavioural dynamics within individual driving episodes. Events are grouped by trip identifier to compute within-trip totals, proportions, and dispersion measures for each SCE category. Unlike the segment level, infrastructure attributes are not re-assigned at this scale; instead, the trip dataset preserves temporal continuity and driver-specific behavioural variability. Safety-outcome proxies are attached after behavioural aggregation, maintaining the chronological relationship between driving patterns and potential safety consequences. The final trip-level dataset comprises **2157 trips**, with **6 behavioural summaries** and used for **trip -level predictors**, enabling micro-scale behavioural analysis independent of fixed spatial boundaries.

Taken together, the segment-level (macro) and trip-level (micro) datasets provide complementary analytical perspectives: the segment level reveals where behavioural instability clusters across the roadway network, while the trip level captures how behavioural patterns evolve within individual journeys. This dual-scale structure enhances the explanatory depth and system coherence of the modelling framework, ensuring that behavioral indicators, infrastructure conditions, crash frequency, and crash severity are consistently represented across both spatial and temporal dimensions

### *4.2.4. Results of Traffic Exposure Estimation and Scaling*

The operationalization of the exposure framework for the Province of Brescia required an empirical scaling procedure because continuous counter-based AADT data were not available across the entire network. Using provincial demographic and motorization statistics together with the telematics dataset, the effective sampling expansion factor was estimated. Brescia has a resident population of **1,260,955** individuals and a motorization rate of **0.64 vehicles per person**, yielding an estimated **807,011 passenger vehicles** in circulation. The telematics dataset included **750 instrumented vehicles**, which corresponds to an empirical expansion factor of:

$$w = \frac{807{,}011}{750} \approx 1{,}076. \tag{4.1}$$

This indicates that each instrumented vehicle in the dataset represents approximately **1,076 vehicles** in the actual circulating fleet. This expansion factor was directly applied to the observed telematics-based segment-hour traffic counts, enabling the reconstruction of population-level exposure estimates across the Brescia road network.

**Segment-Level Allocation**

Segment-level exposure was allocated using the reconstructed AADT estimates derived from the scaled telematics observations. For each road segment $s$and hour $h$, the number of unique instrumented vehicles was extracted and then expanded to population level using the previously obtained factor $w$. The resulting hourly traffic volumes, $T_{s,h} = w \cdot T_{s,h}^{obs}$, were aggregated across all observed days to produce a telematics-based AADT proxy for each segment:

$$AADT_s = \frac{1}{D} \sum_{h \in D} T_{s,h} \,. \tag{4.2}$$

The resulting AADT values exhibit substantial spatial heterogeneity across the Brescia network. Among the 2,515 analyzed segments, exposure ranges from **1,096.63 to 22,026.47 vehicles/day**, with a mean of **1,751.31** and a median of **1,096.63** vehicles/day. The distribution is strongly right-skewed (skewness = 6.99) and heavy-tailed (kurtosis = 66.24), indicating that a small subset of segments carries disproportionately high traffic volumes. Percentile analysis further highlights this hierarchy:

while **75% of segments** record approximately **1,096.63 vehicles/day**, the upper percentiles increase sharply, reaching **2,980.96** (P90), **2,980.96** (P95), and **8,103.08** vehicles/day (P99). These patterns reflect the concentration of mobility demand along major corridors and arterial connectors. Because of this pronounced skewness, a logarithmic transformation was applied to stabilize variance for modeling purposes. The log-transformed exposure values are substantially more symmetric, ranging from **7 to 10**, with reduced dispersion (standard deviation = 0.52), lower skewness (1.98), and lower kurtosis (4.20). This transformation preserves meaningful exposure differences while preventing high-volume segments from exerting disproportionate influence in the Negative Binomial crash-frequency model. Overall, the segment-level results confirm both the structural concentration of traffic across the Brescia network and the importance of exposure normalization for reliable rate-based modeling.

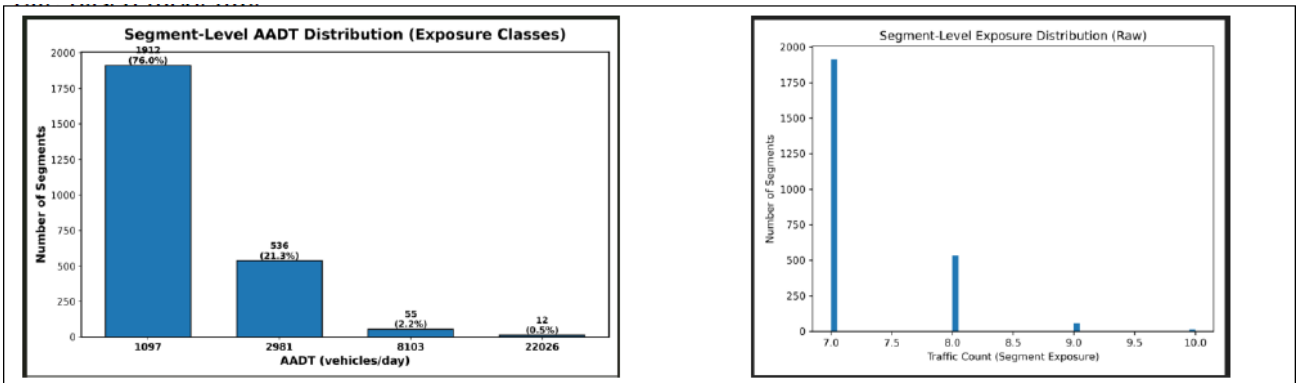



*Figure 0-5: Segment-Level Traffic Exposure: (a) Reconstructed AADT Classes and (b) Raw Log Transformed Exposure*

**Trip-Level Allocation**

Trip-level exposure was allocated using the normalized segment-level values obtained from the reconstructed AADT estimates. For each segment $s$, exposure was first expressed as $E_s = \log(AADT_s)$. Trip-level exposure for each trip $r$was then computed by aggregating these normalized segment-level exposures across all segments traversed within the trip:

$$E_r = \frac{1}{n_r}\sum_{s \in r} E_s\,, \tag{4.3}$$

where $n_r$denotes the number of segments included in trip $r$. The resulting trip-level exposure variable therefore represents the average normalized traffic environment encountered over the duration of each driving episode. Across the 2,157 analyzed trips, two complementary indicators were used to summarize exposure: **the maximum relative exposure** experienced during the trip and **the mean**

**relative exposure** across all traversed segments. The maximum-based indicator (\textit{trip_level_max_relative}) ranges from **1 to 30 vehicles/hour**, with a mean of **11.85** and a median of **10 vehicles/hour**. Dispersion is relatively high (SD = 7.74), and percentile analysis shows that **25% of trips** encounter peak exposure levels above **18 vehicles/hour**, while the upper tail reaches **30 vehicles/hour** (P99). In contrast, the mean-based indicator (\textit{trip_level_mean_relative}) ranges from **1 to 25 vehicles/hour**, with a mean of **8.44** and a median of **8 vehicles/hour**. Variability is lower (SD = 4.99), and percentile growth is smoother, with the P99 equal to **21 vehicles/hour**.

These two indicators capture different dimensions of trip-level traffic intensity. The **mean exposure** reflects the sustained traffic density experienced throughout the trip, representing cumulative interaction opportunity. The **maximum exposure** identifies peak congestion episodes encountered during the driving episode, capturing short-duration but potentially critical high-density conditions. As expected, maximum exposure consistently exceeds mean exposure, but both distributions remain bounded and moderately symmetric, indicating that trip-level exposure is substantially more compressed than segment-level AADT and does not exhibit heavy-tail behavior. In the modeling framework, **mean exposure** was retained for crash frequency analysis because it reflects cumulative interaction potential over the entire trip and exhibits smoother statistical properties. Conversely, **maximum exposure** was used for severity modeling, as severity outcomes are more sensitive to extreme operational conditions than to typical background traffic levels. Together, these two indicators provide a comprehensive representation of traffic intensity at the trip scale.

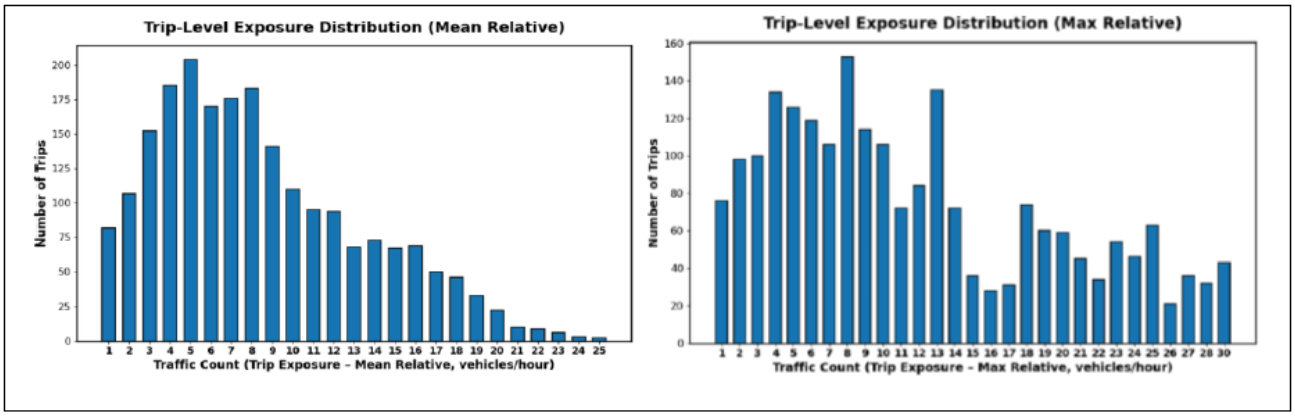



*Figure 0-7: Trip Level Exposure Distributions: (a) Mean Relative Exposure and (b) Maximum Relative Exposure*

Trip level exposure is expressed in vehicles per hour because it reflects the actual traffic intensity encountered during each trip. Unlike AADT, which represents daily demand, the trip level measure is derived directly from segment hour telematics counts, preserving real temporal variation and ensuring comparability across trips of different durations.

### 4.3.Phase 2 Results: Parallel Modelling of Crash Frequency and Crash Severity

Following the construction of the exposure-adjusted, multi-scale analytical datasets in Phase 1, Phase 2 operationalizes the behavioural–infrastructural safety framework through the parallel modelling of crash frequency and crash severity. Models were developed at both the segment and trip levels, enabling safety outcomes to be examined from a network-level performance perspective as well as from a behavioural-variability perspective. Crash frequency was estimated using Negative Binomial (NB) regression and Feedforward Neural Networks (FNN), whereas crash severity was modelled using Logistic Regression and FNN classifiers. This dual-model strategy provides a structured comparison between interpretable econometric relationships and the nonlinear predictive patterns captured by machine-learning models.

To examine potential redundancy among explanatory variables, correlation and multicollinearity analyses were conducted for both the segment-level and trip-level datasets. The correlation heatmaps indicate generally weak pairwise relationships among predictors, with only a few localized clusters and no clear block-structured dependencies. This suggests that the behavioural, infrastructural, exposure-related, and contextual variables capture largely distinct dimensions of the driving environment and driver behaviour, thereby providing complementary explanatory information for the modelling framework.

Multicollinearity was further assessed using the **Variance Inflation Factor (VIF)**. At the **segment level**, the results confirm an overall low level of multicollinearity across the design matrix. Only one predictor, *Vertical_and_horizontal_Road_Signs*, exceeded conventional thresholds (VIF = 12.80), indicating redundancy with other infrastructural indicators, while *Non-Intersection/Straight* showed moderate collinearity (VIF = 6.61). All remaining variables exhibited VIF values ≤ 4.35 (e.g., *Wet_Road_Surface* = 4.35, *Rain_Weather_Conditions* = 3.94, *traffic count* = 3.09), with most predictors clustered within the range of 1–3, which is generally considered indicative of acceptable multicollinearity conditions. Accordingly, *Vertical_and_horizontal_Road_Signs* was removed to reduce redundancy and improve model stability, and the refined design matrix was retained for the

**Negative Binomial (NB)** and **Feed-Forward Neural Network (FNN)** modelling stages. In addition, a **rare-level screening conducted for the segment-level dataset** identified several binarized indicators with very low prevalence (e.g., *Intersection/Crossroads, One_way_carriageway, Rain_Weather_Conditions, Wet_Road_Surface, lane_2,* and *sidewalk_both*), which were excluded during preprocessing to avoid unstable parameter estimates.

At the **trip level**, correlation analysis among the behavioural safety indicators (*Speeding, VDE_flag, Sharp_Turn, SBE, lane_departure,* and *RAE*) similarly reveals weak pairwise relationships, indicating that these variables represent distinct aspects of risky driving behaviour. A moderate association is observed between *VDE_flag* and *Sharp_Turn*, while *lane_departure* shows a modest relationship with *RAE*, whereas *Speeding* and *SBE* remain largely independent of the other behavioural indicators. VIF diagnostics further confirm low multicollinearity across the trip-level predictors. Overall, the correlation and VIF diagnostics indicate limited redundancy among predictors, supporting their inclusion in the modelling framework. **Detailed correlation heatmaps and VIF results are provided in Appendix C (Figures C1–C2 and Tables C1 -C2).**

### *4.3.1. Crash Frequency Modelling Results*

Crash frequency was modelled as a stochastic count process in which exposure was incorporated primarily as an offset term. Descriptive analysis indicated substantial overdispersion in the crash data, with the variance exceeding the mean (**Mean = 1.48, Variance ≈ 4.99**), supporting the use of a **Negative Binomial (NB)** specification rather than a Poisson alternative. The estimated dispersion parameter further confirmed the presence of unobserved heterogeneity across modelling units, indicating that crash occurrence varies across segments and trips beyond what can be explained by observed explanatory variables alone.

#### *4.3.1.1.Negative Binomial Crash Frequency Models*

At the **segment level**, the NB model demonstrated satisfactory goodness-of-fit (**AIC = 8158.2, Log-likelihood = −4049.1**). Behavioural indicators derived from Safety-Critical Events (SCEs) were significant predictors of crash occurrence. In particular, **Speeding** was associated with a notable increase in crash frequency (**IRR = 1.12**), indicating that a one-unit increase in speeding behaviour corresponds to approximately a **12% increase in expected crash frequency**, holding exposure constant. Similarly, **vehicle dynamic instability (VDE_flag)** was positively associated with crash

occurrence (**IRR = 1.13**), suggesting that loss of vehicle stability contributes to higher crash likelihood.

Infrastructure characteristics also played an important role in shaping crash occurrence. Road segments with **three traffic lanes (IRR = 1.31)** exhibited higher crash frequency, while **signalized intersections (IRR = 1.33)** were also associated with increased crash occurrence. A particularly strong effect was observed for **lighted tunnels**, where crash frequency was more than **three times higher (IRR = 3.51)** compared with the baseline condition. Because exposure was incorporated as an offset term, the estimated coefficients represent **intrinsic safety performance differences across segments**, rather than variations driven solely by traffic intensity. In terms of predictive accuracy, the segment-level NB model achieved an **RMSE of 1.07** and **MAE of 0.48**, indicating reasonable predictive performance.

At the **trip level**, the NB model captured behavioural variability across individual driving episodes (**AIC = 10453.0, Log-likelihood = −5219.5, Pseudo $R^2$ = 0.162**). Behavioural indicators were again the dominant predictors of crash occurrence. **Speeding** remained the most influential factor (**IRR = 1.13, p < 0.001**), indicating that higher speeding intensity during trips increases expected crash frequency by approximately **12–13%**. **Sudden Braking Events (SBE)** were also positively associated with crash occurrence (**IRR = 1.07, p = 0.043**), while **vehicle dynamic instability (VDE_flag)** showed a smaller but statistically significant effect (**IRR = 1.03, p = 0.016**). Compared with the segment-level models, behavioural variables played a more prominent role at the trajectory level, highlighting the importance of driver behaviour in shaping crash risk during individual driving episodes. The complete NB estimation results for both modelling scales are reported in **Table 4.3**.

*Table 0-3: Negative Binomial Regression Results for Segment and Trip Levels*

| Negative Binomial Regression Results for Segment level | | | | |
|---|---|---|---|---|
| **Category** | **Independent variable** | **Description** | **Estimate Coeff.** | **P-value** |
| **Model constant** | Constant | Natural log of constant (i.e., α) | -1.782 | **** |
| **Road Infrastructure** | Intersection type | With traffic lights / traffic warden | 0.283 | ** |

| | | | | |
|---|---|---|---|---|
| | Site type | Lighted tunnel | 1.257 | **** |
| **Road Geometry** | N. of lanes | Three lanes | 0.273 | ** |
| **Driving Behaviour** | Speeding | Driving above speed limit | 0.118 | **** |
| | Vertical Dynamic Event (VDE) | Vertical instability event | 0.123 | **** |

| **Negative Binomial Regression Results for Trip Levels** | | | | |
|---|---|---|---|---|
| **Category** | **Independent variable** | **Description** | **Estimate Coeff.** | **P-value** |
| **Model constant** | Constant | Natural log of constant (i.e., α) | -1.041 | **** |
| **Driving Behaviour** | SBE | Sudden Braking Events | 0.065 | * |
| | Speeding | Driving above speed limit | 0.119 | **** |
| | VDE_flag | Vertical Dynamic Event | 0.028 | ** |

**P-value legend: ****: <0.001; ***: <0.005; **<0.010; *<0.1; -:>0.1**

*4.3.1.2.Feed-Forward Neural Network (FNN) Crash Frequency Models*

To capture potential **nonlinear relationships** between behavioural indicators, infrastructure characteristics, and crash occurrence, a **Feed-Forward Neural Network (FNN)** model was implemented alongside the econometric specification. At the **segment level**, the FNN model demonstrated improved predictive performance compared with the NB model, achieving an **RMSE of 0.82** and **MAE of 0.37**, representing approximately a **23% improvement in predictive accuracy** relative to the NB model. Interpretation of the neural network using **SHAP (SHapley Additive Explanations)** revealed that **Speeding** was the most influential predictor of crash frequency, followed by infrastructure-related variables such as **roundabout intersections, road lighting conditions, and signalized intersections**, as well as behavioural indicators including **vehicle dynamic instability (VDE_flag)** and **Sudden Braking Events (SBE)**. These findings suggest that crash occurrence is

shaped by the combined effects of **driver behaviour and roadway environment characteristics** (**Figure 4.6**).

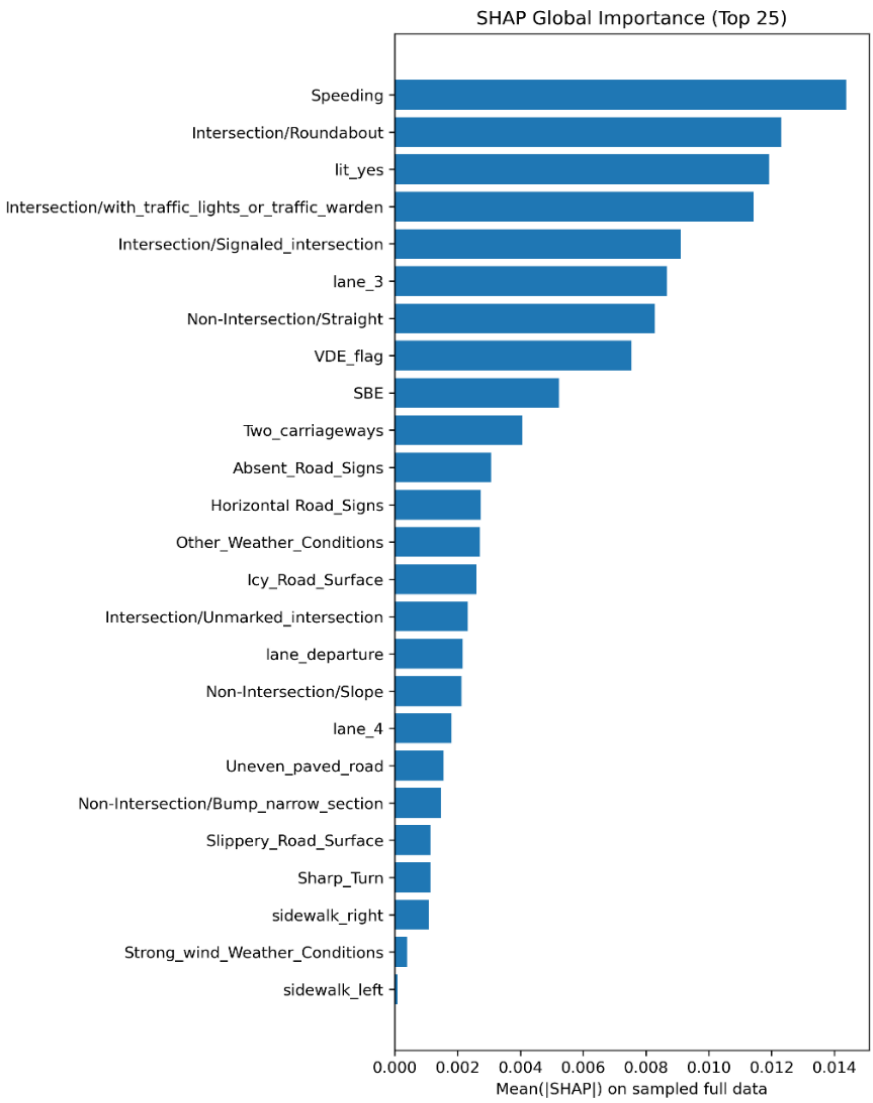



*Figure 0-8: SHAP Global Feature Importance for the FNN Model (Top 25 Variables).*

To further examine the contribution of risky driving behaviours, **permutation importance** was applied to behavioural predictors. This analysis confirmed that **Speeding** and **vehicle dynamic instability (VDE_flag)** were the most influential behavioural variables, followed by **Sudden Braking Events (SBE)**, while behaviours such as **sharp turning, lane departure, and rapid acceleration events (RAE)** contributed comparatively less to overall model performance (**Figure 4.7**).

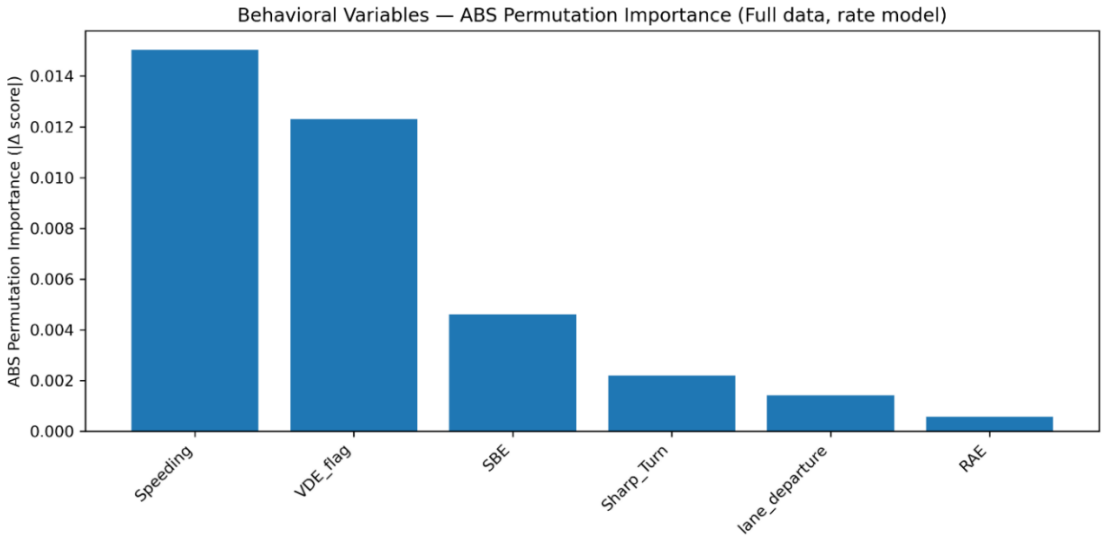



*Figure 0-9: Behavioral Variable Importance Based on Permutation Analysis (FNN Model).*

At the trip level, the FNN model focused primarily on behavioural indicators and was estimated using crash rate (crash_count / mean_relative_traffic) to account for exposure differences across trips. The model achieved an RMSE of 0.461 and MAE of 0.311, indicating acceptable predictive accuracy despite the highly stochastic nature of crash occurrences. Both SHAP and permutation importance analyses consistently identified Speeding as the dominant behavioural predictor, with additional contributions from vehicle dynamic instability (VDE_flag) and Sudden Braking Events (SBE) as seen in fig (4.8). These results indicate that driver behavioural instability becomes the primary driver of crash risk at the trajectory level, where infrastructure characteristics play a relatively smaller role.

