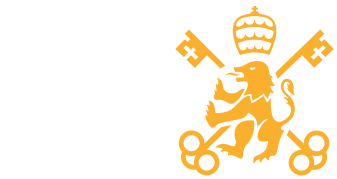


COMILLAS
UNIVERSIDAD PONTIFICIA
ICAI ICADE CIHS

Impact of Vehicle Restraint Systems on Injury Outcomes

A Combined Approach Using Real-World Crash Data and Computational Modelling

by
Manuel Valdano

supervised by
Prof. Dr. Francisco José López Valdés
Prof. Dr. Bengt Pipkorn
Prof. Dr. Jesús Ramón Jiménez Octavio

at
ICAI School of Engineering
Comillas Pontifical University

Madrid, 2024

If everything seems under control, you're not going fast enough - Mario Andretti

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

I acknowledge the use of AI-supported tools, such as ChatGPT (OpenAI, California, United States) and Grammarly (Grammarly, Inc., California, United States), to review my writing at the final stage of this thesis. It is important to note that I independently created and composed all content, using AI tools exclusively to provide assistance in improving the clarity and quality of the writing in this manuscript.

Manuel Valdano
Madrid, November 2024

Acknowledgement

As I complete this thesis, I am deeply grateful to all who have guided and supported me, from my school days through my degree and beyond. Given the limited space here, I extend my gratitude to everyone who has been part of this journey.

Embarking on my PhD journey has been both an inspiring and challenging chapter of my life. Over the past few years, I have learned a lot about my research topic and also about myself. This professional and personal growth period is largely due to the remarkable people who have surrounded me during this time. I am deeply grateful to my supervisors—Francisco López-Valdés, Jesús Jiménez-Octavio, and Bengt Pipkorn—for their support and invaluable guidance. Your mentorship has not only shaped my research journey but has also significantly influenced my personal development. Thank you for giving me the opportunity to pursue my PhD under your guidance.

Although this journey began in 2019, I finally moved from Argentina to Spain in March 2020. Just two days after my arrival, the COVID-19 pandemic began to affect my plans, leading to the closure of the university. A week later, Spain entered the lockdown phase. I am deeply grateful to Mauricio Correa, who, without knowing me, allowed me to stay in his apartment for almost two months until I got a place to move in. It was one of my first times abroad and the first time I had moved away from my parents' house. I don't know how things might have turned out without his kindness.

I would like to express my sincere gratitude to the MOBIOS team (both those who continue with us and those who have embarked on new paths) for sharing invaluable learning moments with me over the past years and providing immense support in completing this thesis. I also extend my heartfelt thanks to all my colleagues at the IIT for enriching my PhD journey and sharing good moments after work. Special thanks go to Luis Sánchez-Merchante, Daniel Pérez-Rapela, and Rocío Suárez-del Fueyo for their help and engaging discussions at the beginning of this journey. A special thank you to Karl-Johan Larsson for his assistance with generating the morphed HBMs, which was crucial to finishing this work.

I am truly grateful to have embarked on this journey in such a fantastic field, allowing me to meet many wonderful and inspiring people. While COVID-19 took away the first two years of conferences, I have cherished reconnecting with everyone at each conference. Special thanks go to the team at CAB from the University of Virginia (UVA), the Autoliv Research and Development team, who warmly welcomed me during my three-month international exchange, and the Chalmers University of Technology team.

I would like to express my gratitude to those who inspired and encouraged me to pursue this path even before I embarked on this journey. Marcela challenged me with countless extra math problems during primary school, igniting my early passion for mathematics. In high school, Alberto, Walter, and Daniel taught me many fascinating lessons that deepened my love for math, physics and engineering. Javier was a pivotal influence during that time, motivating me to look further and pursue a career that would challenge me and bring out the best version of myself. Later, Bruno, Luis, and Marcos initiated my research journey at the university, spending countless hours engaging in stimulating discussions with me. Their guidance and enthusiasm were instrumental in shaping my academic pursuits.

Finally, I wish to express my deepest gratitude to my family, who have been the ultimate support system throughout this journey. I am especially grateful to my partner, Paola, for being an extraordinary source of strength, and I am immensely thankful for her love and understanding. I also want to thank my parents, Ricardo and Rosana, and my siblings, Clara and Gabriel, who, despite the distance from Argentina, have provided their unwavering and

unconditional support. Their love and encouragement have been the cornerstone of my journey, and I could not have reached this point without them. I also want to extend my heartfelt thanks to my uncles and aunts, Leo, Dani, Norma, Silvia, and Gustavo, for their enduring support throughout the years. Lastly, I would like to thank my grandfather, Pepe, who ignited my curiosity and passion for engineering when I was a child, a spark that has propelled me all these years to discover new things and enjoy my career.

Abstract

Road traffic crashes remain a major public health issue, with vehicle safety measures, particularly restraint systems, playing a crucial role in mitigating injury risk during collisions. However, the true efficacy of these systems is often difficult to assess, as real-world crash data involves complex variables such as crash conditions, occupant characteristics, and the specific configurations of safety features in different vehicle models. While previous research has explored vehicle safety performance, these studies have not fully included the specific characteristics of the restraint systems across vehicle models. This dissertation aimed to address this gap by developing a comprehensive methodology to assess the impact of restraint systems on injury outcomes, considering the variability inherent in real-world crash conditions.

A method was developed to identify the presence and configuration of restraint system features (such as pre-tensioning and load-limiting devices), enabling an accurate representation of their real-world implementation and facilitating their integration with crash data. The impact of these restraint systems on injury outcomes was assessed using two complementary approaches. The first approach combined real-world crash data with the restraint systems to estimate the injury risk using a logistic regression method. This provided a statistical analysis of the factors contributing to moderate, severe and fatal injuries across various crash scenarios. For the second approach, a vehicle interior sled model, a human model and real-world crash conditions were combined to assess the impact of restraint systems on predicted injury outcomes. These simulations enabled an isolation of crash variables, offering deeper insights into how restraint systems perform across a wide range of real-world scenarios.

This thesis presents several key findings related to vehicle restraint systems and their impact on injury outcomes. A tool was developed to identify the presence and characteristics of pre-tensioner and load-limiting devices in vehicle restraint systems using force-time history measurements. The tool demonstrated excellent performance in identifying the presence of pre-tensioning devices (F1 score: 0.95) and high accuracy for single-stage load-limiting devices (F1 score: 0.90). However, it showed moderate performance for double-stage load-limiting devices (F1 score: 0.77). Detailed restraint system configurations for 1,318 vehicles were made publicly available via a GitHub repository, enabling broader community validation and integration with crash test databases. For the first approach to assess the impact of restraint systems on injury outcomes, NASS CDS crash data was augmented with detailed information on pre-tensioning and load-limiting devices. This revealed significant associations between these features and injury outcomes. Load-limiting forces were classified into low (below 4.5 kN) and high (above 4.5 kN) to better understand their impact. The presence of these devices was linked to a significant reduction in fatal injuries, with the reduction varying based on load-limiter force level (OR = 0.31 for low-force, OR = 0.42 for high-force devices). Low-force load-limiting devices were associated with a reduced risk of AIS 3+ whole-body injuries (OR = 0.70), and both load-limiting force categories were linked to lower AIS 2+ injury risks in the whole-body in high delta-v crashes. However, no significant associations were found between the presence of pre-tensioning and load-limiting devices and injury risks in specific body regions. For the second approach, a method was developed that employed computational models to estimate the impact of restraint systems on predicted injury outcomes. While the computer modelling-based method did not accurately capture real-world injury distributions, it showed some insights into how variations in crash conditions affect injury outcomes. Significant differences emerged between the Multibody (MB) and Finite Element (FE) models' predictions, with the morphed FE Human Body Model (HBM) demonstrating greater potential in capturing the impact of restraint systems across a range of crash condi-

tions. The method also highlighted that lower load-limiting forces were significantly linked to a reduction in rib fracture risk across both small and large delta-v values. Though the method demonstrated the ability to assess the effects of restraint system configurations on injury outcomes, it does not yet allow for precise predictions of the number of injuries that could be avoided per 1,000 crashes. Future improvements in injury criteria predictions, particularly for low-severity crashes, may enhance the accuracy of these estimations.