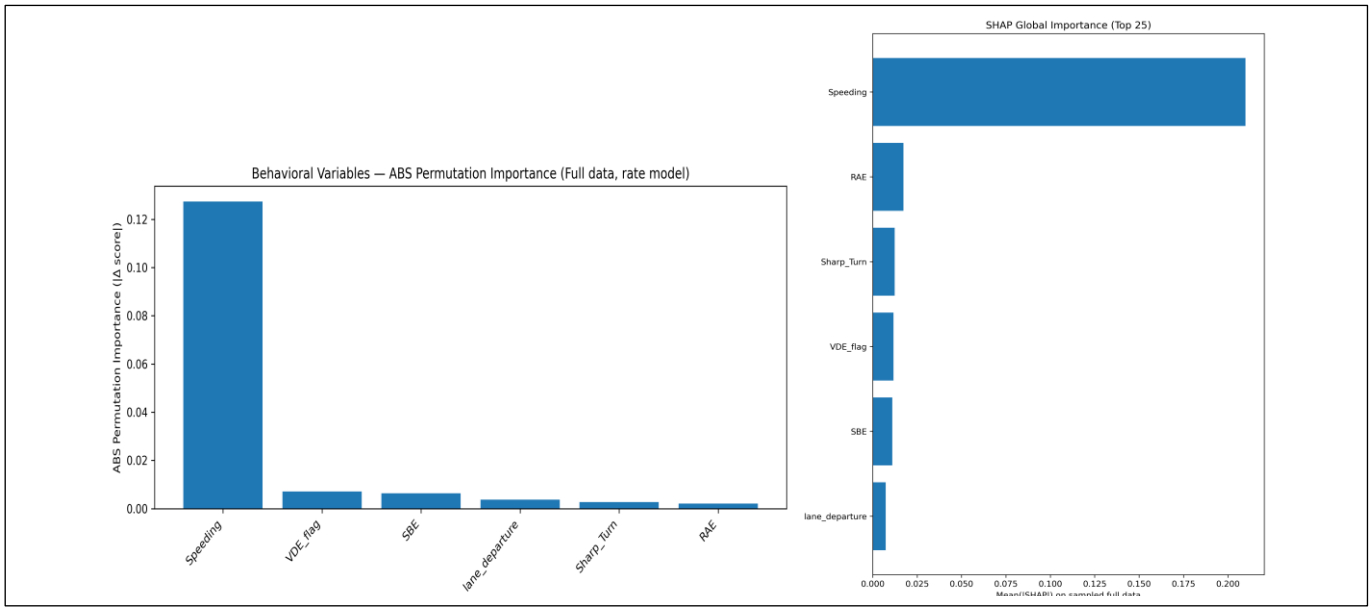



*Figure 0-10: Behavioral Risk Factor Importance: (a) Permutation Importance and (b) SHAP Global Importance.*

A comparative overview of the predictive performance of the econometric and machine learning crash frequency models is presented in **Table 4.4**, highlighting the improved predictive capability of the FNN models, particularly at the segment level where nonlinear interactions between behavioural indicators and roadway characteristics are more pronounced. The complete estimation results for the econometric crash frequency models, including coefficient estimates, standard errors, and statistical significance levels, are provided in **Appendix D (Tables D1–D2)**.

*Table 0-4: Comparison of Crash Frequency Model Performance*

| Analysis Level | Model Type | Prediction Power | RMSE | MAE | Key Observation |
|---|---|---|---|---|---|
| **Segment Level** | Negative Binomial (Econometric) | Pseudo $R^2$ = **0.065** | 1.07 | 0.48 | Provides interpretable rate-based relationships and accounts for exposure effects but captures nonlinear interactions only to a limited extent. |
| **Segment Level** | FNN (Artificial Neural Network) | $R^2$ = **0.678** | 0.82 | 0.37 | Achieves substantially higher predictive accuracy by modelling complex nonlinear interactions between behavioral and roadway factors. |
| **Trip Level** | Negative Binomial (Econometric) | Pseudo $R^2$ = **0.162** | 15.05 | 3.40 | Maintains statistical interpretability, but predictive performance is limited under the high variability of trip-level crash occurrences. |
| **Trip Level** | FNN (Artificial Neural Network) | $R^2$ = **0.314** | 3.82 | 2.52 | Captures nonlinear behavioral variability in trip-level crash risk and improves prediction error compared with the econometric model, though interpretability is reduced. |

### *4.3.2. Crash Severity Modelling Results*

Crash severity was modelled using a binary classification framework in which the outcome variable indicates whether a crash resulted in a severe injury outcome. The severity indicator was constructed using injury and fatality information. Specifically, a crash was classified as severe when it involved at least one fatality or when the number of injured people exceeded a predefined injury threshold ($\tau$). The threshold value was determined through empirical evaluation of alternative candidate values using model performance metrics including the Area Under the Receiver Operating Characteristic Curve (AUC), balanced accuracy, and calibration measures. Based on this validation procedure, $\tau = 3$

was selected as the most appropriate value because it provided a stable class distribution while maximizing predictive performance.

Descriptive analysis indicates moderate class imbalance across modelling units. At the segment level, among 2,515 crash observations, 517 crashes (20.56%) were classified as severe while 1,998 crashes (79.44%) were non-severe. Fatal crashes represented a relatively small share of the dataset (46 cases, 1.83%). At the trip level, the dataset contained 2,157 observations, of which 1,424 crashes (66.02%) were classified as severe and 733 crashes (33.98%) were non-severe. Fatal crashes accounted for 198 observations (9.18%).

Crash severity was first estimated using a Binary Logistic Regression framework, allowing interpretation of the effects of behavioural, infrastructure, and crash configuration variables on the probability of severe outcomes. At the segment level, the Binary Logit model demonstrated good discriminative capability with an AUC of approximately 0.806 and a balanced accuracy of about 0.673, indicating a strong ability to distinguish between severe and non-severe crash outcomes despite class imbalance. The Brier score of approximately 0.178 further suggests acceptable probabilistic calibration of predicted severity probabilities.

Behavioural indicators derived from Safety-Critical Events (SCEs) emerged as important determinants of injury severity. In particular, speeding exhibited a strong positive association with severe crash outcomes (OR = 1.74), indicating that crashes involving speeding behaviour were associated with approximately 74% higher odds of severe injuries compared with crashes without speeding. Similarly, vehicle dynamic instability measured through Vertical Dynamic Events (VDE) significantly increased severity risk (OR = 1.39), suggesting that loss of vehicle control substantially increases the likelihood of severe injury outcomes.

Crash configuration characteristics also played a significant role. Frontal-lateral collisions (OR = 1.71) were associated with higher severity risk, while side collisions (OR = 1.42) and head-on collisions (OR = 1.36) were also linked with elevated injury severity, reflecting the high impact forces and direct vehicle-to-vehicle contact involved in these configurations. In addition, crashes involving multiple vehicles showed a strong association with severity (OR = 1.65), indicating that interactions among multiple vehicles increase injury severity risk.

Road infrastructure characteristics further influenced severity patterns. Crashes occurring on three-lane segments (OR = 1.39) and two-lane roads (OR = 1.37) were associated with increased severity compared with the reference category, while four-lane segments also showed elevated severity risk (OR = 1.19). Similarly, straight non-intersection segments (OR = 1.48) and signalized intersections (OR = 1.33) exhibited higher severity probabilities, potentially reflecting higher travel speeds and complex traffic interactions in these environments. Additional roadway characteristics such as vertical and horizontal road signage (OR = 1.35), sloped segments (OR = 1.16), and lighted tunnels (OR = 1.19) were also associated with increased severity risk. Environmental conditions such as snowy road surfaces (OR = 1.13) also contributed to higher injury severity. In contrast, certain demographic characteristics, such as Driver Age Group B (OR = 0.77), were associated with lower severity risk compared with the reference age group.

At the trip level, the Binary Logit model captured the influence of behavioural dynamics during individual driving episodes. The model achieved an AUC of approximately 0.804, indicating strong discriminative capability, while the Brier score of about 0.166 indicates acceptable probabilistic calibration. Using a classification threshold of 0.63, the model achieved an overall accuracy of 0.714 and a balanced accuracy of approximately 0.737, indicating reasonable predictive performance across both classes.

Behavioural indicators derived from naturalistic driving data were the dominant predictors of severity at the trip level. Speeding exhibited by far the strongest association with severe crash outcomes (OR = 24.53), indicating that trips involving speeding behaviour were associated with more than twenty-four times higher odds of severe crash outcomes. Lane departure events also significantly increased severity risk (OR = 1.23), indicating that lateral vehicle instability during driving episodes contributes to more severe crash outcomes. In addition, Vertical Dynamic Events (OR = 1.31) and sharp turning manoeuvres (OR = 1.16) exhibited marginal statistical significance, suggesting that vehicle instability and aggressive steering behaviour may also contribute to increased crash severity, although their effects are weaker than those of speeding and lane departures.

Overall, the Binary Logit models indicate that driver behaviour, crash configuration, and roadway characteristics jointly influence injury severity outcomes. Across both modelling scales, behavioural instability  particularly excessive speed  consistently emerges as the most influential determinant of severe crash outcomes. These findings highlight the importance of incorporating behavioural

indicators derived from Naturalistic Driving Data into crash severity modelling frameworks, as summarized in Table 4.5.

*Table 0-5: Binary Logistic Regression Results for Crash Severity Estimation at Segment and Trip Levels*

*Binary Logistic Regression Results for Crash Severity Estimation at Segment Levels*

| Category | Independent variable | Description | Estimate Coeff. | Odds Ratio | P-value |
|---|---|---|---|---|---|
| **Driving Behaviour** | Speeding | Driving above speed limit | 0.556 | 1.743 | **** |
| | Vertical Dynamic Event (VDE) | Vertical instability event | 0.332 | 1.394 | **** |
| **Road Geometry** | Lane = 3 | Three-lane roadway | 0.333 | 1.394 | **** |
| | Lane = 2 | Two-lane roadway | 0.311 | 1.365 | **** |
| | Lane = 4 | Four-lane roadway | 0.172 | 1.188 | **** |
| **Road Infrastructure** | Non-Intersection – Straight | Straight road segment | 0.392 | 1.480 | *** |
| | Intersection – Traffic lights / traffic warden | Controlled intersection | 0.286 | 1.331 | **** |
| | Intersection – Unmarked | Uncontrolled intersection | 0.225 | 1.252 | * |
| | Non-Intersection – Lighted tunnel | Tunnel segment | 0.173 | 1.189 | *** |
| | Non-Intersection – Slope | Sloped road segment | 0.147 | 1.158 | ** |
| | Vertical and horizontal road signs | Combined road signage | 0.298 | 1.348 | *** |
| **Crash Characteristics** | Frontal-lateral collision | Collision configuration | 0.534 | 1.706 | * |
| | Side collision | Collision configuration | 0.350 | 1.419 | * |
| | Head-on collision | Collision configuration | 0.304 | 1.355 | * |
| | Number of other vehicles involved | Multi-vehicle involvement | 0.498 | 1.645 | **** |
| **Driver / Vehicle Factors** | Driver Age B | Driver age category B | -0.259 | 0.772 | *** |
| **Road Surface / Environment** | Snowy road surface | Snow-covered surface condition | 0.126 | 1.134 | ** |
| **Vehicle Type** | Scheduled/unscheduled buses in suburban areas A | Vehicle class indicator | 0.118 | 1.125 | * |

***Binary Logistic Regression Results for Crash Severity Estimation at Trip Levels***

| **Category** | Independent variable | Description | Estimate Coeff. | Odds Ratio | P-value |
|---|---|---|---|---|---|
| **Driving Behaviour** | Speeding | Driving above speed limit | 3.200 | 24.533 | **** |
| | Lane departure | Lateral instability event | 0.206 | 1.228 | * |
| | Vertical Dynamic Event (VDE) | Vertical instability event | 0.272 | 1.313 | † |
| | Sharp Turn | Aggressive turning maneuver | 0.152 | 1.164 | † |

**P-value legend: ****: <0.001; ***: <0.005; **<0.010; *<0.1; -:>0.1**

### *4.3.2.1.Feed-Forward Neural Network (FNN) Crash Severty Models*

To capture potential **nonlinear interactions** between behavioural indicators, roadway characteristics, and crash configuration variables, a **Feed-Forward Neural Network (FNN) classifier** was

implemented alongside the econometric severity models. At the segment level, the FNN classifier demonstrated strong predictive performance. Using a classification threshold of **0.17**, the model achieved an **overall accuracy of 87.0%** and a **balanced accuracy of 0.889**, indicating strong classification capability across both severity classes. Probability-based evaluation metrics further confirm the model's predictive performance, with an **AUC of 0.952** and a **Brier score of 0.0649**, indicating excellent discrimination and good probabilistic calibration.

At the segment level, the FNN severity model captured nonlinear relationships among roadway, traffic, and behavioural factors. SHAP interpretation indicated that **vehicle type, lane configuration, and Speeding** were among the most influential predictors of severe crash probability, followed by infrastructure characteristics such as **straight non-intersection segments and signalized intersections**, and behavioural indicators including **vehicle dynamic instability (VDE_flag)** and **Sudden Braking Events (SBE)**. These results suggest that severe crash outcomes are influenced by the combined effects of driver behaviour, traffic interaction, and roadway environment (Figure 4.9).

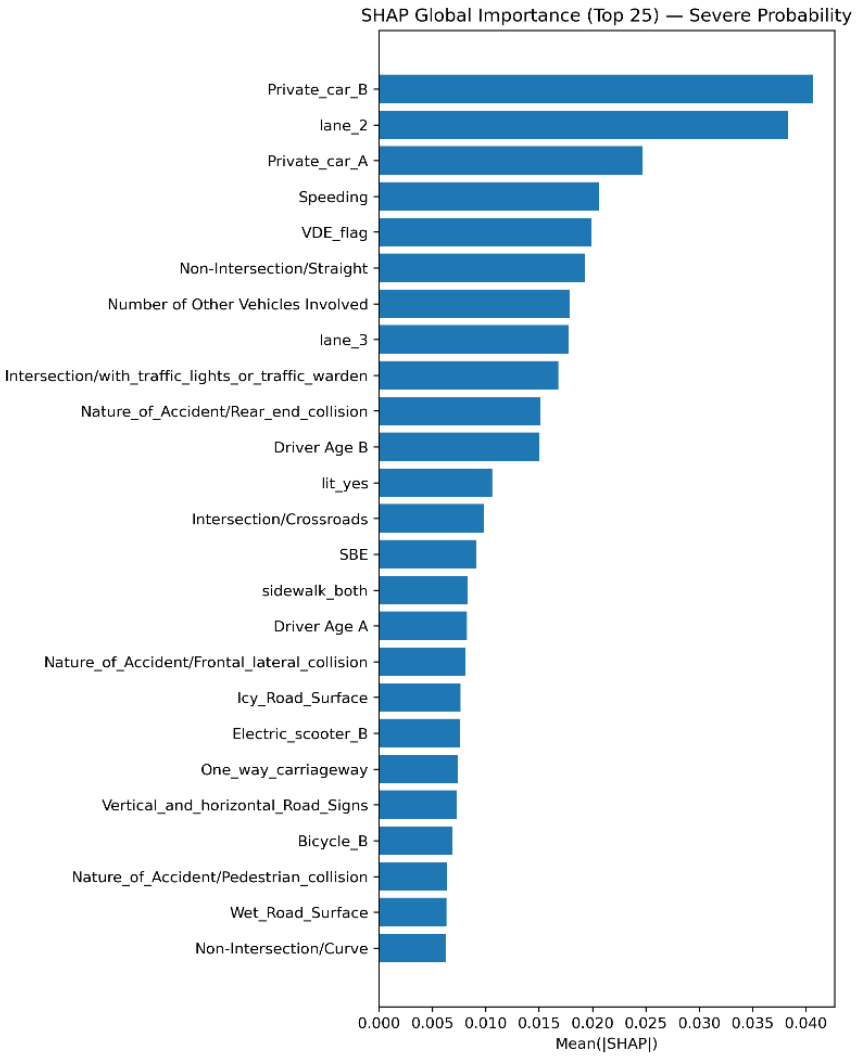



*Figure 0-11: SHAP Global Feature Importance for the Segment-Level FNN Severity Model (Top 25 Variables).*

Permutation importance analysis further confirmed that Speeding produced the largest decrease in model performance when permuted, followed by traffic interaction variables, vehicle type, lane configuration, and behavioural indicators such as VDE_flag and Sharp Turn, reinforcing the role of risky driving behaviour in severe crash outcomes (Figure 4.10).

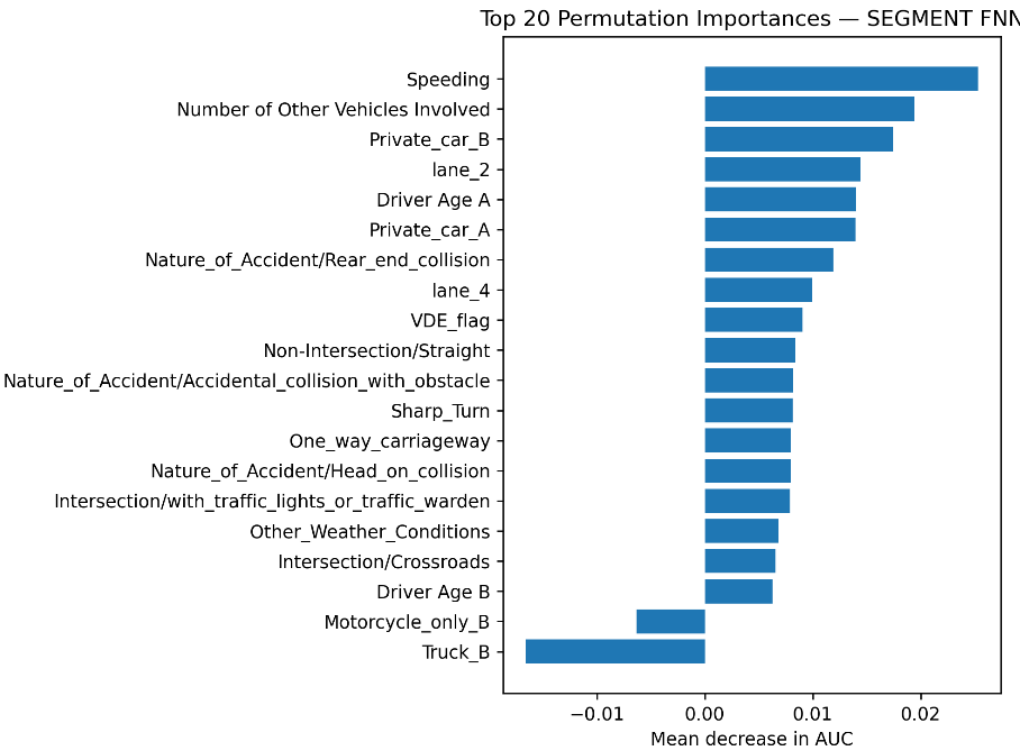



*Figure 0-12: Permutation Importance for the Segment-Level FNN Severity Model (Top 20 Variables).*

At the trip level, the FNN severity model was estimated using six behavioural indicators derived from vehicle motion and driving dynamics: Speeding, Sudden Braking Events (SBE), Rapid Acceleration Events (RAE), Lane Departure, Sharp Turns, and Vehicle Dynamic Instability (VDE_flag). Using a classification threshold of 0.63, the model achieved an overall accuracy of 0.7149. The model correctly identified 593 non-severe crashes and 949 severe crashes, while 140 non-severe crashes were incorrectly classified as severe and 475 severe crashes were misclassified as non-severe. Class-specific performance metrics indicate that the model achieved high precision for severe crash predictions (0.871), although recall for severe cases was moderate (0.666). Probability-based evaluation metrics further confirm the model's performance, with an AUC of 0.809 and a Brier score of 0.164, indicating good discrimination and acceptable probabilistic calibration.

At the trip level, the FNN severity model focused primarily on behavioural indicators derived from naturalistic driving data. Interpretation using permutation importance and SHAP analyses consistently shows that Speeding is the dominant behavioural predictor of severe crash probability. Permutation importance results indicate that Speeding produces the largest decrease in model performance when

permuted, followed by vehicle dynamic instability (VDE_flag), lane departure, Sudden Braking Events (SBE), sharp turns, and Rapid Acceleration Events (RAE).

SHAP global importance analysis confirms this pattern, identifying Speeding as the most influential factor substantially increasing the predicted probability of severe crash outcomes, while other behavioural indicators including lane departure, VDE_flag, sharp turns, SBE, and RAE provide smaller but still meaningful contributions to the model predictions. These results indicate that driver behavioural instability is the primary driver of crash severity at the trajectory level, where roadway infrastructure characteristics play a relatively limited role.

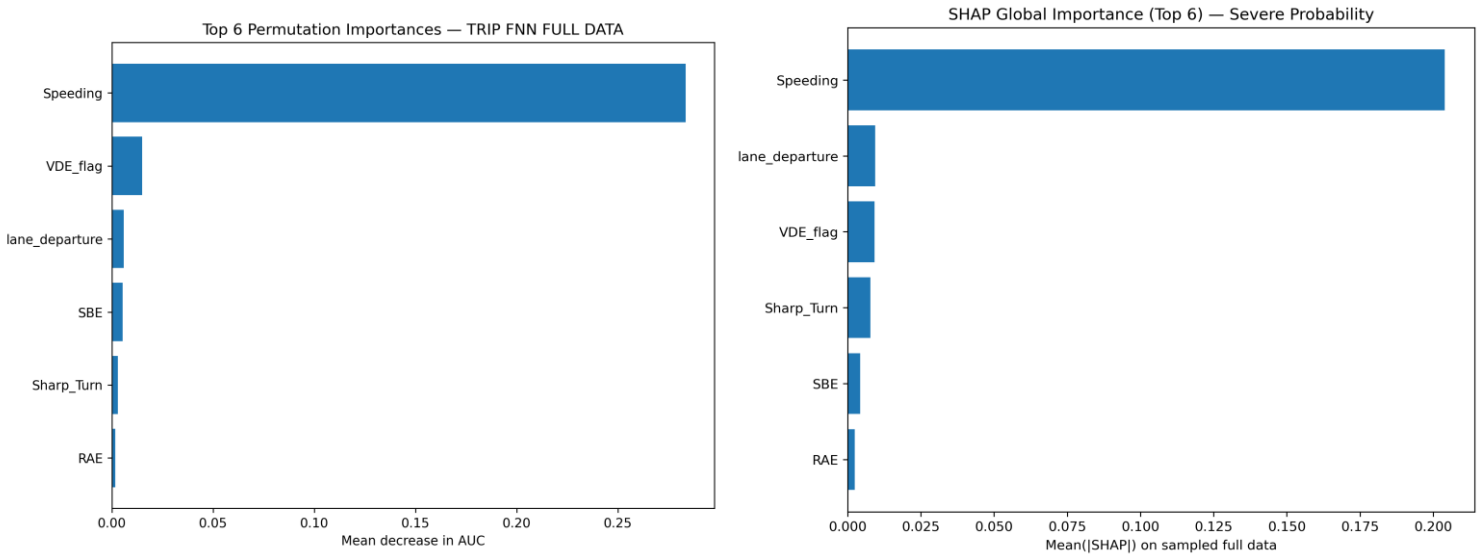



*Figure 0-14: Behavioral Risk Factor Importance: (a) Permutation Importance and (b) SHAP Global Importance.*

A comparative overview of the predictive performance of the econometric and machine learning severity models is presented in **Table 4.6**, which highlights the superior predictive capability of the FNN models, particularly at the segment level where nonlinear interactions between behavioural indicators and roadway characteristics are more pronounced. The complete estimation results for the econometric models, including coefficient estimates, standard errors, and statistical significance levels, are provided in **Appendix E (Tables E1–E2)**.