Resumen

Los incidentes de tráfico siguen siendo un problema importante de salud pública, y las medidas de seguridad de los vehículos, en particular los sistemas de retención, desempeñan un papel crucial en la mitigación del riesgo de lesiones en un choque. Sin embargo, la verdadera eficacia de estos sistemas es a menudo difícil de evaluar, ya que los datos de incidentes en el mundo real implican variables que van desde condiciones del accidente, las características de los ocupantes, a las características del sistema de retención. Aunque investigaciones anteriores han estudiado la seguridad de los vehículos, estos estudios no han incluido las características específicas de los sistemas de retención. Esta tesis tiene como objetivo el desarrollo de una metodología integral para evaluar el impacto de los sistemas de retención en las lesiones, considerando la variabilidad de las condiciones de los incidentes en el mundo real.

Se desarrolló un método para identificar la presencia y configuración de las características de los sistemas de retención (como dispositivos de pretensado y de limitación de carga), lo que permitió una representación precisa de su implementación en el mundo real y facilitó su integración con los datos de incidentes de tráfico. El impacto de estos sistemas de retención en las lesiones se evaluó utilizando dos enfoques complementarios. El primer enfoque combinó datos de incidentes de tráfico reales con los sistemas de retención para estimar el riesgo de lesiones mediante un método de regresión logística. Esto proporcionó un análisis estadístico de los factores que contribuyen a lesiones moderadas, graves y fatales en diversos escenarios de colisión. Para el segundo enfoque, se combinó un modelo del interior de un vehículo genérico, un modelo humano y condiciones de incidentes de tráfico reales para evaluar el impacto de los sistemas de retención en las lesiones estimadas usando el modelo de simulación. Estas simulaciones permitieron aislar variables del accidente, proporcionando una comprensión más profunda de cómo los sistemas de retención funcionan en una amplia gama de escenarios del mundo real.

Esta tesis presenta varios hallazgos relacionados con los sistemas de retención de los vehículos y su impacto en las lesiones. Se desarrolló una herramienta para identificar la presencia y características de dispositivos de pretensado y de limitación de carga en los sistemas de retención de vehículos utilizando mediciones de fuerza en un choque. La herramienta demostró un excelente rendimiento al identificar la presencia de dispositivos de pretensado (puntuación F1: 0.95) y una alta precisión para dispositivos de limitación de carga de una sola etapa (puntuación F1: 0.90). Sin embargo, mostró un rendimiento moderado para dispositivos de limitación de carga de doble etapa (puntuación F1: 0.77). Las configuraciones detalladas de los sistemas de retención para 1,318 vehículos se hicieron disponibles públicamente a través de un repositorio de GitHub. Para el primer enfoque de evaluación del impacto de los sistemas de retención en las lesiones, los datos de incidentes de tráfico reportados en NASS CDS se complementaron con información detallada sobre dispositivos de pretensado y de limitación de carga. Esto reveló asociaciones significativas entre estas características y las lesiones. Las fuerzas de limitación de carga se clasificaron en bajas (por debajo de 4.5 kN) y altas (por encima de 4.5 kN) para comprender mejor su impacto. La presencia de estos dispositivos se vinculó con una reducción significativa en las lesiones fatales, con una reducción que varía según el nivel de fuerza del limitador de carga (OR = 0.31 para dispositivos de baja fuerza, OR = 0.42 para dispositivos de alta fuerza). Los dispositivos de limitación de carga de baja fuerza se asociaron con un menor riesgo de lesiones AIS 3+ en todo el cuerpo (OR = 0.70), y ambas categorías de fuerza de limitación de carga se relacionaron con un menor riesgo de lesiones AIS 2+ en todo el cuerpo en choques con delta-v altos. Sin embargo, no se encontraron asociaciones significativas entre la presencia de dispositivos de pretensado y

de limitación de carga y los riesgos de lesiones en regiones específicas del cuerpo. Para el segundo enfoque, se desarrolló un método que empleó modelos computacionales para estimar el impacto de los sistemas de retención en las lesiones estimadas por el modelo. Aunque el método basado en modelado computacional no capturó con precisión las distribuciones de lesiones del mundo real, proporcionó algunas ideas sobre cómo las variaciones en las condiciones de choque afectan los resultados de las lesiones. Surgieron diferencias significativas entre las predicciones de los modelos de Multibody (MB) y de Elementos Finitos (FE), con el modelo de cuerpo humano (HBM) de FE adaptado a diferentes antropometrías demostró un mayor potencial para capturar el impacto de los sistemas de retención en una variedad de condiciones de choque. El método también destacó que las fuerzas de limitación de carga más bajas estaban significativamente relacionadas con una reducción en el riesgo de fractura de costillas tanto en valores pequeños como grandes de delta-v. Aunque el método demostró la capacidad de evaluar los efectos de las configuraciones de los sistemas de retención en los resultados de las lesiones, aún no permite hacer predicciones precisas del número de lesiones que podrían evitarse por cada 1,000 incidentes de tráfico. Mejoras futuras en las predicciones de criterios de lesiones, especialmente para incidentes de tráfico de baja severidad, podrían mejorar la precisión de estas estimaciones.

Acronyms

AAAM	Association for the Advancement Of Automotive Medicine
AEBS	Automatic Emergency Breaking System
AIC	Akaike Information Criterion
AIS	Abbreviated Injury Scale
ATD	Anthropomorphic Test Device
CISS	Crash Investigation Sampling System
DTW	Dynamic Time Warping
EU	European Union
Euro NCAP	European New Car Assessment Programme
FARS	Fatality Analysis Reporting System
FE	Finite Element
GDP	Gross Domestic Product
GPR	Gaussian Process Regression
GVI	Generic Vehicle Interior
GVI model	Frontal Impact Generic Vehicle Interior Sled Model
HBM	Human Body Model
HFN	Head-Face-Neck
IIHS	Insurance Institute For Highway Safety
IIHS-HLDI	Insurance Institute for Highway Safety and Highway Loss Data Institute
KDE	Kernel Density Estimate
KTH	Knee-Thigh-Hip
LASSO	Least Absolute Shrinkage And Selection Operator
LL	Load Limiter
MAE	Means Absolute Error
MAIS	Maximum Abbreviated Injury Scale
MB	Multibody

MDB	Mobile Deformable Barrier
MPS	Maximum Principal Strain
NASS CDS	National Automotive Sampling System Crashworthiness Data System
NCAP	New Car Assessment Program
NFR	Number Of Fractured Ribs
NHTSA	National Highway Traffic Safety Administration
NN	Neural Networks
OCS	Occupant Classification Systems
OR	Odd Ratio
PDOF	Principal Direction Of Force
PMHS	Post Mortem Human Subject
PP	Pre-Tensioner
PTP	Peak-To-Peak
SD	Standard Deviation
SHBM	Safer Human Body Model
SVR	Support Vector Regressions
THO	Thorax
TTF	Time-To-Fire Or Activation Time
U.S.	United States
UN	United Nations
VIF	Variance Inflation Factor
VIN	Vehicle Identification Number
WHO	World Health Organization
XAI	Explainable Artificial Intelligence

Contents

List of Figures

List of Tables

1 Introduction

This chapter discusses the motivation for this work, presents this thesis's scope, and objectives, and concludes by presenting the overall structure of the work. The motivation begins by addressing the burden of road traffic fatalities. It then provides an overview of vehicle safety assessments, focusing on the role of anthropomorphic test devices and human body models for the development and optimisation of restraint systems. The discussion then shifts to the influence of restraint system configurations on preventing occupant injuries, emphasising their critical role in enhancing safety. A brief overview is provided on the application of metamodels for estimating injury risks, predicting injury patterns, and exploring how these predictions can be carried out. Finally, the thesis scope and objectives are presented, and a description of the thesis layout is provided.

1.1 Literature review and motivation

Road traffic deaths continue to be a significant global public health issue, with 1.19 million fatalities reported in 2021, according to the latest report from the World Health Organization (WHO, 2023a). Road traffic injuries remain the leading cause of death for individuals aged 5-29 years and are the 12th leading cause of death across all age groups (WHO, 2022). Additionally, between 20 and 50 million people suffer non-fatal injuries annually due to road traffic incidents (WHO, 2023b). These injuries often lead to long-term disabilities that significantly diminish the quality of life for the affected victims. Moreover, the economic burden of these incidents is substantial, costing societies between 1 % and 3 % of their Gross Domestic Product (GDP), and in some cases, as much as 6 % (Dalal et al., 2013).

Different initiatives have been launched to address this issue, such as the Decade of Action for Road Safety 2021-2030 and the Vision Zero Initiative. The Global Plan for the Decade of Action was developed by the WHO and the United Nations (UN) regional commissions, aiming to prevent at least 50 % of road traffic deaths and injuries by 2030 (WHO et al., 2021). The Vision Zero Initiative, first introduced in Sweden in 1994, seeks to eliminate road traffic fatalities based on the principle that "No loss is acceptable" (*Vision Zero Initiative - TRIMIS - European Commission* 2017). The European Union (EU) has adopted these principles, aligning its objectives with the Decade of Action by committing to reduce by half road deaths by 2030 and achieve zero fatalities and serious injuries in Europe by 2050 (European Commission, 2019).

The WHO classifies road users into five categories: 4-wheelers, powered 2/3-wheelers, cyclists, pedestrians, and others. The mentioned initiatives aim to reduce deaths and injuries across all categories of road users. Regarding 4-wheeler occupants, 25 % of road traffic deaths globally were observed among these road users, with this figure rising to 50 % in the European region, according to the WHO (2023a).

1.1.1 Vehicle safety assessment

Safer vehicles are one of the aims of the above-mentioned initiatives to enhance road user safety. Regarding 4-wheeler road users, vehicle safety improvements have saved over 600,000 lives in the United States between 1960 and 2012 (Kahane, 2015). The vehicle's structural integrity and the restraint system are the primary components used to protect the occupant. Restraint systems, such as seatbelts and frontal airbags, have proven to be highly effective in reducing fatalities. According to the National Highway Traffic Safety Administration's (NHTSA) National Center for Statistics and Analysis (National Center for Statistics and Analysis, 2019), over 17,000 lives were saved in the United States in 2017 due to the use of these restraint systems. However, an additional 2,500 lives could have been saved if all passengers had been wearing seatbelts.

In the development of restraint systems, Anthropomorphic Test Devices (ATDs) are used to predict the mechanical response of human occupants in vehicle collisions. Government and consumer testing programs and vehicle manufacturers utilise ATDs in a set of crash test configurations based on real-world collisions and evaluate the performance of restraint systems (Hershman, 2001; Hobbs et al., 1998; O'Neill, 2009). Physical measurements, such as accelerations, forces, and deformations, are collected to estimate the risk of a human occupant sustaining an injury in a crash. Injury severity scales are used to classify the potential injuries that an occupant could sustain. One widely used system is the Abbreviated Injury Scale (AIS), developed by the Association for the Advancement of Automotive Medicine (AAAM, 2008), which provides a scale from 1 (least severe) to 6 (most severe) that describes the potential threat to life posed by an injury among other dimensions (Seguí-Gómez et al., 2011).

Using the AIS scale, specific injury severities or groups (such as AIS 2 or AIS 3+) are employed to develop injury criteria, which are then linked to the risk of sustaining an injury through injury risk functions. An injury criterion is a physical measurement extracted from the ATD that is linked to injury risk, such as chest deflection and the risk of a thoracic injury (Horsch et al., 1991). These physical measurements and injury risk are linked using an injury risk function that is derived from experiments using Post Mortem Human Subjects (PMHS) in similar loading conditions. This approach enables the estimation of injury risk in crash scenarios based on measurable physical responses from ATDs. Consequently, the configuration of the restraint system can be optimised to minimise the risk of occupant injuries using crash tests and computer simulations.

1.1.2 Restraint system configuration impact on occupant injuries

This lack of effectiveness in preventing severe thoracic injuries is particularly concerning, given that modern restraint systems incorporate advanced and costly features such as multistage airbags, load limiters, pre-tensioners and knee airbags. Despite these technological advancements, these systems have not yet shown the expected reduction in thoracic injury risk in real-world crashes.

A recent analysis of the National Automotive Sampling System Crashworthiness Data System (NASS CDS) showed that in n crashes with newer model-year vehicles, the risk for the occupant to sustain AIS 2+ and AIS 3+ injuries to the lower extremities and head was significantly reduced while the reduction in AIS 3+ rib or sternum fractures was no significant (Forman et al., 2019). This lack of effectiveness in preventing severe thoracic injuries is particularly concerning, given that modern restraint systems incorporate advanced and costly features such as multistage airbags, load limiters, and pre-tensioners of varying magnitudes and locations. Despite these technological advancements, these systems have not yet shown

the expected reduction in thoracic injury risk in real-world crashes. On the contrary, several other experimental studies using PMHS and ATD, as well as studies using computational human body models (HBMs), which are computational representations of the human anatomy, have suggested that seat belt systems incorporating pre-tensioners and load-limiting devices were effective in the prevention of thoracic injuries (Forman et al., 2009; 2008; Kent et al., 2001; 2007; Michaelson et al., 2008).

Previous research has suggested that while experimental studies, whether involving physical testing or computational simulations, may be pertinent for designing and evaluating the nominal impact of more advanced restraint systems, the actual effectiveness of these systems can only be estimated when they are assessed using real-world injury data (Segui-Gomez et al., 2009). Using data from the fatality analysis reporting system (FARS), NHTSA estimated that a belted driver or front seat passenger had, on average, an estimated 12.8 % lower fatality risk if the seat belt was equipped with a pre-tensioner and a load limiter than if it was not equipped with either. However, the large 95 % confidence interval (2.6 % - 23.0 %) suggests that the estimation needs to be refined, perhaps by increasing the size of the analysed sample (Kahane, 2000).

A non-negligible difficulty for these large epidemiological studies is the challenge of controlling for all the independent variables that might have had an effect on the analysed outcome. In the specific case of thoracic injuries, such a real-world assessment would require that the crash database be used, including whether the vehicles were equipped with pre-tensioners and load limiters. Moreover, it would also need to include the force levels at which these systems act on the passengers, as these values differ between different vehicle models.

NASS CDS has been an essential data source for this epidemiological evaluation, given the detailed information collected in the system. However, the characteristics mentioned above of advanced seat belts are not included in this database, making it difficult to assess their effect using real-world data. Augmenting the database with information on restraint system configurations would enable more comprehensive analyses of their impact on injury outcomes.

NHTSA provides consumers with information on the crashworthiness performance of vehicles through the New Car Assessment Program (NCAP). The NCAP program uses different crash conditions, such as front and side-impact crashes and rollover resistance, to test the crashworthiness of vehicles. The data generated from these tests are shared with consumers through a user-friendly five-star rating system. In contrast, the technical data are available via the NHTSA's research testing database (NHTSA, 2023). This database contains detailed reports, videos, pictures and time history measurements from the tests, which can be utilised to identify the features of the restraint system.

1.1.3 Virtual anthropomorphic test device and human body models

Computer models offer a complementary approach to studying crash data gathered from real-world collisions. This method allows analysts to explore scenarios based on real-world crashes, investigate effects that may not be statistically significant in existing datasets, and examine the performance of restraint systems that have not yet penetrated the vehicle fleet, as well as new solutions that have not been released to the market. Moreover, computer simulations enable the controlled manipulation of variables, allowing analysts to isolate and analyse the impact of specific factors on injury outcomes.