*Table 0-6: Comparison of Crash Severity Model Performance*

| Analysis Level | Model Type | AUC | Balanced Accuracy | Accuracy | Brier Score | Key Observation |
|---|---|---|---|---|---|---|
| **Segment Level** | Binary Logit (Econometric) | 0.806 | 0.673 | 0.539 | 0.178 | Captures statistically significant effects of speeding, vehicle dynamic instability, and |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | | roadway configuration, but shows limited classification performance due to nonlinear relationships and class imbalance. |
| **Segment Level** | FNN (Artificial Neural Network) | 0.952 | 0.889 | 0.870 | 0.065 | Substantially improves severity prediction by capturing nonlinear interactions among behavioral indicators, crash configurations, and roadway attributes. |
| **Trip Level** | Binary Logit (Econometric) | 0.804 | 0.724 | 0.654 | 0.180 | Highlights the strong influence of behavioral indicators particularly speeding on severity outcomes but assumes linear relationships between predictors and crash severity probability. |
| **Trip Level** | FNN (Artificial Neural Network) | 0.809 | 0.738* | 0.715 | 0.164 | Better captures complex behavioral dynamics across trips, improving classification of severe crashes compared with the logit model. |

### *4.3.3. Comparative Interpretation*

The comparison between the statistical and ANN-based models demonstrates important methodological differences in the representation of crash risk mechanisms. Statistical approaches, including Negative Binomial and Binary Logistic Regression models, provided strong interpretability by explicitly quantifying the contribution of explanatory variables through coefficients, incidence rate ratios, and odds ratios. These models were particularly effective for identifying statistically significant relationships among behavioural indicators, roadway characteristics, traffic exposure, and crash outcomes. Variables such as speeding, traffic exposure, roadway geometry, and vehicle instability emerged as highly important because they consistently showed strong statistical association with both crash occurrence and crash severity across multiple model specifications. In contrast, variables such as sharp turning and rapid acceleration exhibited comparatively lower importance, likely because these behaviours occurred less frequently within the dataset and contributed more locally to specific driving situations rather than to generalized network-wide crash patterns. Similarly, environmental variables demonstrated moderate importance, suggesting that while weather and roadway surface conditions influence crash outcomes, their effects are often secondary relative to dominant behavioural and infrastructure-related risk factors.

As summarized in Table 4.7,, segment-level statistical models demonstrated that traffic exposure, roadway geometry, infrastructure characteristics, and behavioural instability jointly influenced crash occurrence and severity patterns across the roadway network. The strong significance of variables such as traffic count, roadway configuration, signalized intersections, and Vertical Dynamic Events (VDEs) indicates that segment-level crash risk is strongly shaped by spatial roadway and operational conditions in addition to driver behaviour. However, despite the explanatory strength of the statistical

models, their capability to represent complex nonlinear interactions among variables remained relatively limited. Importantly, variables identified as having comparatively lower importance should not be interpreted as being unimportant within the crash risk estimation framework. Rather, their lower relative contribution indicates that their effects are either more context-specific, indirectly represented through stronger correlated behavioural patterns, or less dominant under the operational conditions observed in the study area. Therefore, the relative ranking of variables reflects comparative explanatory contribution within the modelling framework rather than absolute safety relevance.

*Table 4-7: Comparative Importance of Explanatory Variables in Segment-Level Statistical and ANN-Based Models*

| Explanatory Variable | Statistical Model (NB / Logit) | ANN Model (FNN) | Comparative Interpretation |
|---|---|---|---|
| **Speeding Events** | Very High | Very High | Consistently emerged as the strongest predictor across both modelling approaches for crash frequency and severity. |
| **Traffic Exposure Variables** | Very High | High | Explicitly incorporated through offsets in statistical models, while ANN models captured exposure effects through nonlinear relationships. |
| **Vertical Dynamic Events (VDE)** | High | High | Significant in both approaches, reflecting vehicle instability and infrastructure interaction effects. |
| **Road Geometry Variables** | High | Moderate–High | Statistical models emphasized interpretable geometric effects, whereas ANN models captured their nonlinear interactions with behaviour. |
| **Traffic Control / Infrastructure Variables** | Moderate–High | High | ANN models captured combined interaction effects among intersections, tunnels, lighting, and behavioural indicators more effectively. |
| **Lane Departure Events** | Moderate | High | Demonstrated stronger nonlinear contribution within ANN severity models compared with statistical approaches. |
| **Sudden Braking Events (SBE)** | Moderate | Moderate–High | ANN models captured behavioural interaction effects associated with abrupt driving responses more effectively. |
| **Environmental Variables** | Moderate | Moderate | Showed similar moderate contribution in both modelling frameworks. |
| **Sharp Turning Events** | Low | Low–Moderate | Limited influence overall, although ANN models captured localized nonlinear effects more effectively. |
| **Rapid Acceleration Events (RAE)** | Low | Low–Moderate | Demonstrated comparatively limited influence across both modelling approaches. |
| **Behavioural Interaction Effects** | Limited | Very High | ANN models substantially improved representation of nonlinear behavioural interactions not explicitly captured in statistical models. |
| **Model Interpretability** | Very High | Moderate–Low | Statistical models provided clearer explanatory relationships through coefficients and odds ratios. |
| **Nonlinear Relationship Modelling** | Limited | Very High | ANN models demonstrated superior capability in capturing complex nonlinear crash risk patterns. |

In contrast, the Feed-Forward Neural Network (FNN) models demonstrated substantially improved predictive capability, particularly at the segment level where interactions between behavioural indicators, infrastructure conditions, and traffic exposure were more complex. The ANN models more effectively captured nonlinear behavioural interaction effects, especially those associated with speeding, lane departure behaviour, vehicle instability, and sudden braking patterns. The higher importance of these variables within the ANN framework suggests that crash risk is strongly influenced not only by the independent contribution of individual behaviours but also by their combined and nonlinear interactions under varying roadway and traffic conditions. At the trip level, where the modelling framework relied primarily on behavioural indicators extracted from Naturalistic Driving Data, ANN models showed stronger capability in representing real-time behavioural variability and trajectory instability dynamics. In particular, speeding consistently remained the dominant predictor because of its direct influence on reaction time reduction, stopping distance, impact force, and overall driving instability. Lane departure and VDE indicators also demonstrated increased importance within ANN models due to their ability to capture loss-of-control dynamics and vehicle instability patterns that may not be fully represented through linear statistical relationships. Conversely, variables such as Rapid Acceleration Events (RAE) and Sharp Turning Events showed comparatively lower contribution because their effects were either less frequent or partially embedded within broader behavioural instability patterns already captured by more dominant variables.

As presented in Table 4.8, the trip-level models differed substantially from the segment-level framework because they focused almost exclusively on behavioural indicators rather than fixed roadway or infrastructure characteristics. Consequently, trip-level crash risk estimation was more strongly associated with dynamic trajectory instability, behavioural interaction effects, and real-time driving variability. This distinction indicates that segment-level models are more effective for identifying spatial and infrastructure-related crash risk patterns across the roadway network, whereas trip-level models provide deeper insight into driver behavioural dynamics and instantaneous crash risk mechanisms during individual driving episodes. Furthermore, the comparison between the statistical and ANN-based frameworks highlights the fundamental distinction between linear and nonlinear representations of crash risk mechanisms. Statistical models primarily estimate predefined and relatively linear relationships between explanatory variables and safety outcomes, making them highly

effective for interpretability, inference, and policy-oriented analysis. In contrast, ANN-based approaches demonstrated stronger capability in capturing nonlinear dependencies, hidden behavioural interactions, and complex multivariate relationships that may not be fully observable through traditional statistical formulations. This is particularly important because real-world crash mechanisms are rarely governed by isolated variables acting independently; instead, crash risk often emerges from the simultaneous interaction of driver behaviour, infrastructure conditions, traffic exposure, and environmental context. Consequently, certain variables that appear only moderately significant within linear statistical models may become substantially more influential within ANN frameworks once nonlinear interaction effects are considered. This finding reinforces the complementary nature of the two modelling paradigms, where statistical models provide transparent explanatory insight while ANN approaches improve the representation of complex behavioural and operational crash risk dynamics. Nevertheless, despite their improved predictive performance, ANN models exhibited lower interpretability compared with the statistical models. Overall, the findings suggest that the statistical models provide stronger explanatory insight for safety interpretation and policy analysis, whereas ANN models offer superior capability for capturing complex nonlinear crash risk patterns and improving predictive accuracy.

*Table 4-8: Comparative Importance of Behavioural Variables in Trip-Level Statistical and ANN-Based Models*

| Behavioural Variable | Statistical Model (NB / Logit) | ANN Model (FNN) | Comparative Interpretation |
|---|---|---|---|
| **Speeding Events** | Very High | Very High | Dominant behavioural predictor across all trip-level models, particularly influential in severity estimation. |
| **Sudden Braking Events (SBE)** | Moderate | Moderate–High | ANN models captured abrupt behavioural instability and interaction effects more effectively. |
| **Lane Departure Events** | Moderate | High | Demonstrated stronger contribution in ANN severity models due to nonlinear instability representation. |
| **Vertical Dynamic Events (VDE)** | Moderate | High | Vehicle instability effects became more pronounced within ANN-based trajectory modelling. |
| **Sharp Turning Events** | Low–Moderate | Moderate | ANN models better captured aggressive manoeuvring dynamics and behavioural variability. |
| **Rapid Acceleration Events (RAE)** | Low | Low–Moderate | Secondary behavioural contribution across both modelling approaches. |
| **Behavioural Interaction Effects** | Limited | Very High | ANN models substantially improved representation of nonlinear behavioural combinations and trajectory dynamics. |
| **Model Interpretability** | Very High | Moderate–Low | Statistical models allowed direct interpretation of behavioural effects through coefficients and odds ratios. |
| **Predictive Capability** | Moderate | High | ANN models achieved superior predictive performance by capturing nonlinear behavioural patterns. |

### *4.3.4. Integrated Interpretation of Frequency and Severity Streams*

The parallel modelling strategy reveals that crash occurrence and crash consequences are governed by partially distinct yet interrelated mechanisms. Frequency models quantify the expected number of crashes under given behavioral and infrastructural conditions, whereas severity models estimate the probability that crashes result in serious outcomes under comparable conditions. Segment-level results emphasize infrastructure-driven safety performance across the network, while trip-level models highlight behavioral variability and driver–environment interaction dynamics.

The outputs of Phase 2 therefore consist of two probabilistic safety indicators for each modelling unit:

- Predicted crash frequency ($\hat{\mu_i}$)
- Predicted severity probability ($\hat{\pi_i}$)

These indicators provide complementary dimensions of safety performance but do not yet constitute a complete risk measure. Their integration and normalization procedures in Phase 3 enables construction of a composite risk index suitable for spatial screening and decision support.

## 4.4.Phase 3: Integrated Risk Index Development, Normalization, and Spatial Risk Mapping

### 4.4.1. *Composite Crash Risk Index Estimation*

Following the Phase 3 modelling framework, the predicted crash frequencies from the crash frequency models were combined with the predicted probabilities of severe outcomes from the crash severity models to compute the Composite Crash Risk Index (CCRI). The predictions were obtained from both econometric models and Artificial Neural Network (ANN) approaches, allowing the framework to capture interpretable statistical relationships as well as nonlinear patterns in the data. For each analytical unit, the CCRI was calculated as the product of predicted crash frequency and predicted severe injury probability. The index was estimated at both segment level and trip level, enabling the assessment of infrastructure-related risk across the road network as well as behaviour-driven risk along individual driving trajectories.

### *4.4.2. Econometric Model–Based Composite Risk*

When the composite risk index was derived from the econometric modelling framework, the resulting risk scores revealed substantial variability across both segments and trips. At the segment level, the composite risk values ranged from 0.000 to 1.000, indicating a wide dispersion of safety conditions

across the analyzed road network. The mean composite risk score was 0.0214, with a standard deviation of 0.0557. These results indicate that most roadway segments exhibit relatively low estimated crash risk values, while a smaller subset concentrates considerably higher risk levels. The coefficient of variation (2.60) further indicates high variability in safety performance across roadway segments, highlighting the presence of localized high-risk areas within the network.

At the trip level, the composite risk index also exhibited considerable heterogeneity across individual vehicle trajectories. The risk values ranged from 0.000 to 1.000, with a mean value of 0.0087 and a standard deviation of 0.0366. The corresponding coefficient of variation (4.19) indicates a highly dispersed distribution, suggesting that only a relatively small proportion of trips contributes disproportionately to the overall crash risk within the dataset. Table 4.9 summarizes the descriptive statistics of the composite crash risk index for the econometric model at both analytical levels.

*Table 0-9: Descriptive Statistics of the Composite Crash Risk Index (Econometric Model)*

| **Statistic** | **Segment Level** | **Trip Level** |
|---|---|---|
| Minimum Risk Value | 0.000 | 0.000 |
| Maximum Risk Value | 1.000 | 1.000 |
| Mean Risk Value | 0.0214 | 0.0087 |
| Standard Deviation | 0.0557 | 0.0366 |
| Coefficient of Variation | 2.60 | 4.19 |
| Variability Level | High | High |

**Note:** *The coefficient of variation (CV = standard deviation / mean) provides a normalized measure of dispersion, enabling comparison of variability across segment-level and trip-level risk distributions.*

### *4.4.3. FNN-Based Composite Risk*

A similar analysis was conducted using the predictions obtained from the Feedforward Neural Network (FNN) models. At the segment level, the composite risk values again ranged from 0.000 to 1.000, indicating considerable variation in safety conditions across the road network. The mean composite risk score was 0.0368, with a standard deviation of 0.0660. These results suggest that most segments exhibit relatively low estimated risk levels, while a smaller subset of segments concentrate

higher levels of crash risk. The coefficient of variation (1.79) indicates substantial variability in safety performance across segments.

At the trip level, the composite risk index ranged from 0.000 to 1.000, with a mean value of 0.0114 and a standard deviation of 0.0455. The corresponding coefficient of variation (3.98) confirms the presence of a highly skewed distribution in which most trips are associated with very low risk values, while a relatively small number of trips account for higher risk levels. Table 4.8 presents the descriptive statistics of the composite crash risk index derived from the FNN-based modelling framework.

*Table 0-10: Descriptive Statistics of the Composite Crash Risk Index (FNN Model)*

| Statistic | Segment Level | Trip Level |
|---|---|---|
| Minimum Risk Value | 0.000 | 0.000 |
| Maximum Risk Value | 1.000 | 1.000 |
| Mean Risk Value | 0.0368 | 0.0114 |
| Standard Deviation | 0.0660 | 0.0455 |
| Coefficient of Variation | 1.79 | 3.98 |
| Variability Level | High | High |

### *4.4.4. Risk Normalization and Road Network Ranking Results*

To enable consistent comparison across the road network, the composite crash risk index was normalized using min–max scaling, which rescales the raw risk values to the interval $[0, 1]$while preserving their relative differences. As in the Composite Crash Risk Index estimation, the normalization was applied to the risk values derived from both the econometric and ANN-based models. The normalized scores were then classified into five ordinal risk categories using a quantile-based approach (20% increments), reflecting the empirical distribution of crash risk across the network and enabling the identification of relatively higher- and lower-risk segments and trips for network screening and prioritization of safety interventions.

### *4.4.5. Econometric Model–Based Risk Classification*

The classification ranges obtained from the normalized composite crash risk index derived from the **econometric modelling framework** are presented in Table 4.8. At the **segment level**, a large proportion of roadway segments exhibited extremely low normalized risk values, resulting in identical lower-bound thresholds for the first three risk categories. As shown in Table 4.8, **Levels 1–3 correspond to segments with near-zero risk values**, while higher-risk segments begin to emerge in **Level 4**, where normalized risk values extend up to **0.03697**. The **highest risk category (Level 5)** includes segments with normalized risk scores ranging from **0.03697 to 1.000**, representing the most critical locations within the network where crash likelihood and severity are comparatively elevated.

At the **trip level**, the normalized risk values exhibit a more continuous distribution across the five classes. **Very low-risk trips** range from **0.00000–0.00209**, while the **low and medium risk categories** extend to **0.00384** and **0.00544**, respectively. **Higher-risk trips** fall within the **0.00544–0.00852** interval (Level 4), while trips exceeding **0.00852** belong to the **very high-risk category (Level 5)**. Trips classified within this highest category represent the most critical driving conditions and therefore require priority attention in safety monitoring and intervention strategies.

*Table 0-11: Quantile-Based Classification of Normalized Crash Risk (Econometric Model)*

| Color code | Level | Segment Level Range | Trip Level Range |
|---|---|---|---|
| **1** | Very Low | 0.00000 – 0.00000 | 0.00000 – 0.00209 |
| **2** | Low | 0.00000 – 0.00000 | 0.00209 – 0.00384 |
| **3** | Medium | 0.00000 – 0.00000 | 0.00384 – 0.00544 |
| **4** | High | 0.00000 – 0.03697 | 0.00544 – 0.00852 |
| **5** | Very High | 0.03697 – 1.00000 | 0.00852 – 1.00000 |

### *4.4.6. FNN-Based Risk Classification*

A similar normalization and classification procedure was applied to the composite crash risk index obtained from the Feedforward Neural Network (FNN) model. The resulting classification ranges are presented in Table 4.9. At the segment level, a large proportion of segments again exhibited extremely small, normalized risk values, producing identical lower-bound thresholds for the first three categories. As shown in Table 4.10, Levels 1–3 correspond to segments with near-zero risk values,

while higher-risk segments appear in Level 4, where normalized values extend up to 0.07633. The highest risk category (Level 5) captures segments with normalized risk scores between 0.07633 and 1.000, representing the most critical parts of the network.

At the trip level, the normalized risk distribution is also highly skewed. Very low-risk trips correspond to values equal to 0.00000, while the medium-risk category extends up to 0.00589. Higher-risk trips fall within the 0.00589–0.01549 interval (Level 4), whereas trips exceeding 0.01549 belong to the very high-risk category (Level 5). Trips within this highest category represent driving conditions associated with elevated behavioral instability or exposure to higher-risk roadway environments. Overall, the quantile-based classification confirms that the FNN-derived composite crash risk index exhibits a highly skewed distribution, with the majority of segments and trips characterized by very low risk values, while a relatively small proportion concentrates substantially higher crash risk.

*Table 0-12: Quantile-Based Classification of Normalized Crash Risk (FNN Model)*

| Color code | Level | Segment Level Range | Trip Level Range |
|---|---|---|---|
| 1 | Very Low | 0.00000 – 0.00000 | 0.00000 – 0.00000 |
| 2 | Low | 0.00000 – 0.00000 | 0.00000 – 0.00000 |
| 3 | Medium | 0.00000 – 0.00000 | 0.00000 – 0.00589 |
| 4 | High | 0.00000 – 0.07633 | 0.00589 – 0.01549 |
| 5 | Very High | 0.07633 – 1.00000 | 0.01549 – 1.00000 |

### *4.4.7. Spatial Distribution of Composite Crash Risk Across the Road Network*

Following the normalization and classification procedures described in the methodology, the composite crash risk indices were spatially integrated with the digital road network using Geographic Information Systems (GIS). This integration enabled the geographic visualization of safety conditions across the study area, producing network-wide maps of crash risk distribution at both the segment level and trip level, thereby supporting the identification of roadway locations and travel trajectories associated with elevated crash risk and capturing safety patterns related to both infrastructure characteristics and driving behavior.

### *4.4.8. Econometric Model–Based Spatial Risk Distribution*

The spatial distribution of the composite crash risk derived from the **econometric modelling framework** reveals noticeable variation in safety conditions across the road network. At the **segment level**, the majority of roadway segments were classified within the lower risk categories, while a smaller subset exhibited elevated risk levels. Specifically, **2,013 segments were categorized as low risk (Level 2)**, whereas **502 segments were identified as high risk (Level 4)**. This distribution suggests that most parts of the network operate under relatively safe conditions, while a limited number of segments concentrate higher estimated crash risk. The spatial distribution of these segment-level risk categories is illustrated in **Figure 4.12**, highlighting the locations where elevated crash risk is concentrated along the network.

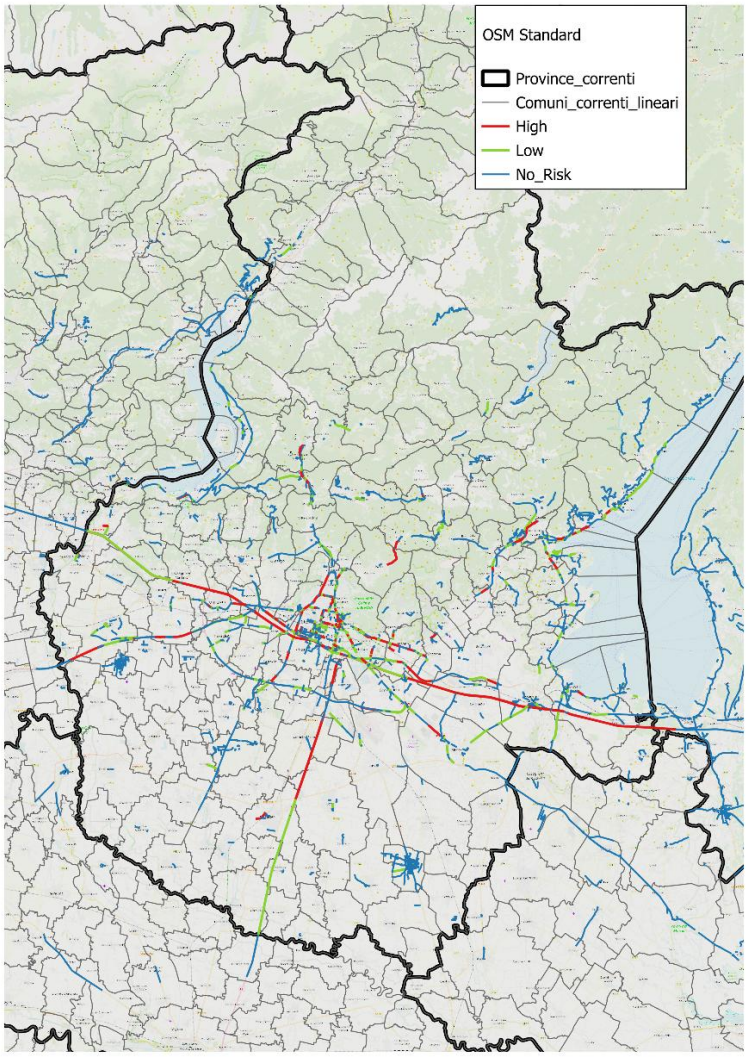



*Figure 0-15: Spatial Mapping of Segment-Level Crash Risk Categories (Econometric Model)*

At the **trip level**, the composite risk index captures behavioural instability and route-specific safety conditions derived from telematics observations. The classification results indicate a relatively balanced distribution of trips across the five risk categories, reflecting variability in driving behaviour and exposure conditions across individual trajectories. Specifically, **433 trips were classified as very low risk (Level 1)**, **430 trips as low risk (Level 2)**, **431 trips as medium risk (Level 3)**, **431 trips as high risk (Level 4)**, and **432 trips as very high risk (Level 5)**. The spatial distribution of these trip-level risk categories is illustrated in **Figure 4.13**.

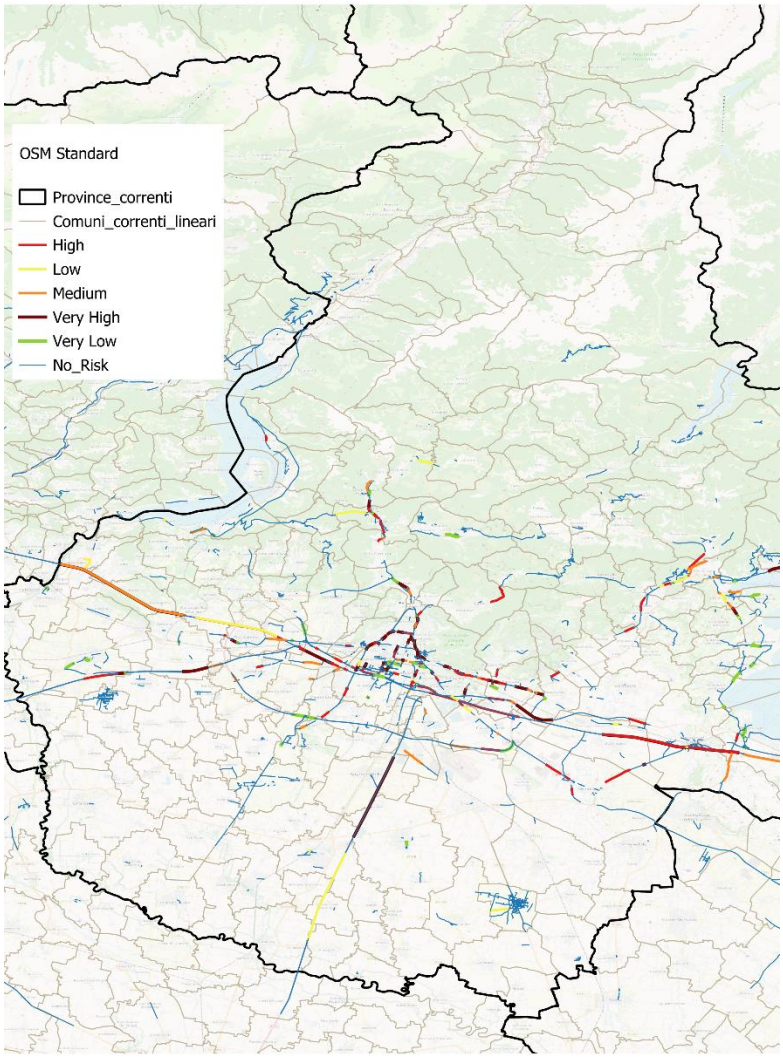



*Figure 0-16: Spatial Mapping of Trip-Level Crash Risk Categories (Econometric Model)*

The spatial visualization further reveals localized clusters of elevated risk along specific roadway corridors and travel paths, suggesting that certain combinations of infrastructure characteristics, traffic conditions, and driving behaviour contribute to higher crash likelihood.

### *4.4.9. FNN-Based Spatial Risk Distribution*

A similar spatial integration procedure was performed using the composite crash risk index derived from the **Feedforward Neural Network (FNN)** model. At the **segment level**, the spatial distribution again shows that the majority of roadway segments fall within lower risk categories, while a smaller subset exhibits higher estimated crash risk. In particular, **2,014 segments were categorized as low risk**, whereas **501 segments were identified as high risk**. These findings confirm that most portions

of the road network operate under relatively safe conditions, while a limited number of segments concentrate elevated risk levels. The spatial distribution of these segment-level risk classes is illustrated in **Figure 4.14**.

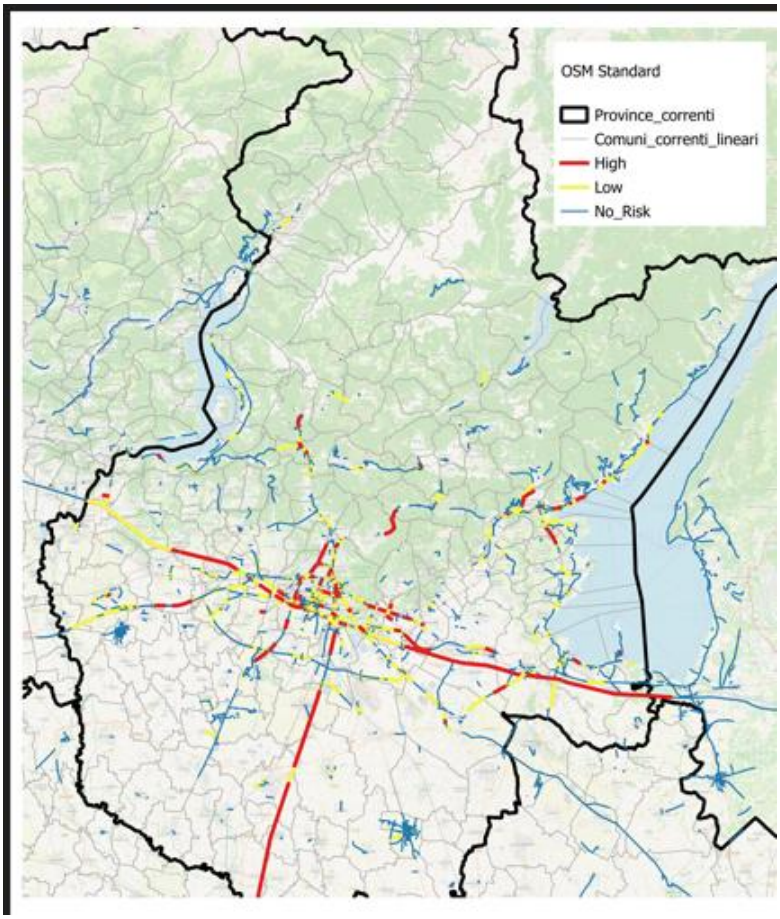



*Figure 0-17: Spatial Mapping of Segment-Level Crash Risk Categories (FNN Model)*

At the **trip level**, the composite risk index captures behavioral variability and route-specific safety conditions derived from telematics observations. The classification results indicate that **1,303 trips were categorized as low risk**, **425 trips as medium risk**, and **429 trips as high risk**. This distribution suggests that most driving trajectories are associated with relatively stable driving conditions, whereas a smaller subset of trips reflects more critical driving behavior or exposure to higher-risk roadway environments. The spatial distribution of these trip-level risk classes is illustrated in **Figure 4.15**.

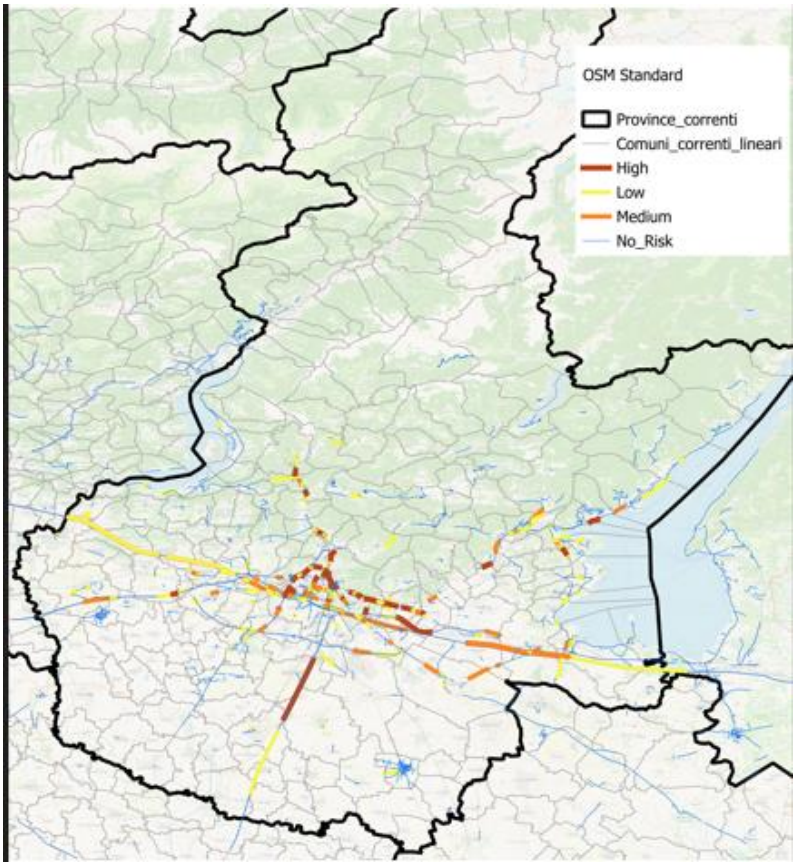



*Figure 0-18: Spatial Mapping of Trip-Level Crash Risk Categories (FNN Model)*

Overall, the spatial analysis confirms that the FNN-based composite crash risk index effectively identifies localized high-risk segments and trips within the network, providing a useful basis for targeted safety monitoring, infrastructure assessment, and risk mitigation strategies. For enlarged risk maps, see Appendix F.

# 5. CHAPTER 5

# DISCUSSION

In this chapter, the results obtained from the proposed behavior-driven and exposure-adjusted road safety framework are discussed. The discussion focuses on four main aspects. First, the role of telematics-derived behavioral indicators and Safety-Critical Events (SCEs) in explaining crash occurrence and crash severity is examined. Second, the comparative contribution of econometric and artificial intelligence models is discussed, highlighting how the two modelling approaches capture different dimensions of crash risk. Third, the implications of integrating crash frequency and severity predictions into a Composite Crash Risk Index are analyzed at both the segment and trip levels. Finally, the broader implications of the findings for proactive road safety management and possible intervention strategies are outlined.

Unlike conventional network screening approaches based solely on historical crash data, the proposed framework integrates behavioral indicators, roadway infrastructure characteristics, contextual factors, and traffic exposure within a unified modelling structure. Consequently, the discussion goes beyond identifying statistically significant variables and instead examines how the interaction between driver behavior, roadway environment, and exposure conditions shapes crash risk across the Brescia road network.

## 5.1. Behavioral indicators as central determinants of crash risk

A first major finding of this research concerns the key role of telematics-derived behavioral indicators in explaining both crash frequency and crash severity. Across both modelling scales and modelling approaches, speeding emerged as the most influential behavioral predictor, followed by indicators of vehicle instability such as Sudden Braking Events (SBE) and vehicle dynamic instability (VDE_flag). This result is broadly consistent with the Naturalistic Driving Studies (NDS) literature, where speed-related behavior and abrupt maneuvers are frequently identified as critical precursors of safety-critical situations and crash outcomes. For example, trajectory-based analyses by Yu et al. (2021) demonstrated that speed variability and following dynamics strongly influence the occurrence of high-risk events, while Bagdadi (2013) showed that harsh braking behaviour represents one of the most reliable indicators of safety-critical driving situations.

However, the present study complements these findings by introducing an important methodological extension. While most previous NDS research has focused primarily on event detection or crash prediction at a single analytical scale, the proposed framework integrates behavioural indicators within a multi-scale risk modelling structure, allowing crash risk to be examined simultaneously at both the segment level (infrastructure perspective) and the trip level (behavioural trajectory perspective). This dual perspective represents a key contribution of the study because it allows behavioural mechanisms to be interpreted in relation to both roadway environments and individual driving episodes, something that remains relatively rare in the existing NDS literature.

The distribution of Safety-Critical Events extracted from the telematics dataset provides an initial insight into the behavioural structure of the study network. Among the 15,311 detected events, approximately 77.6 % correspond to speeding behaviour, substantially exceeding all other event categories. This indicates that excessive speed constitutes the most prevalent behavioural exposure within the observed driving environment. Such a result aligns with the broader road safety literature, where speed choice is consistently identified as one of the primary determinants of crash occurrence and crash severity. Studies such as Zheng et al. (2021) and Arvin et al. (2019) similarly report that speed instability and aggressive speed adjustments significantly increase crash likelihood and crash intensity.

Nevertheless, the interpretation of the SCE distribution requires careful contextualization. The telematics dataset used in this study represents a single-day observation window, which may limit the diversity of behavioral events captured in the dataset. As a result, behaviors such as lane departures, sharp turns, or rapid accelerations may appear less frequent compared with speeding events. This does not necessarily imply that these maneuvers are unimportant contributors to crash risk; rather, it reflects the temporal coverage of the available data. Several NDS studies highlight that rare but critical maneuvers often require longer monitoring periods to be fully captured due to their stochastic nature. Consequently, although speeding dominates the observed event distribution, other SCE types still contribute meaningful explanatory power within the modelling framework.

At the segment level, speeding significantly increased expected crash frequency in the Negative Binomial model and was also identified as the most influential behavioural variable in the Feedforward Neural Network interpretation. Importantly, this effect persists even after controlling for roadway geometry, infrastructure characteristics, environmental conditions, and traffic exposure

variables. This suggests that unsafe speed choice operates as an independent behavioural risk layer rather than merely reflecting roadway design characteristics. Similar observations have been reported in previous NDS research, where behavioural volatility in vehicle dynamics was found to significantly influence crash outcomes even after accounting for roadway context (Arvin et al., 2019).

At the trip level, the behavioural role becomes even more pronounced. Because trip-level models rely primarily on telematics-derived behavioural summaries rather than static infrastructure attributes, they provide a clearer representation of the dynamic processes that shape crash outcomes during individual driving episodes. In these models, speeding again emerged as the dominant explanatory variable, while SBE and vehicle dynamic instability maintained positive contributions to crash probability and severity. Although other behaviours such as lane departure, rapid acceleration, and sharp turning exhibited smaller effects, they still provided meaningful behavioural signals within the trajectory-level risk estimation process.

From a methodological perspective, these findings demonstrate the value of integrating behavioural event mining with traditional crash-based modelling frameworks. Much of the existing NDS literature focuses either on detecting safety-critical events or on predicting crash risk but rarely combines behavioural event extraction with network-level risk modelling. By linking telematics-derived behavioural indicators with crash frequency and severity models and subsequently integrating these outputs within a composite risk framework, the present study extends the existing literature toward a more comprehensive behaviour-driven risk assessment approach.

An additional implication of these findings relates to the proactive nature of behavioural safety indicators. Traditional crash-based safety analyses are inherently retrospective because they rely solely on observed crash records. In contrast, Safety-Critical Events capture near-miss situations and behavioural instabilities that occur before crashes materialize. As emphasized in earlier NDS research, including Antin et al. (2019), the continuous monitoring of driver behaviour provides valuable insights into the mechanisms leading to crashes. The present results support this perspective by demonstrating that telematics-derived SCEs contain substantial predictive information for both crash occurrence and crash severity.

Overall, the results confirm that behavioural instability captured through telematics-derived Safety-Critical Events plays a central role in explaining crash risk across both spatial and operational analytical scales. More importantly, by integrating behavioural indicators within a dual segment-level and trip-level modelling framework, this study extends previous NDS research and provides a more comprehensive understanding of how dynamic driving behaviour interacts with infrastructure, exposure, and environmental conditions to shape crash risk across the road network.

### 5.2. The role of infrastructure, context, and exposure in shaping crash occurrence and consequences

Although behavioural indicators emerged as a key driver of crash risk within the proposed framework, the results demonstrate that crash occurrence and crash consequences cannot be explained solely through driver behaviour. Instead, crash risk emerges from the interaction between behavioural instability, roadway infrastructure characteristics, traffic exposure, and contextual operating conditions. Previous Naturalistic Driving Studies (NDS) have demonstrated the importance of these factors individually; however, most investigations have analysed them separately or within localized experimental contexts. The present study contributes to addressing this limitation by integrating behavioural indicators, infrastructure characteristics, and exposure measures within a unified modelling framework applied at both the segment and trip levels of the road network.

Earlier NDS research has consistently highlighted the influence of roadway and traffic environments on crash risk. For example, Abdelrahman et al. (2018) demonstrated that adverse road surface conditions combined with traffic density significantly increase the probability of crash and near-crash events, while Zurlinden et al. (2020) showed that complex traffic environments increase driver workload and elevate crash likelihood. Similarly, Yu et al. (2021) found that disturbed traffic flow patterns and close vehicle interactions contribute significantly to the emergence of high-risk traffic situations. These studies emphasize the importance of environmental and traffic conditions in shaping crash risk, but they generally examine these factors either in isolation or within localized roadway segments rather than within an integrated network-wide analytical structure.

The results of the present research extend these findings by demonstrating how specific infrastructure configurations influence the translation of behavioural instability into crash outcomes at the network scale. The segment-level crash frequency models indicate that three-lane roads, signalized intersections, and lighted tunnels are associated with higher crash frequencies. These roadway environments represent locations where traffic interactions intensify and where drivers must process

multiple stimuli simultaneously. Intersections concentrate conflicting traffic streams and require rapid decision-making, while tunnels introduce constrained geometric environments characterized by visibility transitions and reduced manoeuvring space. Multi-lane segments similarly increase driving complexity due to overtaking and lane-changing manoeuvres. These results are consistent with earlier studies showing that roadway complexity and interaction density significantly influence crash risk formation (Zurlinden et al., 2020; Yu et al., 2021; Zheng et al., 2021).

A further contribution of the present research lies in the explicit treatment of traffic exposure within the modelling framework. Many earlier studies relied primarily on crash counts or behavioural indicators without clearly separating traffic demand from intrinsic roadway safety performance. However, crash risk is inherently influenced by exposure because higher traffic volumes increase the number of potential vehicle interactions. In the present framework, traffic exposure was explicitly incorporated in the crash frequency models through reconstructed AADT estimates. The resulting exposure distribution revealed strong heterogeneity across the Brescia road network, with a relatively small subset of segments carrying disproportionately high traffic volumes. Accounting for this heterogeneity allows the framework to distinguish between segments with high crash counts due to traffic demand and those exhibiting elevated intrinsic risk even after exposure adjustment. This distinction is essential for reliable network screening and aligns with previous methodological work emphasizing the need to control for exposure when modelling crash occurrence (Jovanis et al., 2011; Arvin et al., 2019).

Beyond crash occurrence, the present study also highlights the importance of infrastructure and contextual conditions in shaping crash severity outcomes. While numerous NDS studies focus on predicting crash or near-crash events, comparatively few incorporate severity modelling within the same analytical framework. For instance, several investigations have focused on behavioural precursors such as driver distraction, glance behaviour, and secondary tasks (Bärgman et al., 2015; Bakhit et al., 2018; Sharmin et al., 2023), while others concentrate on predictive modelling of high-risk events using machine learning techniques (Das & Ahmed, 2021). Although these studies provide valuable insights into behavioural risk mechanisms, they often stop at identifying the likelihood of safety-critical events without explicitly linking them to crash consequences.

The present research addresses this limitation by explicitly modelling crash severity alongside crash frequency within a unified framework. The results indicate that roadway configuration and crash context influence not only whether crashes occur but also how severe their consequences become. Multi-lane road segments and specific collision configurations such as frontal-lateral or head-on crashes were associated with increased probabilities of severe injury outcomes. These findings reinforce the argument that roadway environments influence both the probability of crash occurrence and the physical dynamics that determine crash consequences.

Behavioural indicators remain strongly intertwined with these contextual factors. Consistent with previous NDS research, speeding emerged as one of the most influential predictors of severe crash outcomes. Earlier studies have shown that instability in vehicle dynamics and speed variability significantly increase crash intensity and injury severity (Arvin et al., 2019). The results of the present study confirm this relationship while extending the analysis by integrating behavioural indicators with infrastructure characteristics and exposure conditions within the same modelling architecture. This integrated perspective enables the identification of roadway segments where behavioural instability interacts with roadway design and traffic conditions to generate elevated crash risk.

Another methodological contribution concerns the differentiation between structural and dynamic exposure conditions. At the segment level, exposure reflects long-term traffic demand measured through reconstructed AADT values. At the trip level, however, exposure represents the dynamic traffic environment experienced during individual driving trajectories. The results suggest that mean exposure during a trip influences the number of interaction opportunities and therefore contributes to crash occurrence, whereas peak exposure conditions are more closely associated with severe crash outcomes. This distinction provides new insight into how traffic intensity influences crash risk across different analytical scales.

Overall, these findings reinforce the view that crash risk should be interpreted as a system-level phenomenon resulting from the interaction between driver behaviour and roadway context. While previous studies have examined behavioural risk factors, environmental conditions, or infrastructure characteristics individually, the present research advances the literature by integrating behavioural indicators, infrastructure characteristics, exposure conditions, and severity outcomes within a multi-scale modelling framework. By combining telematics-derived behavioural data with traditional crash

and infrastructure datasets, the proposed approach provides a more comprehensive understanding of how behavioral actions, roadway environments, and traffic demand jointly shape crash risk across the road network.

### 5.3. Comparing econometric and artificial intelligence models

A second important discussion point concerns the comparison between econometric models and artificial intelligence models for crash risk estimation. Traditional econometric approaches, such as the Negative Binomial model for crash frequency and Binary Logistic Regression for crash severity, have long represented the dominant analytical tools in transportation safety research because of their interpretability and their ability to estimate marginal effects, elasticities, odds ratios, and incidence rate ratios (Lord & Mannering, 2010; Mannering & Bhat, 2014). These characteristics are particularly valuable for policy design, since they allow direct quantification of how behavioural, infrastructural, environmental, and exposure-related variables influence crash outcomes. However, such models rely on predefined functional relationships and often assume linear or log-linear associations between explanatory variables and safety outcomes. As a result, they may not fully capture the complex nonlinear mechanisms through which crashes emerge in real-world traffic systems.

In contrast, artificial intelligence models such as Feedforward Neural Networks (FNNs) can learn relationships directly from data and capture nonlinear interactions among behavioural, infrastructural, contextual, and exposure-related variables without imposing restrictive parametric assumptions. The results obtained in this dissertation demonstrate that crash risk formation is influenced by interacting mechanisms that are only partially represented within traditional econometric formulations. This finding is consistent with recent studies showing the growing value of machine learning in road safety analytics (Roussou et al., 2024; Golestani et al., 2025; Khan et al., 2026).

For crash frequency modelling at the segment level, the FNN model demonstrated substantially higher predictive performance than the Negative Binomial model, achieving lower RMSE and MAE values and stronger predictive accuracy. This result suggests that crash occurrence mechanisms involve nonlinear combinations of driver behaviour, roadway characteristics, and traffic exposure. Previous studies have similarly reported improved performance when machine-learning methods are applied to crash prediction tasks (Yeo et al., 2021; Theofilatos et al., 2019). However, many of these investigations focus primarily on maximizing prediction accuracy rather than integrating behavioural indicators derived from Naturalistic Driving Data with roadway infrastructure and exposure variables

within a coherent safety framework. The present study addresses this limitation by embedding machine learning within a behaviour-driven and exposure-adjusted risk modelling structure.

At the trip level, the FNN model again outperformed the Negative Binomial model, although the predictive improvement was smaller. This outcome is expected because trip-level crash occurrence is inherently more stochastic than segment-level crash frequency. Individual trips represent short-term behavioural processes influenced by transient factors such as driver decision-making, traffic conditions, route choice, and environmental variability. As a result, trip-level outcomes exhibit greater randomness and lower structural predictability than fixed roadway segments. Nevertheless, the stronger performance of the FNN model indicates that machine-learning approaches are particularly effective for extracting behavioural signals from trajectory-based telematics data. Similar conclusions have recently been reported in telematics-based driver behaviour studies (Moosavi & Ramnath, 2023; Yedilkhan et al., 2025).

A similar pattern emerged in crash severity modelling. At the segment level, the FNN classifier substantially outperformed the Binary Logit model in terms of AUC, balanced accuracy, and Brier score. These findings indicate that severe crash outcomes are shaped by nonlinear combinations of behavioural indicators, roadway attributes, collision configurations, and exposure conditions. Neural-network approaches have increasingly demonstrated strong performance in injury severity modelling because they can capture complex interactions that are difficult to represent through parametric models (Behboudi et al., 2024; Prajapati et al., 2023; Somvanshi et al., 2026). The SHAP analysis conducted in this dissertation further confirms that speeding behaviour, lane configuration, vehicle type, and roadway context jointly influence severity risk, indicating that severe crash outcomes emerge from combined effects rather than isolated factors.

An important practical distinction concerns interpretability. Econometric models allow coefficients to be translated directly into actionable policy insight. For example, the positive coefficient associated with speeding in the Negative Binomial model indicates that increases in speeding-related Safety-Critical Events are associated with higher expected crash counts, holding other variables constant. In practical terms, this implies that corridors characterized by persistent speeding behaviour should receive priority for speed management interventions such as automated enforcement, variable speed control, traffic calming, or targeted police operations. Likewise, positive effects associated with

signalized intersections, multilane road sections, and tunnel environments suggest that traffic complexity and conflict density increase crash occurrence, thereby identifying locations where engineering redesign or operational improvements may yield substantial safety benefits.

By contrast, machine-learning models provide stronger predictive capability but less transparent structural interpretation. For this reason, the two modelling paradigms should not be viewed as competing alternatives. Rather, they should be considered complementary analytical tools within a unified decision-support framework. Econometric models are particularly valuable for explaining why risk increases and for supporting evidence-based policy communication. Artificial intelligence models are especially valuable for identifying where and when elevated risk is most likely to occur under nonlinear and high-dimensional conditions.

An additional methodological contribution of this dissertation concerns the incorporation of Safety-Critical Events (SCEs) derived from Naturalistic Driving Data directly into both crash frequency and crash severity modelling streams. In much of the earlier literature, SCEs such as sudden braking, harsh turning, lane departure, or rapid acceleration are used primarily as surrogate indicators of conflict frequency or near-crash occurrence rather than as explanatory variables for formal crash consequence modelling (Tivesten & Dozza, 2014; Zheng et al., 2014). By contrast, the present study demonstrates that behavioural instability captured through telematics data provides meaningful explanatory power not only for crash occurrence but also for crash severity outcomes.

Another limitation of much of the existing literature concerns spatial scale. Many studies conduct safety analysis at isolated intersections, selected corridors, or localized hotspots (Abdel-Aty & Pande, 2005; Dong et al., 2014). While such approaches are useful for site-specific interventions, they often fail to capture broader systemic risk patterns across an entire transportation network. Recent studies have increasingly called for network-wide safety frameworks capable of identifying distributed risk concentrations across large road systems (Zeng et al., 2019; Theofilatos et al., 2019; Berhanu et al., 2024). However, relatively few studies combine behavioural indicators, infrastructure variables, exposure measures, and machine-learning approaches within a unified network-level framework.