In recent years, human body models (HBMs) have been developed to further enhance the assessment of passive vehicle safety. These HBMs are computational representations of hu-

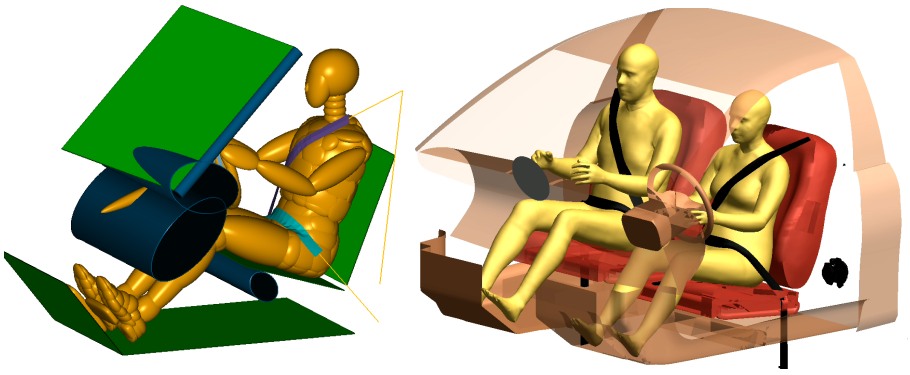
man anatomy, enabling the simulation and analysis of how the human body responds to vehicle collisions. They are designed and validated for omnidirectional injury risk assessment (Gayzik et al., 2011; Kato et al., 2018; Meijer et al., 2013; Östh et al., 2017; Pipkorn et al., 2023), complementing ATDs, which were traditionally developed for specific crash directions (Foster et al., 1977; Lowne et al., 1987; Neilson et al., 1985; Page, 2001; Parent et al., 2013). Furthermore, ATDs were designed to withstand repeated use in research and regulatory crash tests, leading to requirements for repeatability and reproducibility in addition to biofidelity to human response (Crandall et al., 2011). However, previous research has shown that these requirements can limit the ability of ATDs to accurately mimic human responses (Beeman et al., 2012; Higuchi et al., 2019; Lopez-Valdes et al., 2018; López-Valdés et al., 2016; Perez-Rapela et al., 2019; Seacrist et al., 2010). Furthermore, HBMs enable the prediction of injury outcomes at a tissue level and a level of detail that can not be implemented in ATDs.

HBMs are typically employed using two main types of computer models: multibody (MB) models and finite element (FE) models. MB models simplify the human anatomy by representing it as a system of rigid bodies connected by idealised joints. This approach allows for efficient simulation of the human body's overall motion and dynamics during vehicle collisions. Additional elements, such as muscles and ligaments, can be included to increase anatomical fidelity. One widely used MB model is distributed by Simcenter Madymo^{TM} (Helmond, The Netherlands), which represents a 50th percentile male occupant and has been validated for omnidirectional injury risk assessment (Meijer et al., 2013). However, due to the simplifications in their design, MB models primarily apply directional injury criteria similar to those used with ATDs. On the other hand, FE models provide a more detailed and anatomically accurate representation of the human body. FE models divide the human anatomy into thousands or even millions of smaller elements to mimic the mechanical behaviour of tissues, bones, and organs during impact scenarios. This enables using stress-, strain-, pressure- and energy-based injury criteria for specific anatomical structures to predict complex injury mechanisms, which do not apply to ATDs due to their lower anatomical detail (Fahse et al., 2023). However, using FE models has significantly higher computational demands, making simulations more resource-intensive and time-consuming.

Figure 1.1 shows an example of each modelling technique. Subfigure (a) shows an MB HBM driver in a simplified vehicle interior. Subfigure (b) shows an FE HBM driver and front-seat passenger in the interior of a vehicle. These occupants represent a 50th percentile female and 50th percentile male anthropometries, respectively.

Pipkorn et al. (2020) studied moderate to severe injuries in passenger vehicle crashes to establish guidelines for developing HBMs. Their findings suggest that future HBM development should focus on head, thorax, and lower extremities injuries. Additionally, previous research has shown that specific segments of the population, such as elderly, obese, and female occupants, are at a higher risk of death and serious injuries in these body regions (Abrams et al., 2020; 2022; Bose et al., 2011; Carter et al., 2014; Forman et al., 2019; Kahane, 2013b; Noh et al., 2022; Ranmal et al., 2024; Viano et al., 2008). While factors such as crash severity and vehicle type explain some of the differences between male and female occupants (M. Brumbelow, 2023; M. L. Brumbelow et al., 2022), a deeper understanding of how these groups interact with restraint systems could be essential for developing systems that are effective in preventing injuries in these groups when developing adaptative restraint systems (Anja et al., 2011; Boyle et al., 2020; Huang et al., 2015; McCarthy et al., 2001; Miller et al., 1996; Shin et al., 2007; W. Sun et al., 2023).

Kent et al. (2005) categorised the anatomical factors that influence injury risk into three groups: geometric, compositional, and material. In terms of geometric variation, current HBMs are primarily based on the anthropometries used to develop ATDs, which represent



(a) MB HBM example (b) FE HBM example

Figure 1.1: Examples of MB and FE HBMs in a driver and/or front seat passenger position. The MB HBM is the Madyno 50th percentile male occupant. The FE HBM is the VIVA+ HBM in its two occupant variations: the 50th percentile female occupant is seated in the driver position, and the 50th percentile male occupant is seated in the front seat passenger position.

a 5th percentile (small size) female, a 50th percentile (average size) male, and a 95th percentile (large size) male. The THUMS (Kato et al., 2018) and GHBMC (Gayzik et al., 2011) HBMs follow this distribution. In contrast, the VIVA+ HBM (John et al., 2022) has been developed to represent both a 50th percentile male and female occupant. The SAFER HBM (Pipkorn et al., 2023), however, was developed with a single size, representing a 50th percentile male.

Despite these standardised sizes, FE models enable morphing the baseline HBM to specific anthropometries using frameworks like PIPER (Beillas et al., 2016) and parametric HBM morphing (Hwang et al., 2016a; c). Figure 1.2 shows a simplified example of the morphing technique used to adjust the HBM anthropometry. After morphing the HBM, it is essential to validate the new anthropometry to ensure it accurately mimics the behaviour of such occupants (Hwang et al., 2020; Larsson et al., 2021b; Schoell et al., 2015). These models can also be adjusted to simulate drivers and passengers in postures commonly observed in real-world scenarios (Park et al., 2016; Reed et al., 2020; 2002). Compositional and material changes across occupants can also be included on these morphed HBMs in the obtention of the computer model (Larsson et al., 2021a; 2023).

1.1.4 Exploration of injury risk using metamodels or surrogate models

Computer models can be employed to estimate the effectiveness of advanced restraint systems under real-world conditions. As mentioned, existing software and models allow for the simulation of different combinations of occupant anthropometries, vehicles, and impact conditions to assess injury risk. However, two significant challenges may limit the use of computer models for this objective. First, real-world crash databases do not gather all the information necessary to reconstruct such detailed models accurately. Second, even if this information were available, modelling each individual case would be impractical due to the substantial effort required to develop a detailed model for every scenario.

To address these challenges, previous studies have integrated real-world variability into their simulations to investigate its impact on injury outcomes. The selection of an appropriate design of experiments is crucial in minimising the number of simulations needed to capture the effects of these variations effectively (Schneider et al., 2022). Two main approaches

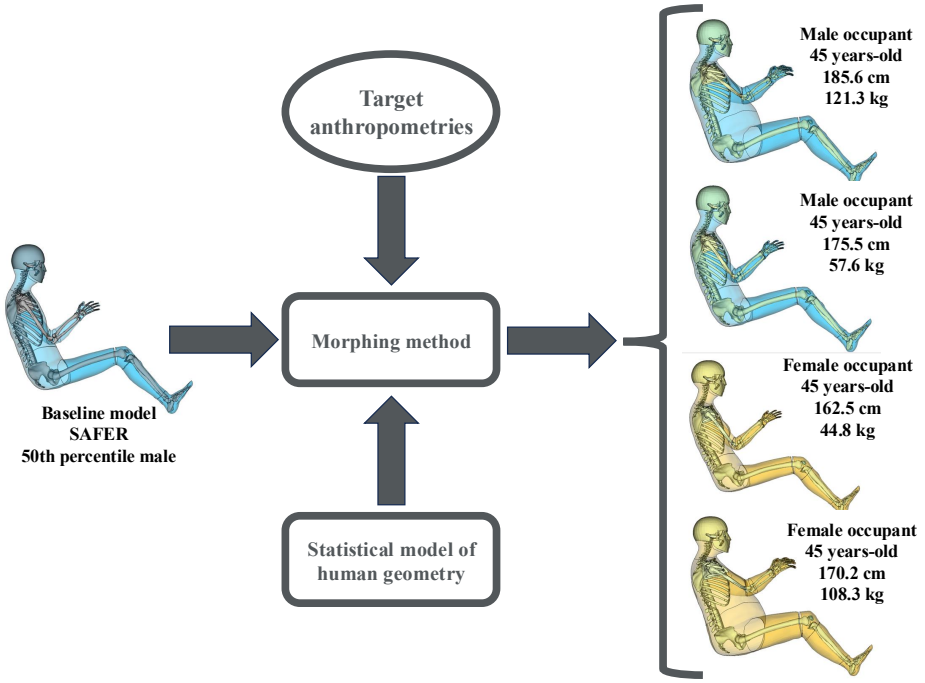


Figure 1.2: Example of the SAFER HBM being morphed into four different anthropometries.

have been applied when selecting the number of cases for large sets of simulations. On one hand, some studies predetermined the number of simulations (Hu et al., 2016; Hu et al., 2015; 2017; 2019; Hwang et al., 2016b; Iraeus et al., 2016; Joodaki et al., 2020a; 2021). On the other hand, the number of simulations can be increased iteratively until a specific condition is met. For example, Larsson et al. (2024), Perez-Rapela et al. (2020), and Schneider et al. (2022) selected the number of simulations based on the convergence of a metamodel (i.e., surrogate models). These metamodels are functions that correlate the input(s) used to construct the simulation model with the output(s) of the model, providing a more efficient way to explore complex relationships.

The convergence of these metamodels is typically assessed based on their prediction error, which is defined as the difference between the actual output of the simulations and the metamodel's prediction for the same input (i.e., the same impact conditions). Different mathematical functions or regression models can be used to develop the metamodel, such as the least absolute shrinkage and selection operator (LASSO), neural networks (NN), support vector regressions (SVR), and Gaussian process regression (GPR). However, the most suitable regression model for a given application is often unknown until the prediction error is measured.

Accurately determining the appropriate number of simulations (i.e., the sample size) and selecting the optimal regression method are critical steps in developing a metamodel that reliably predicts the output of the computational model. To achieve this, the sample size is gradually expanded until the metamodel converges, as indicated by the stabilisation of its prediction error. Once the final metamodels are obtained, the one with the lowest prediction error is selected to replace the computational model, providing an efficient surrogate model for further analysis.

1.1.5 Prediction of injuries and injury patterns

A computer model should aim to predict trends that closely align with those observed in real-world crashes under similar conditions. While achieving precise predictions of injury outcomes remains challenging, the correlation between the outputs of these models and real-world crashes can be assessed in several ways, such as comparing the risk for an injury (Hu et al., 2015; Iraeus et al., 2016; Takhounts et al., 2019) or the incidence of injuries (Bance et al., 2021; Kim et al., 2020; Perez-Rapela et al., 2020; Wang et al., 2021). On the one hand, the risk comparison has been carried out under selected crash conditions (Hu et al., 2015; Takhounts et al., 2019) or through the comparison of an injury risk function obtained from real-world data and the model (Iraeus et al., 2016). On the other hand, the prediction of injuries of a specific severity has been used for calculating global injury metrics (Kim et al., 2020; 2023), predicting victims' triage based on the severity of specific injuries (Bance et al., 2021; Wang et al., 2021), and for predicting the number of injured occupants (Perez-Rapela et al., 2020). However, the methodologies used for these estimations differ across the mentioned studies.

Predicting injuries based on the model output is particularly challenging when translating a physical magnitude measured in the model to actual injuries. This conversion may require the use of an injury risk function to obtain the risk of sustaining an injury of a certain AIS severity (AAAM, 1998; 2008). However, beyond determining the likelihood of a specific injury, it is crucial to develop a method that can assign whether the model predicts an actual injury based on the estimated probability. For instance, let us assume that the model predicts a chest deflection of 60 mm in a particular simulation. This deflection is associated with a 30 % probability of sustaining an AIS 3+ injury to the thorax as given by the Hybrid III 50^{th} percentile male dummy using maximum chest deflection injury criterion (Eppinger et al., 1999). However, when aggregating injury predictions from all simulations to benchmark them against real-world injury distributions, a key question arises: How can we determine whether the 30 % probability of sustaining the injury should be counted as an actual injury?

The studies referenced have employed two primary approaches to address the challenge of determining whether a predicted injury risk should be classified as an actual injury. The first approach, referred to as the "deterministic method," categorises an occupant as either injured or non-injured based on a predefined threshold value. This threshold can be set arbitrarily or optimised to achieve specific objectives (Lubbe et al., 2015). For instance, both Bance et al. (2021) and Wang et al. (2021) adopted a 25 % injury risk threshold to predict if the occupant sustained an injury of a specific severity, applying the threshold to an injury risk function tailored to each AIS level. Another example is provided by Perez-Rapela et al. (2020), which predicted the occurrence of head injuries based on the metamodel's predicted head excursion.

The second approach, referred to as the "probabilistic method," uses the injury probability predicted by the injury risk function to determine whether an injury has occurred. This is performed by randomly sampling from a binomial distribution, which simulates a population of occupants exposed to the same crash conditions (Kim et al., 2020; 2023). This method emulates a pool of occupants who crashed under the same conditions, with only a group of them sustaining injuries according to the injury probability. Then, the method randomly selects an occupant of the pool to obtain the crash outcome.

These two methods are referred to as "deterministic" and "probabilistic" because, while the injury criteria values generated by a computer model remain consistent for the same input conditions, the predicted injury severity may differ depending on the method used. In the deterministic approach, an injury is considered to have occurred if a predefined injury threshold is exceeded. Consequently, this method produces the same injury severity for iden-

tical input conditions every time. In contrast, the probabilistic method introduces a random component into the evaluation process, meaning that the predicted outcome may vary, even when the same input conditions are applied repeatedly. Despite the use of both approaches in prior research, to our knowledge, a rigorous comparison of their outcomes has not yet been undertaken.

1.1.6 Motivation

Road traffic crashes continue to be a significant public health issue, and improving vehicle safety remains a critical strategy in reducing the severity and frequency of injuries sustained in such incidents. Restraint systems play a ritual role in protecting occupants during collisions as a part of the different safety features used in the vehicle fleet. The true efficacy of vehicle safety systems, however, can only be evaluated through analyses that integrate real-world crash data. This integration provides a more comprehensive understanding of how restraint systems function across various crash scenarios and occupant characteristics.

The performance of vehicle restraint systems is influenced by a multitude of factors, including the specific characteristics of the crash and the individual characteristics of the occupants. Designing restraint systems that reliably protect across such diverse conditions requires a nuanced understanding of how these systems behave during real-world collisions. While previous studies have generally addressed the performance of vehicle safety features using crash databases, they have, to our knowledge, lacked a detailed description of restraint systems. This includes how these features may vary across different vehicle models and manufacturing years. By accounting for the detailed, vehicle-specific attributes of restraint systems, it becomes possible to more accurately evaluate their role in mitigating injuries across a wide range of real-world crash conditions. This specificity may be a very important component for the development of more effective and adaptable safety systems.

One of the key challenges in using real-world crash data to evaluate restraint systems is the difficulty in isolating the effects of specific restraint features. This is often due to the limited number of cases where these features are present, which can undermine the statistical significance of the findings.

One of the challenges in using real-world crash data to evaluate restraint systems may rely on the difficulty of isolating the effects of specific restraint features. This may be the result of the limited number of cases where these features are present, which can undermine the statistical significance of the findings. Furthermore, isolating individual factors that contribute to injury outcomes can be difficult in such complex, multifactorial events. To overcome these limitations, the use of computer simulations presents a powerful alternative. By leveraging computational models, controlled environments can be created where individual factors (such as restraint system characteristics) can be isolated and tested under a wide array of crash scenarios. HBMs offer an advanced means of simulating the impact of these factors on injury risk, enabling more precise predictions of injury outcomes.

The ability to predict injury outcomes and patterns through simulations opens new avenues for advancing the comparison of computer model results with real-world crash data. For example, Kim et al. (2020) used estimated injury severities across various body regions to calculate a whole-body injury metric, which was then validated against actual pedestrian injury data. This method illustrates the potential for simulations to complement the findings from existing injury databases, such as the NASS CDS. These databases often reveal complex injury patterns, with some individuals suffering multiple injuries in a single event. These complex injury patterns may be influenced by underlying factors, which can be further ex-

plored through computer models. Therefore, estimating the number and severity of injuries using computer models could complement current methodologies and offer valuable insights for improving vehicle restraint systems.