The methodological framework developed in this dissertation addresses these limitations by integrating crash frequency modelling, crash severity modelling, and behavioural safety indicators

within a unified network-wide analytical structure. Rather than analysing crash mechanisms solely at localized roadway sites, the framework evaluates safety conditions across the entire road network while simultaneously incorporating behavioural indicators derived from telematics data. Furthermore, the analysis is conducted at two complementary analytical scales—roadway segments and individual driving trips—allowing the framework to capture both infrastructure-driven safety mechanisms and behaviour-driven safety mechanisms.

Recent international research trends increasingly emphasize the transition from reactive crash-based analysis toward proactive, AI-supported, and behaviour-driven road safety management. Moosavi and Ramnath (2023) developed a context-aware telematics framework for real-time driver risk estimation, demonstrating the value of behavioural data beyond traditional crash records. Roussou et al. (2024) used machine learning to compare driving behaviour patterns across Germany, Belgium, and the United Kingdom, highlighting behavioural heterogeneity across countries. Golestani et al. (2025) combined telematics big data with interpretable machine learning to examine hotspot risk using weather, behavioural, and spatiotemporal variables. Khan et al. (2026) further emphasized the broader role of artificial intelligence in crash prediction, near-miss detection, and proactive interventions. While these studies represent significant advances, they generally focus on specific dimensions of safety risk such as driver-level prediction, hotspot detection, behavioural classification, or severity estimation. The present dissertation extends beyond these benchmark studies by jointly modelling crash occurrence and crash severity, explicitly incorporating traffic exposure, integrating roadway infrastructure variables, and operating simultaneously at both segment and trip levels. In doing so, the framework provides a broader system-level representation of road safety risk than many existing telematics-based approaches.

A further contribution lies in the synthesis of econometric and artificial intelligence methods within a single analytical pipeline. Econometric models provide interpretable parameter estimates and allow statistical inference regarding the influence of explanatory variables. Machine-learning models, on the other hand, offer stronger predictive capability and can uncover nonlinear relationships that may remain hidden in traditional modelling approaches. The combined use of these paradigms therefore allows the framework to benefit simultaneously from the interpretability of econometric models and the predictive strength of artificial intelligence.

Finally, the modelling outputs are integrated into a Composite Crash Risk Index that directly supports contemporary Vision Zero and Safe System strategies. These policy frameworks recognize that fatalities and serious injuries are preventable, that human errors are inevitable, and that transport systems must be designed to tolerate mistakes without resulting in death or severe harm. Practically, the developed risk index provides a decision-support tool for infrastructure prioritization, targeted enforcement, route-risk screening, and potentially safer route guidance, allowing authorities and navigation systems to compare routes not only by travel time or cost, but also by predicted safety risk.

Overall, the findings of this research suggest that econometric and artificial intelligence models should not be viewed as competing alternatives but rather as complementary analytical tools. Econometric models remain essential for inference and policy interpretation, while machine-learning models provide enhanced predictive performance and the ability to capture complex behavioural–infrastructure interactions. The principal methodological contribution of this dissertation lies in combining these modelling approaches within a multi-level, behaviour-driven, and exposure-adjusted framework that integrates Naturalistic Driving Data with roadway infrastructure and crash outcomes, thereby providing a more comprehensive understanding of crash risk dynamics across transportation networks.

While the comparative findings are encouraging, the interpretation of model performance should also be considered alongside methodological limitations, uncertainty sources, and practical implementation issues. Despite the strong empirical performance of the proposed models, several limitations should be acknowledged. First, the telematics dataset represents a limited observation window of approximately twenty-four hours. Although sufficient to reveal meaningful behavioural patterns, longer observation periods would likely capture greater seasonal, weekly, and weather-related variability. Rare manoeuvres and low-frequency precursors to severe crashes may therefore be underrepresented. Second, traffic exposure was reconstructed through scaled proxy measures because continuous count data were not universally available across the study network. While the adopted methodology provides a defensible approximation of relative demand levels, some uncertainty remains in absolute exposure estimates. Third, machine-learning models may be sensitive to sample composition, class imbalance, and hyperparameter specification. Although validation procedures and explainable AI tools were employed, some uncertainty remains regarding transferability under alternative datasets or operational contexts. Fourth, historical crash records may suffer from

underreporting, particularly for minor crashes, which may affect calibration of supervised severity models. Finally, the transferability of the Brescia case-study results to other jurisdictions should be approached cautiously, since driver culture, infrastructure design standards, enforcement intensity, and fleet composition may differ across regions.

From a practical perspective, the comparative findings generate several direct implications for real-world road safety management. Locations identified as high risk by both econometric and AI models represent particularly robust intervention priorities, since elevated risk is confirmed by two distinct analytical paradigms. Repeated speeding signals detected through telematics data provide opportunities for proactive speed management before crashes accumulate, especially where historical crash data are sparse or delayed. Similarly, recurrent sudden braking clusters may indicate hidden operational problems such as poor sight distance, abrupt merging conditions, inadequate signal coordination, or inconsistent speed environments. These sites may not appear in conventional crash rankings yet may still warrant engineering review. Segment-level risk outputs can support infrastructure investment prioritization, while trip-level risk outputs can support fleet management, driver coaching, insurance analytics, and intelligent in-vehicle warning systems. Finally, the integration of modelling outputs into GIS-based risk maps provides transport authorities with a practical decision-support platform for allocating limited safety budgets toward interventions with the highest expected safety return.

#### *5.3.1 Reflection on Unexpected Findings and Model Limitations*

Although the empirical results generally aligned with theoretical expectations, several findings merit deeper reflection. First, while the Feedforward Neural Network models consistently outperformed the econometric models in predictive accuracy, the magnitude of improvement was smaller at the trip level than at the segment level. This may reflect the inherently stochastic nature of individual trips, where transient behavioural decisions, route choice variability, momentary traffic conditions, and short-term environmental influences introduce randomness that limits model predictability even under flexible machine-learning structures. In contrast, segment-level conditions are relatively more stable and therefore allow stronger structural patterns related to infrastructure, traffic exposure, and recurrent behavioural tendencies to be learned more effectively.

Second, the role of behavioural indicators differed across analytical scales. At the segment level, behavioural variables such as speeding and sudden braking were influential, but their effects were partially moderated by roadway geometry, traffic demand, and infrastructure configuration. This is expected because segment-level crash outcomes represent aggregated safety performance, where repeated behavioural patterns are embedded within broader contextual conditions. Conversely, at the trip level, behavioural indicators became more directly connected to crash risk because each trip captures immediate driver actions and real-time operational decisions. Therefore, behaviours such as speeding, harsh braking, and instability signals naturally exhibit stronger behavioural meaning at the trip scale, where risk formation is more closely tied to moment-to-moment driving dynamics.

Third, some behavioural indicators such as lane departure and sharp turning, although theoretically associated with elevated crash risk, appeared less frequent and occasionally exhibited weaker statistical influence than expected. This result likely reflects the short observation window of the telematics dataset and the relatively rare occurrence of such manoeuvres during routine daily driving. Their lower frequency should therefore not be interpreted as a lack of safety relevance, but rather as a data availability constraint associated with limited temporal coverage. Under longer observation periods, these manoeuvres would likely provide stronger explanatory value.

Fourth, the dominance of speeding across both frequency and severity models deserves particular reflection. While speeding is widely recognized as a major safety determinant, its repeated prominence in this study suggests that speed choice may also act as a proxy for broader behavioural aggressiveness, risk tolerance, or adaptation to roadway conditions. In this sense, speeding may represent not only an isolated violation, but also a broader behavioural signature associated with unsafe driving styles.

Fifth, exposure-related and infrastructure variables remained highly influential in several models, confirming that behavioural risk does not operate independently from roadway context. In practical terms, unsafe behaviour may produce substantially different safety outcomes depending on whether it occurs on urban streets, multilane corridors, tunnels, rural roads, or intersections. For example, sudden braking on a congested urban street may generate minor rear-end conflicts, whereas the same manoeuvre on a high-speed corridor may produce much more severe consequences. This reinforces

the importance of integrated behavioural–infrastructural modelling rather than driver-only approaches and supports the multi-factor structure adopted in this dissertation.

Sixth, the Composite Crash Risk Index also revealed that high crash frequency does not always coincide with high crash severity. Some segments exhibited relatively frequent crashes with moderate consequences, while other locations showed fewer crashes but substantially higher severity potential. This confirms the conceptual importance of jointly modelling occurrence and consequence rather than relying solely on crash counts for safety prioritization.

Several methodological limitations should also be acknowledged. The telematics dataset covered approximately one day of observation, which may not fully capture weekly, seasonal, or adverse-weather variability in traffic demand and driving behaviour. Traffic exposure was reconstructed using proxy scaling methods due to incomplete count data, introducing some uncertainty in absolute demand estimates. Historical crash databases may also underreport minor crashes or non-injury collisions, potentially affecting calibration of supervised frequency and severity models. Finally, while the Brescia case study provides a valuable and heterogeneous European application, direct transferability to other jurisdictions should be approached cautiously because of differences in driver culture, enforcement intensity, fleet composition, land-use structure, and roadway design standards.

Despite these limitations, the principal conclusions of the study remain robust. The consistency of results across econometric and artificial intelligence models, together with the repeated importance of behavioural indicators such as speeding and sudden braking, provides strong evidence that telematics-derived behaviour contains substantial predictive information for proactive road safety assessment.
Overall, these unexpected findings and limitations do not weaken the main contributions of the study. Rather, they highlight the complexity of crash risk formation and reinforce the need for future research using longer observation periods, richer multi-source datasets, panel-based telematics observations, and external validation across multiple jurisdictions.

### 5.4. Integrated Interpretation of Crash Frequency and Severity Streams

A central conceptual contribution of this research lies in explicitly treating crash frequency and crash severity as two complementary yet analytically distinct dimensions of road safety performance. Previous studies have widely recognized that the mechanisms governing crash occurrence differ from

those influencing crash consequences (Lord & Mannering, 2010; Mannering & Bhat, 2014). Crash frequency models generally describe how often crashes occur under given infrastructural, operational, and traffic conditions, while crash severity models estimate the probability that those crashes will result in serious or fatal outcomes.

The modelling results obtained in this study confirm this distinction. Crash frequency appears to be more strongly associated with traffic exposure, operational complexity, and recurring behavioral instability, whereas crash severity is influenced more strongly by the intensity of driver behavior and crash configuration. This differentiation between occurrence and consequence mechanisms is consistent with previous road safety research, which indicates that exposure-related factors and infrastructure characteristics tend to influence crash likelihood, while behavioral intensity and collision dynamics are more closely associated with injury severity outcomes (Elvik et al., 2009; Theofilatos et al., 2019).

This distinction is also evident in the empirical results obtained in the present study. Some roadway attributes that increase crash occurrence do not necessarily correspond to the highest severity outcomes. Similarly, locations with relatively low crash frequencies may still generate severe consequences when crashes occur. Such discrepancies between crash occurrence patterns and injury severity distributions have been documented in previous network safety studies and reinforce the importance of analyzing these two dimensions separately before integrating them into a comprehensive safety risk framework.

Behavioral indicators derived from Naturalistic Driving Data further illustrate this difference. Speeding, for instance, contributes significantly to both crash frequency and crash severity. However, its effect is substantially larger in the severity models, particularly at the trip level where the binary logit model revealed a very large odds ratio associated with speeding behavior. This finding suggests that while multiple behaviors may trigger conflict situations, excessive speed plays a particularly critical role in determining whether those conflicts escalate into severe injury outcomes. This observation aligns with extensive road safety research demonstrating that speed is one of the most influential determinants of crash severity due to its direct relationship with impact energy during collisions (Aarts & van Schagen, 2006).

For this reason, the integration of predicted crash frequency and predicted severity probability into a Composite Crash Risk Index is methodologically justified. A location should not be considered safe simply because crash counts are low if the few crashes that occur result in severe outcomes. Conversely, locations experiencing minor conflicts should not necessarily receive the same intervention priority as sites where fewer crashes produce more serious consequences. By combining the probability of crash occurrence with the probability of severe outcomes, the Composite Crash Risk Index provides a more balanced representation of safety performance.

This integrated risk perspective is consistent with modern road safety management frameworks that emphasize the importance of considering both crash probability and consequence severity when prioritizing safety interventions (Mannering et al., 2020). In this study, the predicted crash frequency obtained from the crash frequency models was combined with the predicted probability of severe outcomes derived from the crash severity models to compute the Composite Crash Risk Index for each analytical unit. The resulting indicator therefore represents the expected severity-weighted crash occurrence under given behavioral, infrastructural, and environmental conditions.

### 5.5. Interpretation of the Composite Crash Risk Index at Segment and Trip Levels

The Composite Crash Risk Index reveals substantial heterogeneity in safety conditions across both roadway segments and individual driving trajectories. The spatial integration of the normalized risk indices with the digital road network, performed using Geographic Information Systems (GIS), enabled the visualization of safety conditions across the study area and provided additional insight into how crash risk is distributed geographically across the network. The resulting risk maps, presented in Chapter 4, illustrate the spatial variability of crash risk at both the segment and trip levels, highlighting locations and trajectories associated with elevated safety concerns. For enlarged risk maps, see Appendix F.

At the segment level, the econometric-based risk scores exhibit considerable dispersion across the road network. Although the mean risk value remains relatively low (0.0214), the standard deviation (0.0557) and coefficient of variation (2.60) indicate a highly uneven distribution of risk across roadway segments. This pattern suggests that most segments operate under relatively safe conditions, while a smaller subset concentrates disproportionately higher levels of crash risk. The spatial mapping results further confirm this pattern, showing that high-risk segments are not uniformly distributed but instead appear as localized clusters along specific roadway corridors.

Such concentration of risk in a limited number of locations is consistent with the principles of network safety analysis, where safety problems are rarely distributed uniformly across the transportation system. Instead, a relatively small number of critical segments often account for a significant share of the total crash risk within the network. The classification results obtained in this study reinforce this interpretation. A large proportion of roadway segments were categorized as low risk, while a smaller subset was classified as high risk. Specifically, 2,013 segments were identified as low risk and 502 segments were classified as high risk within the analyzed network using the econometric model. These findings suggest that targeted safety interventions focusing on high-risk segments may yield substantial improvements in overall network safety without requiring uniform investment across the entire roadway system.

At the trip level, however, the crash risk distribution exhibits a different structure. The mean risk value for trips is considerably lower (0.0087), but the coefficient of variation is substantially higher (4.19), indicating even greater relative variability across individual driving trajectories. Unlike the segment-level distribution, which concentrates risk in a limited number of locations, the trip-level results show a more continuous dispersion of risk across driving episodes. The classification results further illustrate this behavioral variability. Trips are distributed relatively evenly across the five risk categories, with 433 trips classified as very low risk, 430 as low risk, 431 as medium risk, 431 as high risk, and 432 as very high risk. The spatial visualization confirms that these variations are not confined to specific locations but occur across different routes and driving contexts, reflecting the dynamic nature of driver behavior captured through telematics data.

The spatial mapping results derived from the Feedforward Neural Network (FNN) model further reinforce these findings. At the segment level, the FNN-based composite risk index similarly identified 2,014 segments as low risk and 501 segments as high risk, confirming that only a limited portion of the network concentrates elevated crash risk. At the trip level, the FNN-based classification produced three risk categories, with 1,303 trips categorized as low risk, 425 trips as medium risk, and 429 trips as high risk. The corresponding spatial maps show that high-risk trajectories are distributed across multiple travel paths, suggesting that behavioral variability and situational factors play a significant role in crash risk formation.

The dual-scale structure of the proposed framework therefore provides two complementary perspectives on road safety performance. Segment-level analysis highlights locations where safety

problems are structurally embedded within the road network, often due to infrastructure characteristics, traffic conditions, or geometric design features. However, these structural conditions do not operate independently; their safety implications are strongly influenced by how drivers interact with the roadway environment. Driver behavior and driver-related characteristics such as speed choice, maneuvering strategies, and responses to complex traffic situations can amplify or mitigate the risks associated with specific infrastructure configurations. As a result, the same roadway design may produce different safety outcomes depending on the behavioral patterns of drivers operating within that environment.

This interaction between driver behavior and infrastructure has long been recognized as a fundamental component of crash causation, where crashes emerge from the dynamic interaction between drivers, vehicles, and roadway environments (Lord & Mannering, 2010). However, many previous network screening approaches have primarily focused on infrastructure and crash data without explicitly incorporating behavioral indicators derived from real-world driving observations. For example, the Risk-Based Network-Wide Road Safety Screening framework proposed by Winlaw et al. (2019) integrates crash occurrence likelihood, crash severity probability, and traffic exposure to compute a network-wide risk score for each road unit. While this framework represents an important advancement in risk-based infrastructure safety assessment, its primary focus remains on infrastructure-level crash prediction models derived from historical crash data.

Despite the substantial progress made in recent years, much of the existing literature on telematics-based road safety modelling and Naturalistic Driving Studies remains concentrated in North America, where large-scale datasets such as SHRP2 have traditionally dominated empirical research. Consequently, evidence from European urban and regional contexts remains comparatively limited, particularly regarding integrated frameworks that combine behavioural indicators, crash outcomes, infrastructure conditions, and exposure measures. Considering the rapid emergence of recent studies (2024–2026), the international research direction is clearly moving toward proactive, AI-supported, and behaviour-driven safety management. In this context, this dissertation contributes by applying and validating a proactive road safety framework in the Province of Brescia, thereby extending the applicability of telematics-based safety modelling to a heterogeneous European road environment and demonstrating its relevance beyond the North American setting.

Recent literature has increasingly emphasized the use of telematics data, Naturalistic Driving Studies (NDS), and artificial intelligence to shift road safety analysis from reactive crash-based assessment toward proactive risk prediction. Moosavi and Ramnath (2023) developed a context-aware framework using telematics variables to estimate driver risk in real time, demonstrating the value of behavioural and environmental information beyond traditional crash records. However, their focus was primarily on driver-level risk estimation, whereas this dissertation extends the analysis to both road-segment and trip-level safety risk, linking behavioural indicators with network screening and infrastructure conditions. Roussou et al. (2024) applied machine learning to compare driving behaviour patterns across Germany, Belgium, and the United Kingdom, highlighting behavioural heterogeneity across countries. In contrast, the present study moves beyond behavioural classification by explicitly connecting risky driving behaviours to crash frequency, crash severity, and composite network risk.

Osman and Hajij (2021) demonstrated that advanced deep learning techniques can outperform conventional approaches in crash prediction for advanced driver support systems. While their contribution focused on predictive accuracy, this dissertation combines predictive models with interpretable econometric methods, integrating Negative Binomial Regression, Logistic Regression, and Feedforward Neural Network approaches to balance explanation and prediction. Similarly, Castermans et al. (2024) showed that connected vehicle data can improve accident prediction. The present research advances this concept by transforming telematics observations into explicit Safety-Critical Events (SCEs) such as sudden braking, lane departure, speeding, and instability, which serve as proactive behavioural precursors of crash risk.

More recent studies further reinforce this transition. Khan et al. (2026) discussed the broader role of artificial intelligence in crash risk prediction, near-miss detection, and proactive interventions. This dissertation operationalizes that vision through a measurable framework that converts behavioural and contextual data into a Composite Crash Risk Index. Somvanshi et al. (2026) used Naturalistic Driving Data and neural networks to predict crash severity, whereas the present study jointly models both crash occurrence and crash severity, recognizing that risk depends on both probability and consequence. Golestani et al. (2025) combined telematics big data with interpretable machine learning to examine hotspot risk through behavioural, weather, and spatiotemporal factors. In comparison, this dissertation integrates those dimensions with explicit traffic exposure and roadway infrastructure variables within a unified modelling structure.

Liang et al. (2025) proposed real-time lane-level crash detection on freeways using sparse telematics data, mainly focusing on immediate operational crash identification. By contrast, the present framework is designed for strategic and network-level safety management, supporting long-term prioritization of high-risk corridors and trips. Berhanu et al. (2024) linked crash prediction with spatial network analysis for safer route planning. This dissertation extends that idea by introducing route comparison based not only on travel time or congestion, but also on predicted crash risk. Yedilkhan et al. (2025) used deep learning models to identify risky driver behaviour from telematics sensor data, while this research further links detected behaviours to observed crash outcomes and infrastructure context.

Overall, the findings indicate that the framework developed in this dissertation advances literature in several important ways. It explicitly incorporates behavioural indicators derived from Naturalistic Driving Data, including Safety-Critical Events such as sudden braking, rapid acceleration, lane departures, speeding, and vehicle instability. It simultaneously operates at two complementary analytical scales road segments and individual driving trips allowing the identification of both infrastructure-related risk concentrations and behaviour-driven safety dynamics. It jointly models crash frequency and crash severity, recognizing that road safety risk depends not only on how often crashes occur but also on their likely consequences. Furthermore, it combines econometric and artificial intelligence approaches, integrating interpretability with predictive power and capturing nonlinear relationships among behavioural, infrastructural, environmental, and exposure-related factors.

Finally, the modelling outputs are synthesized into a Composite Crash Risk Index that directly supports contemporary Vision Zero and Safe System strategies. These policy frameworks recognize that fatalities and serious injuries are preventable, that human errors are inevitable, and that transport systems must be designed to tolerate mistakes without resulting in death or severe harm. Practically, the developed risk index provides a decision-support tool for infrastructure prioritization, targeted enforcement, route-risk screening, and real-time safer route guidance, allowing users or navigation systems to compare routes not only by travel time or monetary cost, but also by predicted safety risk. Overall, this research contributes a unified, scalable, and behaviour-informed framework for proactive road safety assessment and network-level risk management.

# 6. CHAPTER 6

# CONCLUSIONS, CONTRIBUTIONS, AND FUTURE RESEARCH DIRECTIONS

## 6.1. Addressing the Research Gaps

This dissertation contributes to addressing several important limitations identified in the contemporary literature on telematics-based road safety modelling, Safety-Critical Events (SCEs), and crash risk estimation. Existing studies have often been constrained to localized roadway sections, specific manoeuvres, or driver-level behavioural classification tasks, thereby limiting their capacity to support strategic network-wide safety management. Many studies also focus primarily on crash occurrence or near-miss detection without explicitly integrating crash severity, while others analyse frequency and severity separately rather than within a unified risk framework. In addition, much of the empirical evidence remains concentrated in North American Naturalistic Driving Study programmes, reducing the geographical diversity and transferability of findings to European transport environments.

This research advances beyond the current state of the art by proposing and empirically validating a multi-scale behaviour-driven crash risk framework that jointly integrates telematics-derived behavioural indicators, roadway infrastructure characteristics, environmental conditions, and traffic exposure variables. Unlike prior approaches centred only on drivers or isolated road segments, the framework simultaneously operates at two complementary analytical scales:

- Segment level, supporting hotspot screening, infrastructure prioritization, and network planning;
- Trip level, capturing behavioural instability and dynamic risk during individual journeys.

A further innovation is the explicit joint modelling of crash frequency and crash severity, recognizing that road safety risk depends not only on how often crashes occur, but also on the seriousness of their likely consequences. This responds directly to a persistent methodological gap in prior literature.

The dissertation also advances beyond purely predictive black-box systems by combining econometric models, artificial intelligence, and explainable AI techniques. Negative Binomial and Logistic Regression models provide interpretability and policy relevance, while Feedforward Neural Networks capture nonlinear interactions. SHAP values and permutation importance were further

applied to identify dominant behavioural and infrastructural predictors. This creates a balanced modelling architecture that integrates explanation, prediction, and operational usefulness.

Finally, the study introduces a Composite Crash Risk Index (CCRI) that synthesizes crash frequency, severity probability, and exposure into a single operational metric for network ranking, spatial mapping, route-risk comparison, and proactive intervention planning. In contrast to traditional crash-count screening methods, the CCRI provides a richer and forward-looking representation of road safety conditions.

### 6.2.Key Findings and Methodological Contribution

The empirical results demonstrate that behavioural indicators extracted from Naturalistic Driving Data significantly improve crash risk estimation beyond conventional crash-based approaches alone. Variables such as speeding, sudden braking, lane departure, and vehicle instability consistently emerged as important determinants of both crash occurrence and severe outcomes. These findings support the central hypothesis of the dissertation that telematics-derived behaviour can serve as a proactive safety signal rather than relying solely on historical crash records.

The proposed framework further showed that artificial intelligence models improved predictive performance, while econometric models preserved interpretability and causal insight. This confirms the value of hybrid modelling strategies in modern transport safety analytics.