Further research into real-world crash data has already begun to uncover key injury trends among vehicle occupants. For instance, a study by Suarez-del Fueyo et al. (2020) clustered NASS CDS data for belted occupants to examine patterns of AIS 3+ injuries across different body regions, accounting for both occupant and crash severity characteristics. Their findings highlighted several trends, including a higher incidence of knee, thigh, and hip injuries among obese occupants and a reduction in severe injuries in newer vehicle models in lateral impacts. Notably, thoracic injuries were identified as the most common severe injury, particularly among elderly occupants. However, this study did not account for the type or effectiveness of the restraint systems used, which could enhance the understanding of their impact on preventing injuries. Including this information could provide valuable insights into how changes in restraint systems influence injury outcomes, potentially improving safety in some regions while inadvertently increasing risk in others for some conditions. This data could be instrumental in optimising future adaptive restraint systems for robust occupant protection.

1.2 Scope and objectives of the thesis

The aim of this PhD dissertation is to develop a method for assessing the impact of restraint systems on injury outcomes, considering the variability of real-world crash conditions. First, the presence and characteristics of the restraint system features (the pre-tensioning and load-limiting devices) implemented in vehicles were identified from crash test data. Second, two approaches were used to evaluate how the identified restraint systems impacted the injury outcomes in real-world crash conditions. The first approach combined real-world crash data with the restraint systems to estimate the injury risk using a logistic regression method. The second approach combined a vehicle interior sled model, a human model and real-world crash conditions to assess the impact of restraint systems on predicted injury outcomes.

Thus, to achieve the aim of this dissertation, the following three specific objectives were proposed:

- Objective 1: To develop a method for identifying the presence and characteristics of restraint system features.
- Objective 2: To assess the effect of restraint systems on injury risk using real-world crash data.
- Objective 3: To assess the effect of restraint systems on predicted injury outcomes based on computer modelling incorporating real-world crash conditions.

1.2.1 Thesis outline

This dissertation is organised into five chapters. The first chapter presents the introduction and outlines the scope of the dissertation. The second chapter addresses the first and second objectives mentioned above. Chapters three and four are dedicated to exploring the third objective. Finally, the fifth chapter provides a discussion of the dissertation, its conclusions, and recommendations for future research. In addition to the main chapters, seven appendices are included at the end of the manuscript, offering supplementary information. Furthermore,

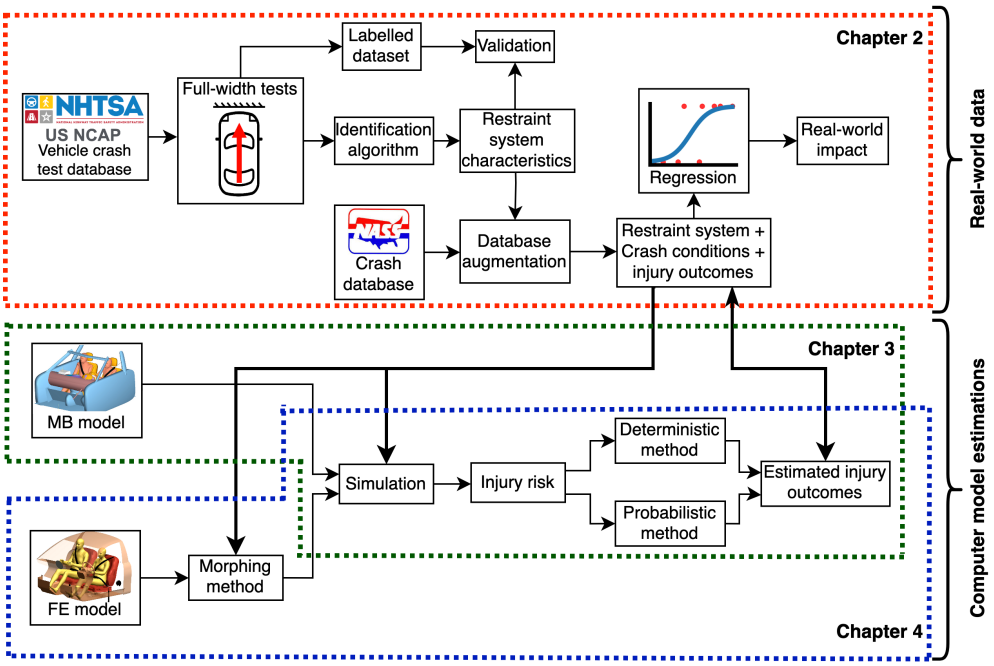


Figure 1.3: Flowchart illustrating the chapters and methodologies employed throughout the dissertation.

Figure 1.3 presents a flowchart highlighting the materials, methods, and the corresponding chapters in which they were applied.

Chapter 2 address the first and second objectives of the dissertation. This chapter details the development and validation of the method for automatically identifying restraint system configurations. It also includes a discussion on the evolution of identified restraint system characteristics over the last 40 years. This chapter offers a comprehensive description of the database augmentation process and presents the results of the multivariate logistic regression analysis for moderate and severe injuries. The overall structure and flow of this chapter are illustrated within the red dashed rectangle in Figure 1.3.

Using a sample from real-world crashes is not without its limitations, such as the potential lack of statistically significant results in the available sample studied or the difficulty in isolating individual factors. To overcome these challenges, Chapters 3 and 4 focus on the third objective of the dissertation by developing a method to assess the impact of restraint system characteristics on injury outcomes using computer models (and therefore, generating an ad-hoc population of crashes to carry out this assessment).

The developed method is designed to estimate how restraint systems impact injury outcomes in real-world conditions. To achieve this, Appendix C provides a detailed description of the distributions for the variables that describe the crash configurations observed in Chapter 2. To ensure the method accurately predicts injury outcomes, Chapter 3 and Chapter 4 present a comparison of the injury outcomes estimated from deterministic and probabilistic approaches with the injury data presented in Chapter 2. These two approaches have been applied by previous studies to estimate injury outcomes in the real world (Bance et al., 2021; Kim et al., 2020; Perez-Rapela et al., 2020; Wang et al., 2021). On the one hand, the deterministic approach applied threshold values based on the lower limits used at Euro NCAP for the Assessment Protocol for Adult Occupant Protection (Euro NCAP, 2023a). On the other hand, the probabilistic method was based on Kim et al. (2020) but used just two injury severity levels.

This comparison took into account the number of estimated injured occupants, the injury patterns sustained, and the conditions under which these injuries were predicted. Benchmarks of the estimations of the method are presented using the Madymo human model for MB simulations and the SAFER HBM, morphed into different anthropometries, for FE simulations. Chapter 3 focuses on the application of the MB model, while Chapter 4 details the use of the FE model. The method applied in these chapters is illustrated within the green and blue dashed rectangles in Figure 1.3.

2 Characterisation of restraint systems and assessment of their influence on occupant injury risk in real-world crashes

This chapter focuses on identifying the presence and characteristics of restraint system features and assessing their impact on injury risk using real-world data. To achieve this objective, augmenting the real-world crash database with detailed information on the restraint system configurations was necessary. Thus, the presence and characteristics of restraint system features were identified from crash test data, which required developing and validating an algorithm to identify these elements accurately. Then, the augmented dataset using the identified restraint systems was utilised in a multivariate logistic regression model to study the influence of these characteristics on injury outcomes. Additionally, the chapter provides a description of how the presence and characteristics of restraint system features have evolved over the past forty years, which enabled a second validation of the identified method.

2.1 Introduction

To effectively assess the impact of restraint systems on the real world, the developed method must be applicable across all types of crashes. While it is possible to study every crash scenario, this research specifically concentrated on a targeted subset of frontal crashes. The selection of this subset was guided by two primary factors: the availability of crash test data to analyse the features and configuration of the restraint systems, and the focus on crash types associated with the highest rates of fatalities and injuries among occupants.

In 2021, frontal crashes accounted for approximately 65 % of fatal outcomes and 60 % of injuries in passenger vehicle crashes in the United States (US) (National Center for Statistics and Analysis, 2023). Meanwhile, NHTSA provides public access to test databases (NHTSA, 2023) that contain engineering data collected from various research and testing activities, including the NCAP program, biomechanics studies, sled tests, and compliance testing. These databases include detailed reports, videos, images, and time history measurements from the tests. This study focused specifically on crash tests conducted under the NCAP program. Vehicle-to-barrier crashes carried out within this program were selected as the most representative impact condition for frontal crashes.

Vehicle restraint systems are compromised of several integrated components that collaborate to protect occupants during a crash. Key components include seat belts, load limiters, pre-tensioners, frontal airbags, side airbags, head restraints, child restraint systems, collapsible steering columns, knee bolsters, knee airbags, and anti-submarining pans. According to National Center for Statistics and Analysis (2019), almost 15,000 lives were saved in the US in 2017 due to the use of seat belts.

The seat belt provides a better restraint when tightly adjusted, but a three-point belt with a retractor sacrifices some initial restraining ability for daily comfort (Håland, 2006). The introduction of the retractor pre-tensioner, first installed by Mercedes-Benz in 1983, addresses this by quickly tightening the belt in response to a crash, effectively reducing upper-torso load during frontal collisions. The most significant advancement in the three-point seat belt is shoulder belt force limitation (Håland, 2006). Renault introduced a 6 kN retractor load limiter in 1995, later reduced to 4 kN in 1998 with an optimised airbag to reduce chest deflection.

While different features influence occupant behaviour during a crash, the presence of load-limiting and pre-tensioning devices can be specifically identified through time-history measurements taken from the seat belt during crash tests. Figure 2.1 illustrates the identification of the pre-tensioning and load-limiting features within the restraint system. The presence of a pre-tensioner device is indicated by a sharp increase in seat belt force captured by the load cell, which is typically triggered by a pyrotechnic charge, among other mechanisms. As for the load limiter, its presence is characterised by the stabilisation of force measured in the load cell, resulting from the plastic deformation of a component designed to limit the maximum force exerted on the occupant by the seat belt.

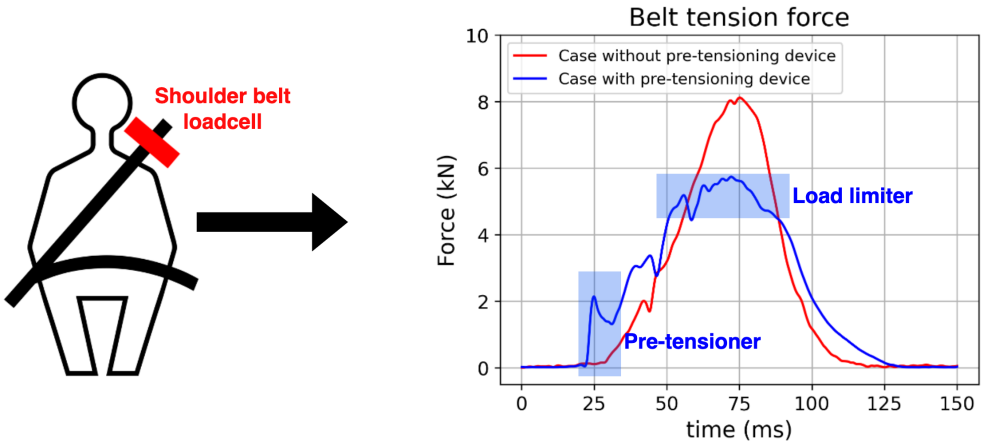


Figure 2.1: Identification of pre-tensioning and load limiting features from seat belt loadcell time-history measurements.

These measurements also enable the identification of specific characteristics of the restraint system features, such as the activation time (time-to-fire, or TTF) and the force applied by the pretensioner device, as well as the load-limiting force provided by the load limiter. Consequently, the method for identifying restraint system configurations was developed to focus on these characteristics, utilising data from vehicle-to-barrier crash tests conducted under the NCAP program.

The NCAP program is dedicated to evaluating vehicles sold in the US, meaning the identified characteristics of restraint systems are specific to the US vehicle fleet. Consequently, the analysis of restraint system effectiveness was conducted within the US context. The NASS CDS, which collected crash data nationwide from 1979 to 2015, provides a comprehensive record of police-reported crashes where at least one vehicle required towing due to damage. To augment the data of the NASS CDS, restraint system characteristics identified from NCAP crash tests were merged into the dataset. Both the NASS CDS and NCAP databases

include Vehicle Identification Numbers (VINs) for vehicles involved in crashes, enabling the use of specific metadata within the VIN for this data augmentation process.

2.2 Methods

This section of the chapter is divided into three parts. The first part explains the method to characterise the pre-tensioning and load-limiting devices based on force time history measurements. The second part explains the method to carry out a descriptive analysis of the adoption of the mentioned features and the evolution of the identified characteristics. The last part explains the method to aggregate the data from the restraint system characteristics to the NASS CDS database and assess the influence of such characteristics on AIS 2+ and AIS 3+ real-world injuries. Figure 2.2 shows a brief summary of the methodologies applied in the chapter to identify the restraint system characteristics and assess their impact on occupant injuries using real-world crashes.

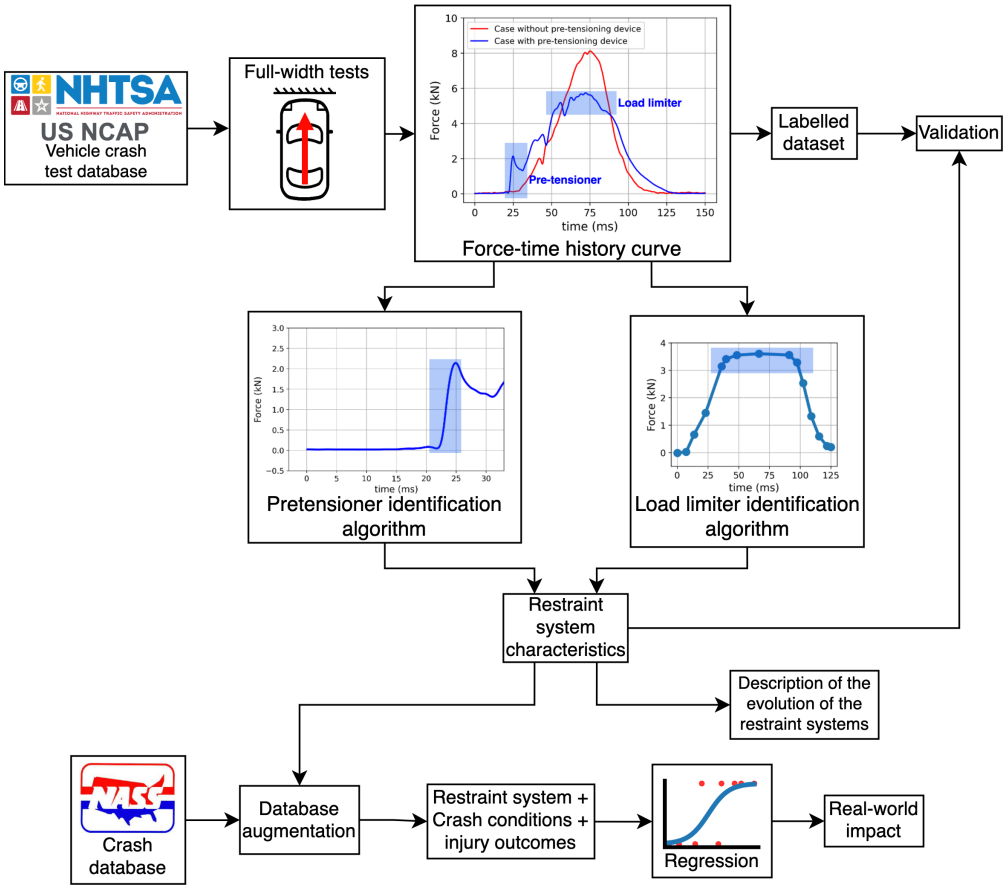


Figure 2.2: Flowchart illustrating the methodologies employed in this chapter.

2.2.1 Characterisation of pre-tensioning and load-limiting devices in the restraint system

Data source

Data were gathered from NHTSA's vehicle crash test database, including vehicles tested from 1980 onwards. Crash tests were filtered to retain the 56 km/h full-width frontal tests carried out under the NCAP program. The metadata and ISO-MME files of these crash tests were downloaded from the NHTSA webpage (NHTSA, 2023) to be used to develop characterisation algorithms of the restraint system used in the tested vehicles.