For clarity, the principal contributions of this dissertation are summarized as follows:

**Theoretical and Scientific Contributions**

- Development of a unified behavior-driven road safety framework integrating telematics, infrastructure, environment, and exposure within one analytical structure.
- Extension of crash risk theory through the simultaneous treatment of frequency, severity, and exposure rather than isolated modelling streams.
- Validation of Safety-Critical Events as meaningful surrogate indicators of network-level crash risk.
- Expansion of telematics-based safety research into a European context, specifically the Province of Brescia, helping address geographical imbalance in prior studies.

**Methodological Contributions**

**Integrating the data into the following:**

- Introducing  a multi-scale modelling architecture combining segment-level and trip-level risk estimation.
- Integration of interpretable econometric methods with AI models to balance explanation and prediction  integrating .
- Development of the Composite Crash Risk Index (CCRI) for unified ranking, classification, and spatial risk mapping.

**Practical and Policy Contributions**

- Provision of a decision-support tool for proactive safety planning, enabling authorities to prioritize corridors, intersections, and high-risk zones before severe crashes accumulate.
- Support for Vision Zero and Safe System strategies by focusing on preventable risk precursors rather than waiting for crash outcomes.
- Potential application in safer route guidance systems, allowing route comparison not only by travel time or distance, but also by predicted safety risk.
- Transferable and scalable framework that can be recalibrated for future datasets, cities, or connected vehicle ecosystems.

Overall, this dissertation demonstrates that integrating behavioral telematics data with exposure-adjusted statistical and AI models represents a substantial methodological advancement over traditional reactive crash analysis. The research therefore provides both a scientific contribution to transport safety modelling and a practical pathway toward data-driven, proactive road safety management.

## 6.3.Practical Implications for Road Safety and Route Choice

The proposed framework offers several practical implications for road safety management and intelligent transportation systems. At the network level, the composite crash risk index enables transportation authorities to **identify high-risk road segments and prioritize safety interventions** more effectively. By incorporating behavioral indicators into the analysis, the framework supports a

**proactive safety assessment approach**, allowing emerging risk conditions to be identified even in locations where historical crash data are limited.

In addition, the methodology provides a potential foundation for **safety-aware route choice applications**. By integrating crash risk indicators into navigation systems or routing algorithms, it would be possible to evaluate whether a particular road link is relatively safe or potentially hazardous based on observed driving behavior and contextual conditions. Such systems could enable route selection not only according to travel time or distance, but also according to **estimated safety levels**, allowing drivers and mobility platforms to consider safety as an additional decision criterion. This approach could contribute to safer mobility patterns and support data-driven traffic management strategies.

### 6.4.Limitations and Future Research Directions

While the findings of this research provide valuable insights into behavior-driven crash risk estimation, some considerations related to the available data and scope of analysis should be acknowledged. The Naturalistic Driving Data used in this study were derived from **a single day of observations**, which limits the ability to capture longer-term temporal variations in driving behavior, traffic demand, and environmental conditions. In addition, the characteristics of the available dataset may not fully represent the diversity of drivers, vehicles, and contextual factors present across the entire road network.

Future research could further enhance the robustness of the analysis by incorporating **larger and more diverse Naturalistic Driving Study datasets**, including **longer observation periods and additional behavioral indicators**. The integration of richer telematics information, such as **in-vehicle and outward-facing camera data capturing videos or images of driver behavior and surrounding traffic environments**, could provide deeper insights into driver–vehicle–environment interactions and further strengthen the interpretation of Safety-Critical Events.

Importantly, the methodological framework proposed in this study was designed to be **flexible and scalable**, allowing new variables and data sources to be incorporated without altering its fundamental structure. As additional behavioral, environmental, or traffic-related inputs become available, they can be integrated into the modelling framework to improve prediction accuracy and expand analytical capabilities. In this sense, the proposed approach provides a **robust foundation for future research**,

enabling continuous refinement as new telematics technologies and Naturalistic Driving datasets become available.

Overall, this research demonstrates the potential of integrating behavioral telematics data with exposure-adjusted modelling techniques to support **proactive and network-wide road safety assessment**. By combining econometric and artificial intelligence approaches and jointly considering crash frequency and crash severity within a composite risk framework, the study advances the methodological foundations of crash risk estimation. The findings highlight the value of behavior-based safety indicators in identifying hazardous road conditions and provide a promising pathway toward **data-driven, safety-oriented transportation planning and management**.

## REFERENCES


1. AASHTO. (2020). An Overview of Resources, Updates, and Plans for the Next Edition AASHTO Highway Safety Manual Second Edition Updates. (2024). Available at: https://www.highwaysafetymanual.org/Documents/HSM2_Full_Update_Presentation_20240624.pdf [Accessed 25 Feb. 2026].
2. Abdel-Aty, M., & Pande, A. (2005). Identifying crash propensity using specific traffic speed conditions. *Journal of safety Research, 36*(1), 97-108. *https://doi.org/10.1016/j.jsr.2004.11.002*
3. Ahlström, C., Georgoulas, G., & Kircher, K. (2021). Towards a context-dependent multi-buffer driver distraction detection algorithm. *IEEE transactions on intelligent transportation systems, 23*(5), 4778-4790. https://doi.org/10.1109/TITS.2021.3060168
4. Ahlstrom, C., Victor, T., Wege, C., & Steinmetz, E. (2011). Processing of eye/head-tracking data in large-scale naturalistic driving data sets. *IEEE transactions on intelligent transportation systems*, *13*(2), 553-564. https://doi.org/10.1109/TITS.2011.2174786
5. Ahmed, M. M., Khan, M. N., Das, A., & Dadvar, S. E. (2022). Global lessons learned from naturalistic driving studies to advance traffic safety and operation research: A systematic review. *Accident Analysis & Prevention, 167,* 106568. https://doi.org/10.1016/j.aap.2022.106568
6. Ali, E. M., Ahmed, M. M., & Wulff, S. S. (2019). Detection of critical safety events on freeways in clear and rainy weather using SHRP2 naturalistic driving data: Parametric and non-parametric techniques. *Safety Science, 119,* 141-149. https://doi.org/10.1016/j.ssci.2019.01.007
7. Ali, Y., Haque, M. M., & Zheng, Z. (2022). An Extreme Value Theory approach to estimate crash risk during mandatory lane-changing in a connected environment. *Analytic methods in accident research, 33,* 100193. https://doi.org/10.1016/j.amar.2021.100193
8. Antin, J. F., Lee, S., Perez, M. A., Dingus, T. A., Hankey, J. M., & Brach, A. (2019). Second strategic highway research program naturalistic driving study methods. *Safety Science, 119,* 2-10. https://doi.org/10.1016/j.ssci.2019.01.016

*9.* Arbabzadeh, N., & Jafari, M. (2017). A data-driven approach for driving safety risk prediction using driver behavior and roadway information data. *IEEE transactions on intelligent transportation systems, 19*(2), 446-460. https://doi.org/10.1109/TITS.2017.2700869

*10.* Arvin, R., & Khattak, A. J. (2020). Driving impairments and duration of distractions: assessing crash risk by harnessing microscopic naturalistic driving data. *Accident Analysis & Prevention, 146,* 105733. https://doi.org/10.1016/j.aap.2020.105733

*11.* Arvin, R., Kamrani, M., & Khattak, A. J. (2019). The role of pre-crash driving instability in contributing to crash intensity using naturalistic driving data. *Accident Analysis & Prevention, 132,* 105226. https://doi.org/10.1016/j.aap.2019.07.002

12. Axhausen, K. W., Zimmermann, A., Schönfelder, S., Rindsfüser, G., & Haupt, T. (2002). Observing the rhythms of daily life: A six-week travel diary. *Transportation*, *29*(2), 95-124. https://doi.org/10.1023/A:1014247822322

*13.* Azizi, L., & Hadi, M. (2021). Using traffic disturbance metrics to estimate and predict freeway traffic breakdown and safety events. *Transportation research record, 2675(10),* 723-733. https://doi.org/10.1177/03611981211012422

14. Bagdadi, O. (2013). Assessing safety critical braking events in naturalistic driving studies. *Transportation research part F: traffic psychology and behaviour*, *16*, 117-126. https://doi.org/10.1016/j.trf.2012.08.006

15. Bagdadi, O., & Várhelyi, A. (2011). Jerky driving—an indicator of crashproneness?. *CrashAnalysis & Prevention*, *43*(4), 1359-1363. https://doi.org/10.1016/j.aap.2011.02.009

*16.* Bakhit, P. R., Guo, B., & Ishak, S. (2018). Crash and near-crash risk assessment of distracted driving and engagement in secondary tasks: a naturalistic driving study. *Transportation research record, 2672(38),* 245-254. https://doi.org/10.1177/0361198118772703

*17.* Bärgman, J., Lisovskaja, V., Victor, T., Flannagan, C., & Dozza, M. (2015). How does glance behavior influence crash and injury risk? A 'what-if'counterfactual simulation using crashes and near-crashes from SHRP2. *Transportation Research Part F: Traffic Psychology and Behaviour, 35,* 152-169. https://doi.org/10.1016/j.trf.2015.10.011

18. Behboudi, N., Moosavi, S., & Ramnath, R. (2024). Recent advances in traffic accident analysis and prediction: A comprehensive review of machine learning techniques. *arXiv preprint arXiv:2406.13968*. https://doi.org/10.48550/arXiv.2406.13968

19. Berhanu, Y., Schröder, D., Wodajo, B. T., & Alemayehu, E. (2024). Machine learning for predictions of road traffic accidents and spatial network analysis for safe routing on accident and congestion-prone road networks. Results in Engineering, 23, 102737.
20. Bharadwaj, N., Edara, P., & Sun, C. (2019). Risk factors in work zone safety events: a naturalistic driving study analysis. *Transportation research record, 2673*(1), 379-387. https://doi.org/10.1177/0361198118821630
21. Bird, F. E. (1984). *Management guide to loss control*. NOSA Safety Centre (PO Box 26434, Arcadia 0007). https://doi.org/10.12691/ajcea-3-5-5
22. Bugusa, Y., & Patil, S. (2019). An Improved Accident Crash Risk Prediction Model Based On Driving Outcomes Using Ensemble Of Prediction Algorithms. *International Journal of Scientific & Technology Research, 8*(11), 603-611. https://www.researchgate.net/publication/338885545_An_Improved_Accident_Crash_Risk_Prediction_Model_Based_On_Driving_Outcomes_Using_Ensemble_Of_Prediction_Algorithms
23. Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (No. 53). Cambridge university press. https://books.google.com/books?hl=en&lr=&id=d6BTfLBbxjcC&oi=fnd&pg=PR15&dq=Cameron,+A.+C.,+%26+Trivedi,+P.+K.+(1998).+Regression+analysis+of+count+data.+Cambridge+University+Press.+https://doi.org/10.1017/CBO9781139051521&ots=VjviUpqnIa&sig=iMAoE0H8mJ8SIRpkQkXTcV4H37c
24. Castermans, T., Capel, E. H., & Meessen, J. F. Road accident prediction in south-eastern England: the advantages of using connected vehicle data.
25. Das, A., & Ahmed, M. M. (2021). Machine learning approach for predicting lane-change maneuvers using the SHRP2 naturalistic driving study data. *Transportation research record, 2675(9),* 574-594. https://doi.org/10.1177/03611981211003581
26. Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., ... & Knipling, R. R. (2006). The 100-car naturalistic driving study, Phase II-results of the 100-car field experiment. https://rosap.ntl.bts.gov/view/dot/37370
27. Dong, C., Clarke, D. B., Yan, X., Khattak, A., & Huang, B. (2014). Multivariate random-parameters zero-inflated negative binomial regression model: An application to estimate crash frequencies at intersections. *Accident Analysis & Prevention*, *70*, 320-329. https://doi.org/10.1016/j.aap.2014.04.018

*28.* Dozza, M., & González, N. P. (2013). Recognising safety critical events: Can automatic video processing improve naturalistic data analyses?. *Accident Analysis & Prevention, 60,* 298-304. https://doi.org/10.1016/j.aap.2013.02.014

29. Elvik, R., Høye, A., Vaa, T., & Sørensen, M. (2009). *The handbook of road safety measures*. Emerald Group Publishing Limited. https://doi.org/10.1108/9781848552517-012

30. European Transport Safety Council (ETSC). (2019). *Briefing: EU Strategic Action Plan on Road Safety*. [online] Available at: https://etsc.eu/briefing-eu-strategic-action-plan-on-road-safety-2/.

31. Eurostat (2021). *Road accidents: number of fatalities continues falling*. [online] Europa.eu. Available at: https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20210624-1 [Accessed 11 Mar. 2026].

32. Golestani, A., Rezaei, N., Malekpour, M. R., Ahmadi, N., Ataei, S. M. N., Khosravi, S., ... & Farzadfar, F. (2025). Predicting errors in accident hotspots and investigating satiotemporal, weather, and behavioral factors using interpretable machine learning: An analysis of telematics big data. Plos one, 20(7), e0326483.

33. Gong, H., Chen, C., Bialostozky, E., & Lawson, C. T. (2012). A GPS/GIS method for travel mode detection in New York City. *Computers, environment and urban systems*, *36*(2), 131-139. https://doi.org/10.1016/j.compenvurbsys.2011.05.003

34. Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1, No. 2, pp. 1-800). Cambridge: MIT press. https://synapse.koreamed.org/pdf/10.4258/hir.2016.22.4.351

35. Guo, F., Zheng, X., Guo, M., Chen, Y., Han, C., & Li, J. (2025). Assessing and interpreting driving risks through trajectory data analysis across vehicle types. *Transportmetrica A: Transport Science*, 1-30. https://doi.org/10.1080/23249935.2025.2552245

*36.* Habibovic, A., Tivesten, E., Uchida, N., Bärgman, J., & Aust, M. L. (2013). Driver behavior in car-to-pedestrian incidents: An application of the Driving Reliability and Error Analysis Method (DREAM). *Accident Analysis & Prevention, 50,* 554-565. https://doi.org/10.1016/j.aap.2012.05.034

37. Hashmi, F., Hassan, M. U., Zubair, M. U., Ahmed, K., Aziz, T., & Choudhry, R. M. (2024). Near-miss detection metrics: an approach to enable sensing technologies for proactive construction safety management. *Buildings*, *14*(4), 1005. https://doi.org/10.3390/buildings14041005

38. Hauer, E. (1997). *Observational before/after studies in road safety. Estimating the effect of highway and traffic engineering measures on road safety*. https://trid.trb.org/View/1136089

39. Hauer, E. (2001). Overdispersion in modelling crashs on road sections and in Empirical Bayes estimation. *CrashAnalysis & Prevention*, *33*(6), 799-808. https://doi.org/10.1016/S0001-4575(00)00094-4
40. Heinrich, H. W. (1941). Industrial CrashPrevention. A Scientific Approach. https://www.cabidigitallibrary.org/doi/full/10.5555/19432701767
41. Huang, T., Wang, S., & Sharma, A. (2020). Highway crash detection and risk estimation using deep learning. *Accident Analysis & Prevention, 135,* 105392. https://doi.org/10.1016/j.aap.2019.105392
42. Hussain, F., Li, Y., Arun, A., & Haque, M. M. (2022). A hybrid modelling framework of machine learning and extreme value theory for crash risk estimation using traffic conflicts. *Analytic methods in accident research, 36,* 100248. https://doi.org/10.1016/j.amar.2022.100248
43. Istat.it. (2022). *Incidenti stradali in Italia 2022*. [online] Available at: https://www.istat.it/comunicato-stampa/incidenti-stradali-in-italia-anno-2022/ [Accessed 24 Feb. 2026].
44. Jovanis, P. P., Aguero-Valverde, J., Wu, K. F., & Shankar, V. (2011). Analysis of naturalistic driving event data: Omitted-variable bias and multilevel modeling approaches. *Transportation research record, 2236*(1), 49-57. https://doi.org/10.3141/2236-06
45. Kashkanova, A. (2025). Methods for assessing the risk of crashs in the driver-vehicle-road-environment system. *Archiwum Motoryzacji*, *109*(3), 41-61. https://doi.org/10.14669/AM/207102
46. Khan, M. S., Nyamagoud, B., Malipatil, S., & Meti, A. (2026). Artificial Intelligence for Road Safety Analytics: Crash Risk, Near-Miss Detection, and Proactive Interventions. Deep Science Publishing.
47. Khattak, Z. H., Rios-Torres, J., Fontaine, M. D., & Khattak, A. J. (2023). Inferring safety critical events from vehicle kinematics in naturalistic driving environment: Application of deep learning Algorithms. *Journal of Intelligent Transportation Systems*, *27*(4), 423-440. https://doi.org/10.1080/15472450.2022.2048655
48. Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data. https://rosap.ntl.bts.gov/view/dot/62931

49. Kluger, R., Smith, B. L., Park, H., & Dailey, D. J. (2016). Identification of safety-critical events using kinematic vehicle data and the discrete fourier transform. *Accident Analysis & Prevention*, *96*, 162-168. https://doi.org/10.1016/j.aap.2016.08.006
50. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444. https://doi.org/10.1038/nature14539
51. Li, Y., Dalhatu, S., & Yuan, C. (2024). Analysing freeway diverging risks using high-resolution trajectory data based on conflict prediction models. *Transportation safety and environment*, *6*(1), tdad002. https://doi.org/10.1093/tse/tdad002
52. Liang, Shixiao, Chengyuan Ma, Pei Li, Haotian Shi, Jiaxi Liu, Hang Zhou, Keke Long, Bofeng Cao, Todd Szymkowski, and Xiaopeng Li. "Real-Time Lane-Level Crash Detection on Freeways Using Sparse Telematics Data." arXiv preprint arXiv:2511.18148 (2025).
53. Lipar, P., Lakner, M., Maher, T., & Žura, M. (2011). Estimation of road centerline curvature from raw GPS data. *The Baltic Journal of Road and Bridge Engineering*, *6*(3), 163-168. https://bjrbe-journals.rtu.lv/article/view/bjrbe.2011.21
54. Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation research part A: policy and practice*, *44*(5), 291-305. https://doi.org/10.1016/j.tra.2010.02.001
55. Mannering, F. L., & Bhat, C. R. (2014). Analytic methods in crashresearch: Methodological frontier and future directions. *Analytic methods in crashresearch*, *1*, 1-22. https://doi.org/10.1016/j.amar.2013.09.001
56. Mannering, F. L., Shankar, V., & Bhat, C. R. (2016). Unobserved heterogeneity and the statistical analysis of highway crashdata. *Analytic methods in crashresearch*, *11*, 1-16. https://doi.org/10.1016/j.amar.2016.04.001
57. Mannering, F., Bhat, C. R., Shankar, V., & Abdel-Aty, M. (2020). Big data, traditional data and the tradeoffs between prediction and causality in highway-safety analysis. *Analytic methods in accident research*, *25*, 100113. https://doi.org/10.1016/j.amar.2020.100113
58. Miaou, S. P. (1994). The relationship between truck crashs and geometric design of road sections: Poisson versus negative binomial regressions. *CrashAnalysis & Prevention*, *26*(4), 471-482. https://doi.org/10.1016/0001-4575(94)90038-8
59. Moosavi, S., & Ramnath, R. (2023). Context-aware driver risk prediction with telematics data. Accident Analysis & Prevention, 192, 107269.

60. Moradi, S. (2022). Real-time crash risk analysis using deep learning. https://www.researchgate.net/profile/Saeid_Moradi9/publication/361101213_REAL-TIME_CRASH_RISK_ANALYSIS_USING_DEEP_LEARNING_Saeid_MORADI_Department_of_Civil_Engineering_Transport_Programme/links/629cb4f1416ec50bdb0e093f/REAL-TIME-CRASH-RISK-ANALYSIS-USING-DEEP-LEARNING-Saeid-MORADI-Department-of-Civil-Engineering-Transport-Programme.pdf
61. Naji, H. A., Xue, Q., Lyu, N., Duan, X., & Li, T. (2022). Risk levels classification of near-crashes in naturalistic driving data. *Sustainability*, *14*(10), 6032. https://doi.org/10.3390/su14106032
62. Nguyen, H. M. V., Oh, S. C., Saha, G., Blunt, T. L., Ball, M. J., Duffy, M. E., ... & Park, B. B. (2024, May). An analysis of vehicular telematics and geo-location data to maximize road safety. In *2024 Systems and Information Engineering Design Symposium (SIEDS)* (pp. 136-141). IEEE. https://doi.org/10.1109/SIEDS61124.2024.10534635
63. Osman, O. A., & Hajij, M. (2021). Development and comparative analysis of advanced deep learning techniques for crash prediction in advanced driver support systems. Transportation research record, 2675(12), 730-740.
64. Paefgen, J., Staake, T., & Thiesse, F. (2013). Evaluation and aggregation of pay-as-you-drive insurance rate factors: A classification analysis approach. *Decision Support Systems*, *56*, 192-201. https://doi.org/10.1016/j.dss.2013.06.001
65. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *bmj*, *372*. https://doi.org/10.1136/bmj.n71
66. Paramasivan, K., Subburaj, R., Sharma, V. M., & Sudarsanam, N. (2022). Relationship between mobility and road traffic injuries during COVID-19 pandemic—The role of attendant factors. *PloS one*, *17*(5), e0268190. https://doi.org/10.1371/journal.pone.0268190
67. Passmore, J., Yon, Y., & Mikkelsen, B. (2019). Progress in reducing road-traffic injuries in the WHO European region. *The Lancet Public Health*, *4*(6), e272-e273. https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(19)30074-X/fulltext
68. Pei, X., Wong, S. C., & Sze, N. N. (2012). The roles of exposure and speed in road safety analysis. *Crashanalysis & prevention*, *48*, 464-471. https://doi.org/10.1016/j.aap.2012.03.005
69. Prajapati, G., Kumar, L., & Patil, S. R. S. (2023). Road accident prediction using machine learning. *Journal of Scientific Research and Technology*, 48-59. https://doi.org/10.5281/zenodo.7961680

70. Precht, L., Keinath, A., & Krems, J. F. (2017). Effects of driving anger on driver behavior–Results from naturalistic driving data. *Transportation research part F: traffic psychology and behaviour*, *45*, 75-92. https://doi.org/10.1016/j.trf.2016.10.019
71. Razi, A., Chen, X., Li, H., Wang, H., Russo, B., Chen, Y., & Yu, H. (2023). Deep learning serves traffic safety analysis: A forward-looking review. *IET Intelligent Transport Systems*, *17*(1), 22-71. https://doi.org/10.1049/itr2.12257
72. Ronan (2020). *Road Safety Annual Report 2020*. [online] ITF. Available at: https://www.itf-oecd.org/road-safety-annual-report-2020.
73. Roussou, S., Garefalakis, T., Michelaraki, E., Brijs, T., & Yannis, G. (2024). Machine learning insights on driving behaviour dynamics among Germany, Belgium, and UK drivers. Sustainability, 16(2), 518.
74. Safer (2019). *Safer roads: what are the EU measures to reduce car accidents? | Topics | European Parliament.* [online] Topics | European Parliament. Available at: https://www.europarl.europa.eu/topics/en/article/20190307STO30715/safer-roads-what-are-the-eu-measures-to-reduce-car-accidents.
75. Savolainen, P. T., Mannering, F. L., Lord, D., & Quddus, M. A. (2011). The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. *CrashAnalysis & Prevention*, *43*(5), 1666-1676. https://doi.org/10.1016/j.aap.2011.03.025
76. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, *61*, 85-117. https://doi.org/10.1016/j.neunet.2014.09.003
77. Shaon, M. R. R., Qin, X., Afghari, A. P., Washington, S., & Haque, M. M. (2019). Incorporating behavioral variables into crash count prediction by severity: A multivariate multiple risk source approach. *Accident Analysis & Prevention*, *129*, 277-288. https://doi.org/10.1016/j.aap.2019.05.010
78. Sharmin, S., Ivan, J. N., Marsh, K. L., Paxton, A., & Tucker, A. (2023). Driver psychology latent classes as predictors of traffic incident occurrence in naturalistic driving study data. *Transportation research record*, *2677*(2), 839-857. https://doi.org/10.1177/03611981221108985
79. Shen, J., & Yang, G. (2020). Crash risk assessment for heterogeneity traffic and different vehicle-following patterns using microscopic traffic flow data. *Sustainability*, *12*(23), 9888. https://doi.org/10.3390/su12239888

80. Shichrur, R., & Ratzon, N. Z. (2023). Optimal duration of in-vehicle data recorder monitoring to assess bus driver behavior. *Sensors*, *23*(21), 8887. https://doi.org/10.3390/s23218887
81. Silva, P. B., Andrade, M., & Ferreira, S. (2020). Machine learning applied to road safety modeling: A systematic literature review. *Journal of traffic and transportation engineering (English edition)*, *7*(6), 775-790. https://doi.org/10.1016/j.jtte.2020.07.004
82. Skaug, L., Nojoumian, M., Dang, N., & Yap, A. (2025). Road crash analysis and modeling: A systematic review of methods, data, and emerging technologies. *Applied Sciences*, *15*(13), 7115. https://doi.org/10.3390/app15137115
83. Somvanshi, S., Liu, J., Chakraborty, R., Tamakloe, R., & Das, S. (2026). Predicting Crash Severity using Naturalistic Driving Data and Neural Networks. International Journal of Intelligent Transportation Systems Research, 1-18.
84. Stopher, P., FitzGerald, C., & Xu, M. (2007). Assessing the accuracy of the Sydney Household Travel Survey with GPS. *Transportation*, *34*(6), 723-741. https://doi.org/10.1007/s11116-007-9126-8
85. Theofilatos, A., Chen, C., & Antoniou, C. (2019). Comparing machine learning and deep learning methods for real-time crash prediction. *Transportation research record*, *2673*(8), 169-178. https://doi.org/10.1177/0361198119841571
86. Theofilatos, A., Chen, C., & Antoniou, C. (2019). Comparing machine learning and deep learning methods for real-time crash prediction. *Transportation research record*, *2673*(8), 169-178. https://doi.org/10.1177/0361198119841571
87. Tian, R., Li, L., Chen, M., Chen, Y., & Witt, G. J. (2013). Studying the effects of driver distraction and traffic density on the probability of crash and near-crash events in naturalistic driving environment. *IEEE Transactions on Intelligent Transportation Systems*, *14*(3), 1547-1555. https://doi.org/10.1109/TITS.2013.2261988
88. Tivesten, E., & Dozza, M. (2014). Driving context and visual-manual phone tasks influence glance behavior in naturalistic driving. *Transportation research part F: traffic psychology and behaviour*, *26*, 258-272. https://doi.org/10.1016/j.trf.2014.08.004
89. Wali, B., & Khattak, A. J. (2020). Harnessing ambient sensing & naturalistic driving systems to understand links between driving volatility and crash propensity in school zones–A generalized hierarchical mixed logit framework. *Transportation research part C: emerging technologies*, *114*, 405-424. https://doi.org/10.1016/j.trc.2020.01.028

*90.* Wang, Z., Ramyar, S., Salaken, S. M., Homaifar, A., Nahavandi, S., & Karimoddini, A. (2017, June). A collision avoidance system with fuzzy danger level detection. In *2017 IEEE intelligent vehicles symposium (IV)* (pp. 283-288). IEEE. https://doi.org/10.1109/IVS.2017.7995733

91. Wijnen, W., & Stipdonk, H. (2016). Social costs of road crashes: An international analysis. *Accident Analysis & Prevention*, *94*, 97-106. https://doi.org/10.1016/j.aap.2016.05.005

92. Winlaw, M., Steiner, S. H., MacKay, R. J., & Hilal, A. R. (2019). Using telematics data to find risky driver behaviour. *Accident Analysis & Prevention*, *131*, 131-136. https://doi.org/10.1016/j.aap.2019.06.003

*93.* Winlaw, M., Steiner, S. H., MacKay, R. J., & Hilal, A. R. (2019). Using telematics data to find risky driver behaviour. *CrashAnalysis & Prevention*, *131*, 131-136. https://doi.org/10.1016/j.aap.2019.06.003

*94.* World Health Organization (WHO). (2019). *Global status report on road safety 2018*. World Health Organization. https://books.google.com/books?hl=en&lr=&id=uHOyDwAAQBAJ&oi=fnd&pg=PR6&dq=Ronan,+C.+(2020).+Global+status+report+on+road+safety+2018.+World+Health+Organization.+https://www.who.int/publications/i/item/9789241565684&ots=2T5hWwxcYX&sig=JzHV4jOUO7sXckIkNsyjEokErME

95. World Health Organization (WHO). (2023). *Road traffic injuries*. [online] World Health Organization. Available at: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.

*96.* Wu, K. F. K., & Wang, L. (2021). Exploring the combined effects of driving situations on freeway rear-end crash risk using naturalistic driving study data. *Accident Analysis & Prevention, 150,* 105866. https://doi.org/10.1016/j.aap.2020.105866

*97.* Wu, K. F., & Jovanis, P. P. (2013). Defining and screening crash surrogate events using naturalistic driving data. *Accident Analysis & Prevention, 61,* 10-22. https://doi.org/10.1016/j.aap.2012.10.004

98. Wu, K. F., & Thor, C. P. (2015). Method for the use of naturalistic driving study data to analyze rear-end crash sequences. *Transportation Research Record*, *2518*(1), 27-36. https://doi.org/10.3141/2518-04

*99.* Xu, D. (2021). *Innovative Applications of Naturalistic Driving Data for In-Depth Analysis of Freeway Interchange Deceleration Lane Design and Work Zone Mobility* (Doctoral dissertation, Auburn University).

https://search.proquest.com/openview/9dfb072180a3a8f8a74ebfba7cc77daf/1?pq-origsite=gscholar&cbl=18750&diss=y

100. Yedilkhan, D., Agybetov, N., & Amirgaliyev, B. (2025). Driver Behaviour Analysis Using Telematics Sensor Data and Deep Learning Models. Procedia Computer Science, 272, 594-600.

101. Yeo, J., Lee, J., & Jang, K. (2021). The effects of rainfall on driving behaviors based on driving volatility. *International journal of sustainable transportation*, *15*(6), 435-443. https://doi.org/10.1080/15568318.2020.1756543

*102.* Yu, R., Han, L., & Zhang, H. (2021). Trajectory data-based freeway high-risk events prediction and its influencing factors analyses. *Accident Analysis & Prevention, 154,* 106085. https://doi.org/10.1016/j.aap.2021.106085

103. Yu, R., Han, L., Abdel-Aty, M., Wang, L., & Zou, Z. (2024). Improving model robustness of traffic crash risk evaluation via adversarial mix-up under traffic flow fundamental diagram. *Accident Analysis & Prevention*, *194*, 107360. https://doi.org/10.1016/j.aap.2023.107360

104. Zeng, Q., Gu, W., Zhang, X., Wen, H., Lee, J., & Hao, W. (2019). Analyzing freeway crash severity using a Bayesian spatial generalized ordered logit model with conditional autoregressive priors. *CrashAnalysis & Prevention*, *127*, 87-95. https://doi.org/10.1016/j.aap.2019.02.029

105. Zeng, Y., Qiang, Y., Zhang, N., Yang, X., Zhao, Z., & Wang, X. (2024). An influencing factors analysis of road traffic accidents based on the analytic hierarchy process and the minimum discrimination information principle. *Sustainability*, *16*(16), 6767. https://doi.org/10.3390/su16166767

*106.* Zhang, C., He, J., King, M., Liu, Z., Chen, Y., Yan, X., ... & Zhang, H. (2021). A crash risk identification method for freeway segments with horizontal curvature based on real-time vehicle kinetic response. *Accident Analysis & Prevention, 150,* 105911. https://doi.org/10.1016/j.aap.2020.105911

*107.* Zhang, H., & Khattak, A. (2011). Spatiotemporal patterns of primary and secondary incidents on urban freeways. *Transportation research record, 2229*(1), 19-27. https://doi.org/10.3141/2229-03

108. Zheng, L., Ismail, K., & Meng, X. (2014). Traffic conflict techniques for road safety analysis: open questions and some insights. *Canadian journal of civil engineering*, *41*(7), 633-641. https://doi.org/10.1139/cjce-2013-0558

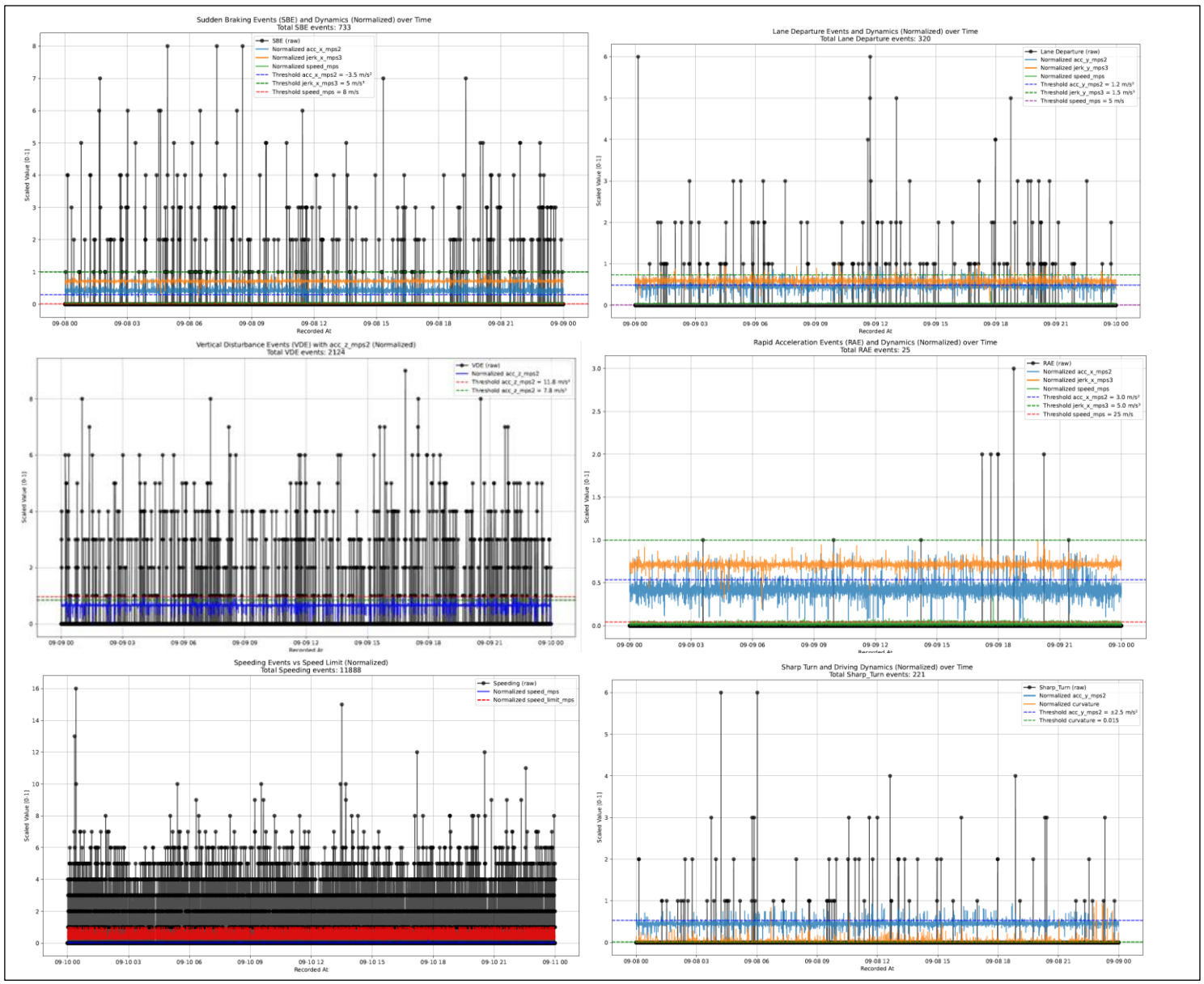
*109.* Zheng, Q., Xu, C., Liu, P., & Wang, Y. (2021). Investigating the predictability of crashes on different freeway segments using the real-time crash risk models. *Accident Analysis & Prevention, 159,* 106213. https://doi.org/10.1016/j.aap.2021.106213

110. Zheng, Z., Ahn, S., Chen, D., & Laval, J. (2011). Applications of wavelet transform for analysis of freeway traffic: Bottlenecks, transient traffic, and traffic oscillations. *Transportation Research Part B: Methodological*, *45*(2), 372-384. https://doi.org/10.1016/j.trb.2010.08.002

111. Zurlinden, H., Baruah, A., & Gaffney, J. (2020). Towards linking driving complexity to crash risk. *Journal of road safety, 31*(1), 66-80. https://search.informit.org/doi/abs/10.3316/informit.072631340590916

## APPENDICES

**Appendix A:** detailed time-series visualizations illustrating the threshold-based detection process

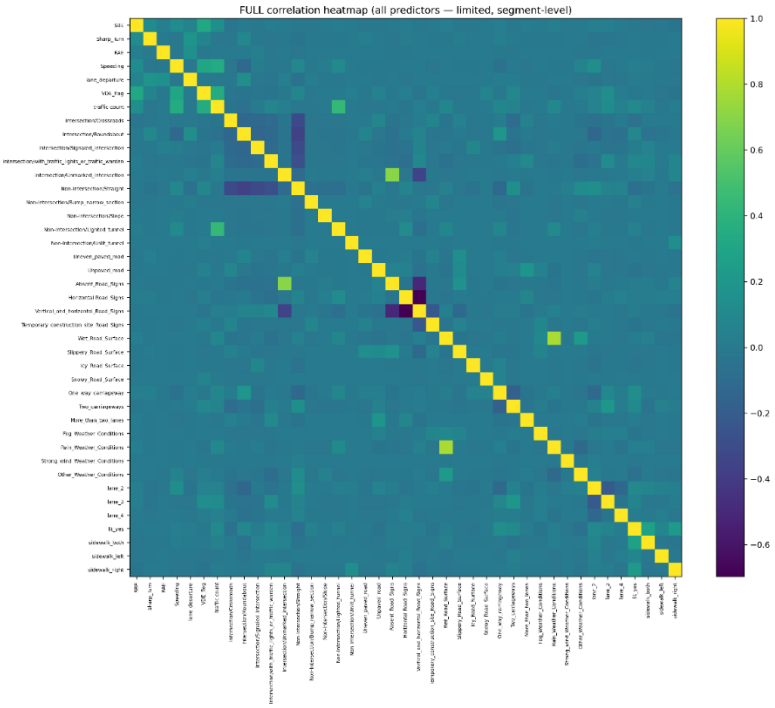


## Appendix B :Categorized Explanatory Variables Table

| *Category* | *Variable* | *Short Explanation* |
|---|---|---|
| ***Driving Behaviour*** | Speeding | Indicator of driving above the posted speed limit derived from telematics observations |
| | SBE | Sudden Braking Event detected from vehicle acceleration patterns |
| | RAE | Rapid Acceleration Event indicating aggressive acceleration behavior |
| | Sharp_Turn | Sudden or aggressive turning maneuver indicating unstable vehicle control |
| | lane_departure | Vehicle deviates from the current lane unexpectedly |
| | VDE_flag | Vertical Dynamic Event representing instability due to vertical road disturbances |
| ***Road Geometry*** | lane_2 | Two-lane roadway segment |
| | lane_3 | Three-lane roadway segment |
| | lane_4 | Four-lane roadway segment |
| | More_than_two_lanes | Road segment containing more than two travel lanes |
| ***Road Infrastructure*** | Intersection_Traffic_Light | Intersection controlled by traffic lights |
| | Intersection_Unmarked | Intersection without traffic control devices |
| | Non-Intersection_Straight | Straight road segment not located at an intersection |
| | Non-Intersection_Tunnel | Road segment located within a tunnel |
| | Non-Intersection_Bump_narrow_section | Segment containing traffic calming bumps or narrow sections |
| | Vertical_and_horizontal_Road_Signs | Presence of road signs guiding traffic movement |
| | Absent_Road_Signs | Lack of traffic regulatory or warning signs |
| | Traffic_Control_Present | Presence of infrastructure regulating vehicle movement |
| | Traffic_Control_Absent | Lack of traffic control infrastructure |
| ***Traffic / Exposure*** | traffic_count | Traffic flow volume representing vehicle exposure on the segment |
| | AADT_proxy | Proxy for Annual Average Daily Traffic representing traffic exposure |
| ***Environmental / Weather*** | rain_Weather_Conditions | Rain conditions affecting roadway safety |
| | wind_Weather_Conditions | Wind conditions potentially influencing vehicle stability |
| | fog_Weather_Conditions | Fog reducing driver visibility |
| | cloudy_Weather_Conditions | Cloudy weather condition |
| | clear_Weather_Conditions | Clear weather conditions |
| ***Lighting Conditions*** | lit_yes | Road segment with artificial lighting present |

| | | |
|---|---|---|
| | lit_no | Road segment without lighting infrastructure |
| ***Road Surface Condition*** | Dry_Road_Surface | Dry pavement condition |
| | Wet_Road_Surface | Wet pavement due to precipitation |
| | Snowy_Road_Surface | Snow-covered road surface |
| | Icy_Road_Surface | Presence of ice on pavement surface |
| | Slippery_Road_Surface | Reduced friction surface condition |
| | Uneven_Road_Surface | Road irregularities affecting vehicle stability |
| ***Driver Characteristics*** | Driver_Age_A | Age of the driver of vehicle A involved in the event |
| | Driver_Age_B | Age of the driver of vehicle B involved in the event |
| | Young_Driver | Indicator for young driver category |
| | Elderly_Driver | Indicator for elderly driver category |
| ***Vehicle Characteristics*** | Passenger_car_A | Passenger car involved in the crash |
| | Passenger_car_B | Passenger car involved as second vehicle |
| | Motorcycle_A | Motorcycle involved in the event |
| | Motorcycle_with_passenger_A | Motorcycle carrying passenger |
| | Bus_A | Bus involved in the crash |
| | Truck_A | Heavy truck vehicle involved |
| | Emergency_or_police_vehicle_B | Emergency service vehicle involved |
| | Special_vehicle_A | Special purpose vehicle category |
| | Commercial_vehicle | Vehicle used for commercial transport |
| ***Crash Type / Accident Characteristics*** | Nature_of_Accident_Pedestrian_collision | Crash involving a pedestrian |
| | Nature_of_Accident_Collision_with_vehicle | Collision between two moving vehicles |
| | Nature_of_Accident_Collision_with_parked_vehicle | Crash involving a parked vehicle |
| | Nature_of_Accident_Side_collision | Side impact collision |
| | Nature_of_Accident_Rear_end | Rear-end collision |
| | Nature_of_Accident_Head_on | Head-on vehicle collision |
| | Nature_of_Accident_Single_vehicle | Crash involving only one vehicle |
| | Nature_of_Accident_Loss_of_control | Crash caused by loss of vehicle control |

## Appendix C (Figures C1–C2 and Tables C1-C2).




**Figures C1.** Pairwise Correlation Heatmap of Segment-Level Behavioral Predictors

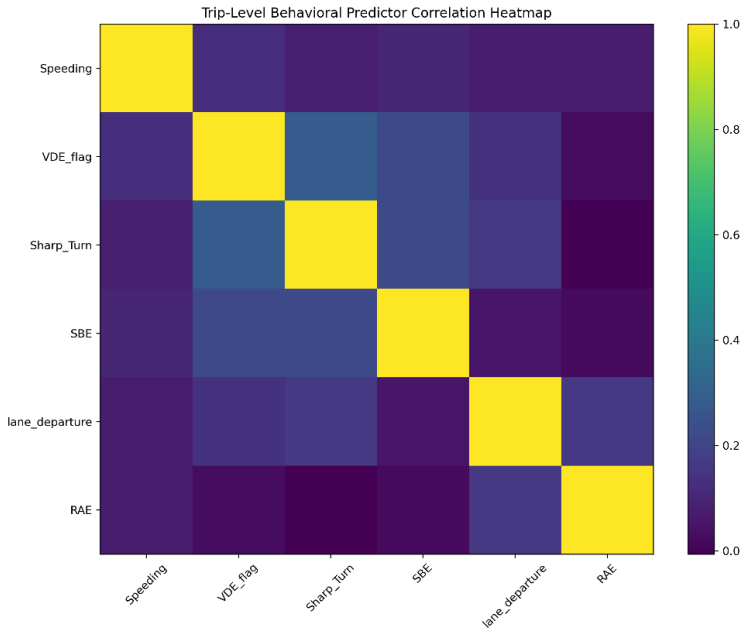



**Figure C2.** Pairwise Correlation Heatmap of Trip-Level Behavioral Predictors

*Table C1: VIF Values for Predictors Included in the Segment-Level Models*

| variable | VIF |
|---|---|
| Non-Intersection/Straight | 6.606619 |
| Wet_Road_Surface | 4.348872 |
| Rain_Weather_Conditions | 3.944642 |
| traffic count | 3.094659 |
| Intersection/Roundabout | 2.725849 |
| Absent_Road_Signs | 2.393501 |
| Intersection/Signaled_intersection | 2.244654 |
| Intersection/Crossroads | 2.230503 |
| Intersection/Unmarked_intersection | 2.189562 |
| Intersection/with_traffic_lights_or_traffic_warden | 2.133719 |
| lane_2 | 2.062852 |
| lit_yes | 1.654797 |
| Horizontal Road_Signs | 1.621807 |
| Other_Weather_Conditions | 1.543287 |
| VDE_flag | 1.529613 |
| Speeding | 1.519718 |
| Non-Intersection/Lighted_tunnel | 1.425141 |
| One_way_carriageway | 1.416756 |
| Two_carriageways | 1.3137 |
| lane_3 | 1.284739 |
| SBE | 1.206135 |
| sidewalk_both | 1.179533 |
| sidewalk_right | 1.130559 |
| Slippery_Road_Surface | 1.128461 |
| lane_departure | 1.105338 |
| Unpaved_road | 1.105329 |
| Fog_Weather_Conditions | 1.095741 |
| Temporary_construction_site_Road_Signs | 1.085891 |
| Sharp_Turn | 1.080736 |
| Non-Intersection/Slope | 1.080299 |
| More_than_two_lanes | 1.077911 |
| lane_4 | 1.071756 |
| Non-Intersection/Bump_narrow_section | 1.066003 |
| Uneven_paved_road | 1.046134 |
| RAE | 1.032496 |

| | |
|---|---|
| Non-Intersection/Unlit_tunnel | 1.028044 |
| sidewalk_left | 1.02488 |
| Icy_Road_Surface | 1.019387 |
| Snowy_Road_Surface | 1.012568 |
| Strong_wind_Weather_Conditions | 1.003959 |

***Tables C2.*** *VIF Values for Predictors Included in Trip-Level Models*

| **Variable** | **VIF** |
|---|---|
| VDE_flag | 1.141745 |
| Sharp_Turn | 1.139859 |
| SBE | 1.083317 |
| lane_departure | 1.065889 |
| RAE | 1.032205 |
| Speeding | 1.030032 |

## Appendix D (Tables D1–D2).

**Table D1**: complete estimation results for the NB models at segment level

```
STEP 1 — FIT FULL-DATA NB MODEL
============================================================
                 Generalized Linear Model Regression Results
==============================================================================
Dep. Variable:            crash_count   No. Observations:                 2515
Model:                            GLM   Df Residuals:                     2485
Model Family:        NegativeBinomial   Df Model:                           29
Link Function:                    Log   Scale:                          1.0000
Method:                          lbfgs   Log-Likelihood:                -4049.1
Date:                Fri, 06 Mar 2026   Deviance:                       214.34
Time:                        23:07:06   Pearson chi2:                     311.
No. Iterations:                     7   Pseudo R-squ. (CS):           0.06483
Covariance Type:            nonrobust
```

| **Category** | **Variable** | **Coefficient** | **Std. Error** | **z-value** | **P-value** |
|---|---|---|---|---|---|
| Constant | Intercept | -1.782**** | 0.065 | -27.382 | 0.000 |
| Road Infrastructure | Intersection/Roundabout | 0.019 | 0.099 | 0.192 | 0.847 |
| Road Infrastructure | Intersection/Signaled intersection | 0.083 | 0.103 | 0.804 | 0.421 |
| Road Infrastructure | Intersection with traffic lights/warden | 0.283*** | 0.109 | 2.593 | 0.010 |
| Road Infrastructure | Intersection/Unmarked intersection | 0.367 | 0.343 | 1.070 | 0.285 |
| Road Geometry | Non-Intersection/Straight | 0.029 | 0.075 | 0.380 | 0.704 |
| Road Geometry | Bump / Narrow section | -0.124 | 0.445 | -0.278 | 0.781 |
| Road Geometry | Slope | 0.020 | 0.349 | 0.056 | 0.955 |
| Road Infrastructure | Lighted tunnel | 1.257**** | 0.280 | 4.481 | 0.000 |

| | | | | | |
|---|---|---|---|---|---|
| Road Condition | Uneven paved road | -0.088 | 0.289 | -0.306 | 0.760 |
| Road Condition | Unpaved road | 0.004 | 0.259 | 0.015 | 0.988 |
| Road Infrastructure | Absent road signs | -0.112 | 0.251 | -0.446 | 0.656 |
| Road Infrastructure | Horizontal road signs | -0.142 | 0.130 | -1.090 | 0.276 |
| Road Infrastructure | Temporary construction signs | 0.035 | 0.371 | 0.093 | 0.926 |
| Environment | Slippery road surface | 0.065 | 0.327 | 0.200 | 0.842 |
| Road Geometry | Two carriageways | 0.053 | 0.086 | 0.613 | 0.540 |
| Road Geometry | More than two lanes | -0.128 | 0.352 | -0.363 | 0.717 |
| Environment | Fog weather conditions | 0.250 | 0.395 | 0.632 | 0.527 |
| Environment | Other weather conditions | -0.008 | 0.090 | -0.089 | 0.929 |
| Road Geometry | Lane = 3 | 0.273** | 0.112 | 2.433 | 0.015 |
| Road Geometry | Lane = 4 | 0.268 | 0.182 | 1.474 | 0.140 |
| Environment | Lighting present | 0.120* | 0.065 | 1.850 | 0.064 |
| Infrastructure | Sidewalk left | -0.147 | 0.383 | -0.383 | 0.702 |
| Infrastructure | Sidewalk right | 0.035 | 0.183 | 0.192 | 0.848 |
| Driving Behaviour | Sudden Braking Events (SBE) | 0.013 | 0.024 | 0.559 | 0.576 |
| Driving Behaviour | Sharp Turn | -0.018 | 0.028 | -0.664 | 0.507 |
| Driving Behaviour | Rapid Acceleration Event (RAE) | 0.006 | 0.026 | 0.234 | 0.815 |
| Driving Behaviour | Speeding | 0.118**** | 0.025 | 4.652 | 0.000 |
| Driving Behaviour | Lane departure | 0.001 | 0.028 | 0.017 | 0.986 |
| Driving Behaviour | Vertical Dynamic Event (VDE) | 0.123**** | 0.024 | 5.043 | 0.000 |

**Tables D1:** complete estimation results for the econometric crash frequency models at Trip level

```
Generalized Linear Model Regression Results
==============================================================================
Dep. Variable:            crash_count   No. Observations:                 2157
Model:                            GLM   Df Residuals:                     2150
Model Family:        NegativeBinomial   Df Model:                            6
Link Function:                    Log   Scale:                          1.0000
Method:                          lbfgs   Log-Likelihood:                -5219.5
Date:                Sun, 08 Mar 2026   Deviance:                       893.21
Time:                        12:19:22   Pearson chi2:                 1.11e+03
No. Iterations:                    12   Pseudo R-squ. (CS):            0.1619
Covariance Type:            nonrobust
```

| Category | Variable | Coefficient | Std. Error | z-value | P-value |
|---|---|---|---|---|---|
| Constant | Intercept | -1.041**** | 0.035 | -30.146 | 0.000 |
| Driving Behaviour | Sudden Braking Event (SBE) | 0.065** | 0.032 | 2.024 | 0.043 |
| Driving Behaviour | Sharp Turn | 0.031 | 0.128 | 0.241 | 0.809 |
| Driving Behaviour | Lane Departure | 0.148 | 0.097 | 1.525 | 0.127 |
| Driving Behaviour | Speeding | 0.119**** | 0.008 | 15.101 | 0.000 |
| Driving Behaviour | Vertical Dynamic Event (VDE) | 0.028** | 0.011 | 2.407 | 0.016 |
| Driving Behaviour | Rapid Acceleration Event (RAE) | 0.302 | 0.299 | 1.010 | 0.312 |

## Appendix E (Tables E1–E2).

**Tables E1:** Complete estimation results for the Logit model at Segment level

```
===== FULL DATA PERFORMANCE =====
[[1968   30]
 [ 417  100]]
```

```
              precision    recall  f1-score   support

           0     0.8252    0.9850    0.8980      1998
           1     0.7692    0.1934    0.3091       517

    accuracy                         0.8223      2515
   macro avg     0.7972    0.5892    0.6036      2515
weighted avg     0.8137    0.8223    0.7770      2515

AUC: 0.8021459564012755
Brier: 0.12621965366992566
```

**COEFFICIENTS:**

| Variable | Coefficient | Odds Ratio | P-value | Significance |
|---|---|---|---|---|
| **Speeding** | 0.556 | 1.743 | <0.001 | **** |
| **Frontal-lateral collision** | 0.534 | 1.706 | 0.029 | * |
| **Number of other vehicles involved** | 0.498 | 1.645 | <0.001 | **** |
| **Non-Intersection / Straight** | 0.392 | 1.480 | 0.001 | *** |
| **Side collision** | 0.350 | 1.419 | 0.042 | * |
| **Lane = 3** | 0.333 | 1.394 | <0.001 | **** |
| **Vertical Dynamic Event (VDE)** | 0.332 | 1.394 | <0.001 | **** |
| **Lane = 2** | 0.311 | 1.365 | <0.001 | **** |
| **Head-on collision** | 0.304 | 1.355 | 0.022 | * |
| **Vertical & horizontal road signs** | 0.298 | 1.348 | 0.002 | *** |
| **Intersection with traffic lights** | 0.286 | 1.331 | <0.001 | **** |
| **Driver Age B** | -0.259 | 0.772 | 0.005 | *** |
| **Unmarked intersection** | 0.225 | 1.252 | 0.033 | * |
| **Lighted tunnel** | 0.173 | 1.189 | 0.002 | *** |
| **Lane = 4** | 0.172 | 1.188 | <0.001 | **** |
| **Non-Intersection slope** | 0.147 | 1.158 | 0.006 | ** |
| **Snowy road surface** | 0.126 | 1.134 | 0.010 | ** |

**Tables E2:** Complete estimation results for the Logit model at Trip Level

```
===== FULL DATA PERFORMANCE =====
[[594 139]
 [479 945]]
              precision    recall  f1-score   support

           0     0.5536    0.8104    0.6578       733
           1     0.8718    0.6636    0.7536      1424

    accuracy                         0.7135      2157
   macro avg     0.7127    0.7370    0.7057      2157
weighted avg     0.7636    0.7135    0.7210      2157

AUC: 0.8040452503947146
Brier: 0.16590610279712328

 COEFFICIENTS:
```

| Variable | Coefficient | Std. Error | z-value | P-value | Odds Ratio | 95% CI (Coef) | 95% CI (Odds Ratio) | Significance |
|---|---|---|---|---|---|---|---|---|
| **Speeding** | 3.200 | 0.186 | 17.183 | 0.000 | 24.533 | [2.835 , 3.565] | [17.030 , 35.340] | **** |
| **Constant** | 1.698 | 1.595 | 1.064 | 0.287 | 5.462 | [-1.429 , 4.825] | [0.240 , 124.556] | |
| **RAE (Rapid Acceleration Event)** | 0.716 | 31.789 | 0.023 | 0.982 | 2.047 | [-61.590 , 63.022] | [1.79E-27 , 2.35E+27] | |
| **VDE_flag (Vertical Dynamic Event)** | 0.272 | 0.153 | 1.779 | 0.075 | 1.313 | [-0.028 , 0.572] | [0.973 , 1.772] | |
| **SBE (Sudden Braking Event)** | 0.219 | 0.127 | 1.731 | 0.083 | 1.245 | [-0.029 , 0.467] | [0.971 , 1.595] | |
| **Lane Departure** | 0.206 | 0.088 | 2.348 | 0.019 | 1.228 | [0.034 , 0.377] | [1.035 , 1.459] | * |
| **Sharp Turn** | 0.152 | 0.083 | 1.833 | 0.067 | 1.164 | [-0.011 , 0.314] | [0.990 , 1.369] | |

```
===== FULL DATA PERFORMANCE =====
[[594 139]
 [479 945]]
              precision    recall  f1-score   support

           0     0.5536    0.8104    0.6578       733
           1     0.8718    0.6636    0.7536      1424

    accuracy                         0.7135      2157
   macro avg     0.7127    0.7370    0.7057      2157
weighted avg     0.7636    0.7135    0.7210      2157

AUC: 0.8040452503947146
Brier: 0.16590610279712328
```

COEFFICIENTS:

| Variable | Coefficient | P-value | Odds Ratio | Significance |
|---|---|---|---|---|
| **Speeding** | 3.200 | 0.000 | 24.533 | **** |
| **Constant** | 1.698 | 0.287 | 5.462 | |
| **RAE (Rapid Acceleration Event)** | 0.716 | 0.982 | 2.047 | |
| **VDE_flag (Vertical Dynamic Event)** | 0.272 | 0.075 | 1.313 | |
| **SBE (Sudden Braking Event)** | 0.219 | 0.083 | 1.245 | |
| **Lane Departure** | 0.206 | 0.019 | 1.228 | * |
| **Sharp Turn** | 0.152 | 0.067 | 1.164 | |

## Appendix F: Risk Maps in Enlarged Format(Fig F1 - Fig F4)

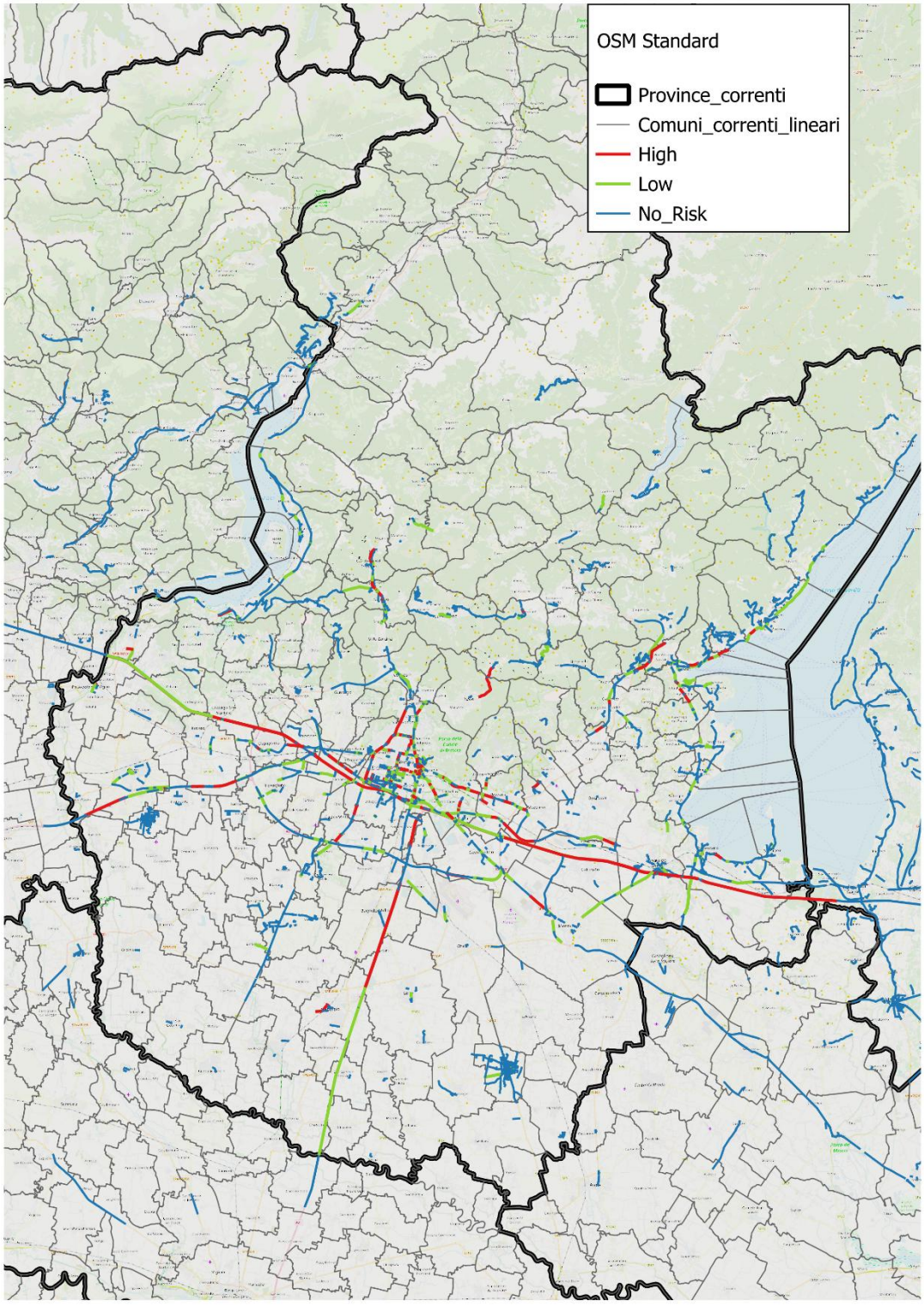



*Fig F1. Spatial Mapping of Segment-Level Crash Risk Categories (Econometric Model)*

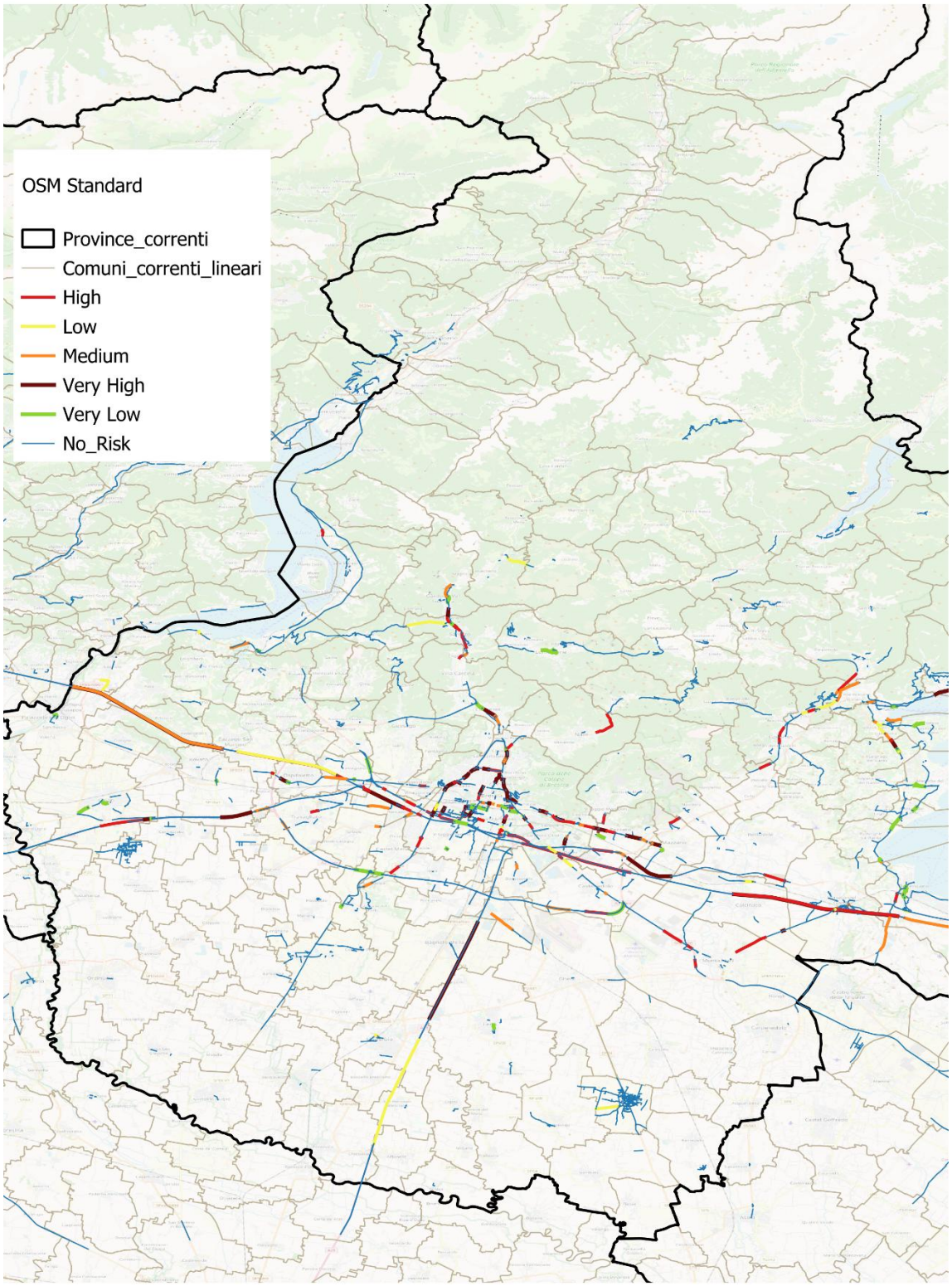



**Fig.F2 Spatial Mapping of Trip-Level Crash Risk Categories (Econometric Model)**

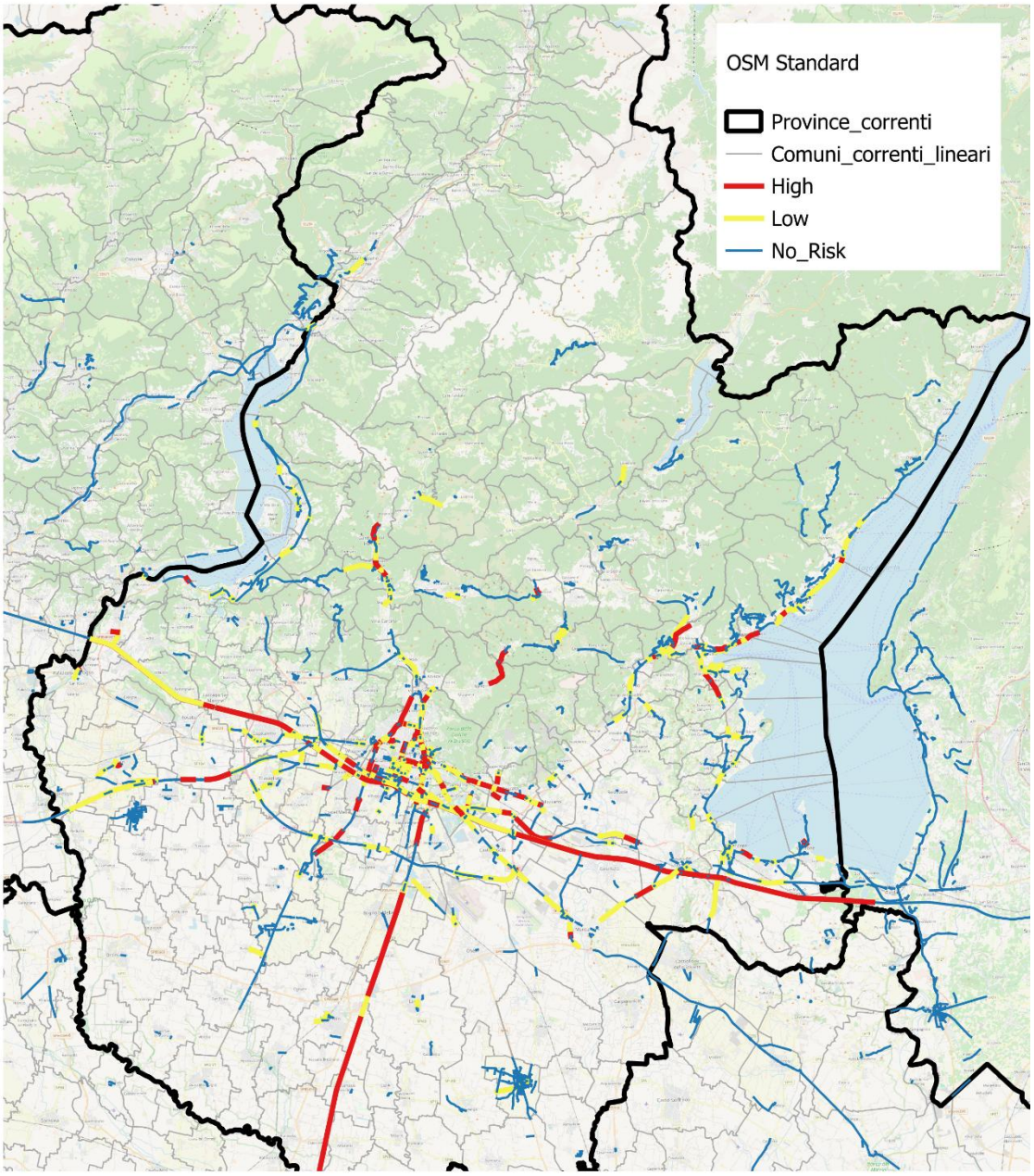



*Fig.F3 Spatial Mapping of Segment-Level Crash Risk Categories (FNN Model)*

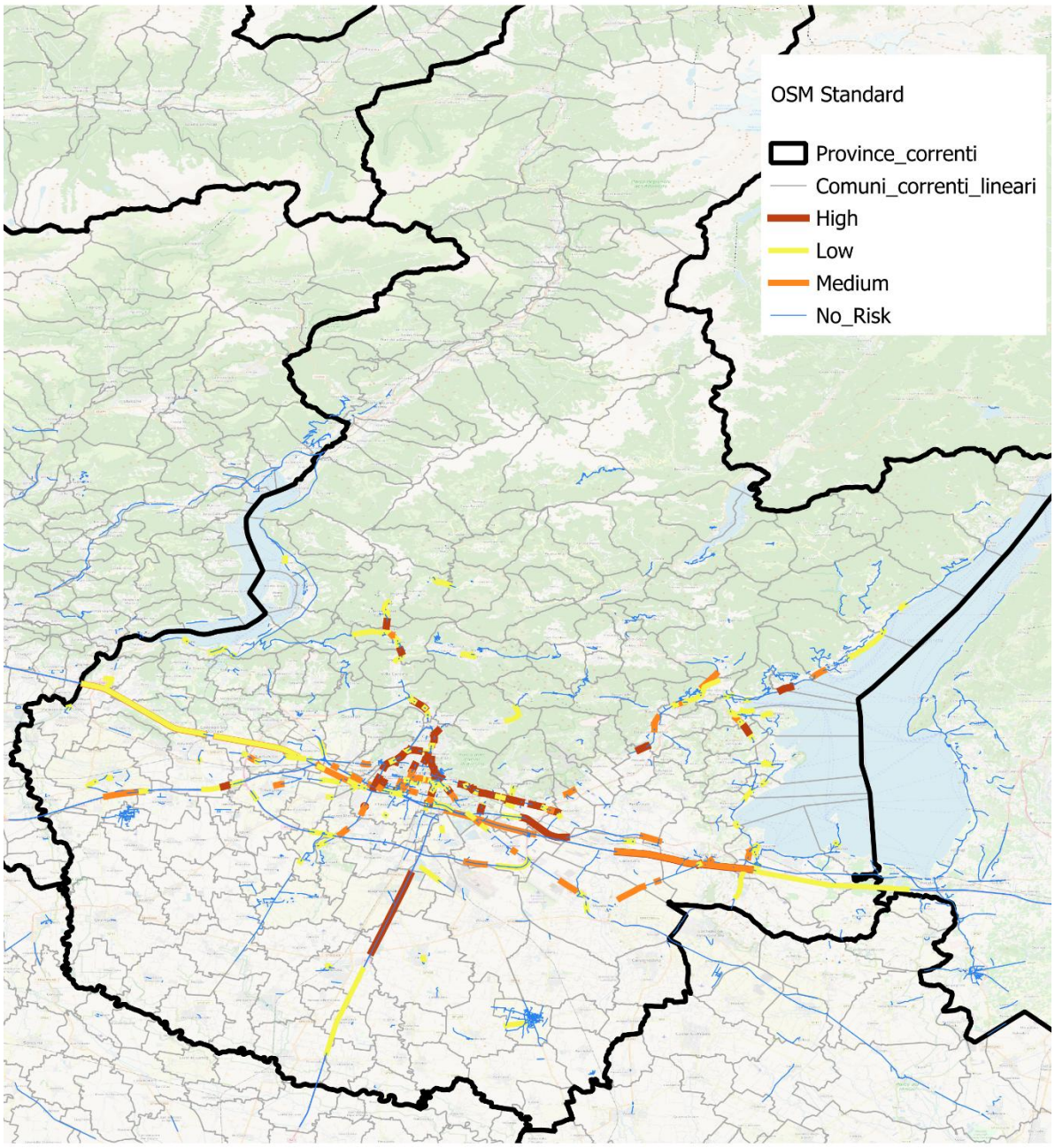



**Fig.F4** ***Spatial Mapping of Trip-Level Crash Risk Categories (FNN Model)***