Force-time history curves measured at the driver's and passenger's lap and shoulder belts were obtained from the ISO-MME files. The channel code and comments in these files were used to identify where the force-time history curve was measured in the tests. Instrumentation metadata and the comment section of the ISO-MME files were used to look for keywords suggesting questionable and lost data or channel failure, and these readings were then disregarded. The retained seat-belt force signals were re-sampled to 10 kHz using a linear interpolation method.

Validation subset

A subset of randomly selected 60 % of the seat-belt force-time history curves was chosen to serve as the validation subset of the characterisation algorithms. The curves in the validation subset were inspected manually by a group of 9 people to determine the characteristics of the restraint systems in the validation subset. These individuals were finalising the second year of the Master's program on Mobility and Safety at Universidad Pontificia Comillas, and all of them had completed several courses on injury biomechanics, restraint systems and crashworthiness. In addition, specific instructions about the procedure were provided to the group, and further assistance was provided when needed. The instructions defined what characteristics and how they should be classified according to the time-history signal of the seat-belt forces. These instructions are shown in Appendix B.

As the actual configuration of the pre-tensioning and load-limiting devices used in vehicles was not available, after the aforementioned manual inspection was completed, the identified values were reviewed to prevent misclassifications. These values were then used to tune the identification algorithms and to assess their performance in identifying the presence and characteristics of pre-tensioner and load-limiting devices.

Pre-tensioner characteristics: time-to-fire and peak force

An algorithm was developed to identify three characteristics of the pre-tensioner: whether a pre-tensioner was present, the magnitude of the pre-tensioner force, and TTF. The lap- and shoulder-belt force-time history curves were analysed independently to identify these characteristics in each section of the seat belt. This analysis was restricted to the first 30 ms of the crash. A CFC 180 filter was applied to the signals to eliminate undesired noise.

A sliding time window was applied to each force-time history curve to measure the peak-to-peak (PTP) force difference inside the time considered by the sliding window. The application of this technique resulted in a second curve, which showed the PTP difference in the time window, as illustrated in Figure 2.3. The time window analysed in the figure is shown as a shadowed red rectangle in the original force-time history curve. The PTP force in this period was plotted at the time represented by the red line in the resultant curve on the right subfigure. A negative value was obtained if the highest value was observed before the lowest

value. The widths of the sliding windows were between 0.5 and 5.0 ms, with increments of 0.5 ms. The sliding window width was selected to maximise the pre-tensioner identification performance, i.e., the F1-score described below. The criteria for selecting the sliding window width will be explained in detail later in this chapter. The PTP force-time history curve was used to identify the maximum value and time of the peak.

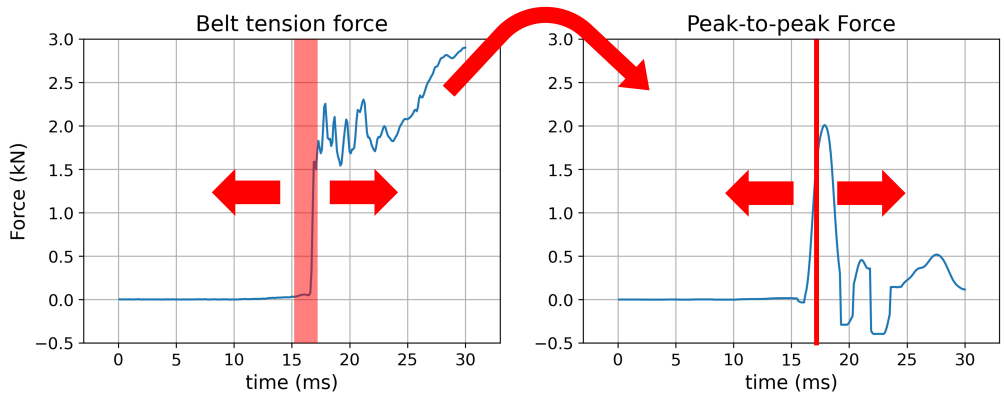


Figure 2.3: In red, an example of the sliding window applied to the first 30 ms of a force-time history curve. The plot on the left side shows the original force-time history curve, and the plot on the right side shows the result of applying the sliding window.

A binary classificatory algorithm was then used to classify the systems into two basic categories: restraint systems with and without a pre-tensioner. For this algorithm, a threshold was defined, which was used to classify those curves with a maximum value higher than the threshold as a restraint system with a pre-tensioner device affecting that seat-belt section. The force-time history curves in the validation subset were used to find the best threshold for each sliding window width.

Various metrics, such as recall, precision, accuracy and the F1-score, were utilised to evaluate the effectiveness of the aforementioned binary classification algorithm. In this study, we defined the presence of a pre-tensioner as a positive case. Recall (Ting et al., 2010), also known as sensitivity, is the proportion of true positive cases to total positives (i.e. cases with a pre-tensioner that were correctly classified as such by the algorithm over the total number of restraint systems with pre-tensioners). Precision (Ting et al., 2010), or confidence, is the proportion of correctly classified positive cases among all predicted positive cases. Accuracy (Sammut et al., 2010a) measures the proportion of all positive or negative cases that the algorithm correctly classified. The F1-score (Sammut et al., 2010b) measures the algorithm's overall performance, considering precision and recall. This score is the harmonic mean of precision and recall, providing a single metric that balances these two measures and is calculated with the following equation:

$$F1\text{-}score = \frac{2 \cdot Recall \cdot Precision}{Recall + Precision} \tag{2.1}$$

The F1-score was maximised for each sliding window to define the threshold force for detecting a pre-tensioner. This metric was different for each sliding window. The optimal window width and the force threshold were selected to maximise the F1-score across all window widths.

After classifying the force-time history curves between those without and with a pre-tensioner pulling on the belt section, the TTF and the force of the pre-tensioning device were

determined for the curves included in the second group. A similar process as above was used to identify these parameters, but in this process, the criterion used to identify the TTF and the pre-tensioning force in the data was the minimisation of the means absolute error (MAE) between the algorithm and the validation subset.

Note that as the criteria for choosing the optimal sliding window width were different according to the characteristics under study, three widths were obtained as the optimal ones.

The location of a pre-tensioner can be at the retractor, buckle or anchor plate, and multiple pre-tensioner devices could be simultaneously used for a vehicle occupant. Although the algorithm determines if a pre-tensioning force was observed at the lap or shoulder belt, the actual location of the pre-tensioning device was not determined. Thus, the results are shown in terms of the belt section where a sudden increase of the tension force (i.e., a pre-tension) was observed, which could be either the lap belt, shoulder belt or both.

Load-limiter characteristics: force level

A second algorithm was developed to detect and characterise load-limiting forces from the shoulder-belt force-time history curves. This process was divided into two phases: first, the shape of the force-time history curve was characterised; second, the load-limiting device characteristics were extracted from the previous step.

Before applying the algorithm, the curves were filtered with a CFC 180. After filtering, the curves were re-sampled using the Visvalingam-Whyatt algorithm (Visvalingam et al., 1993) to obtain a similar shape curve with fewer points, which may not have resulted in evenly distributed points in time. The simplified force-time history curve points were used to obtain groups according to the point distance in the force and time axes, and these points were grouped to obtain the load-limiting force levels. The outcome of this process resulted in three categories: no force limiting in the restraint system; a single-stage force limiting feature in the restraint system (when only one group of points maintaining similar distance between them was found); or a dual-stage force limiting feature in the restraint system (when two groups of points with similar distance were identified).

Specifically, the requirements so that a group of points was considered to have a similar distance were the following:

- The PTP force between the group points must be smaller than 1.4 kN.
- The group points must be obtained from a continuous time series, i.e. all points from the simplified force-time history curve between the minimum and maximum time of the group must belong to the group.
- The average force of the group points must be higher than 1.5 kN.
- The duration of the group must be longer than 15 ms.

After applying the above conditions, it was noted that the algorithm had difficulties identifying the first load-limiting force in restraint systems with dual load-limiting levels. However, it correctly identified the second stage of load limiting. Thus, a time-series K-means clustering method was used to identify, based on their shape, which of the original curves could be associated with a dual-stage load-limiting device. The clustering metric used to measure the distance between the time series was the dynamic time warping (DTW) distance. The number of clusters was defined using the elbow method (Syakur et al., 2018), which involves plotting the sum of distances of samples to their closest cluster centre against the number

of clusters to identify a point where the rate of decrease sharply slows. This "elbow" point indicates an optimal balance between minimising variance within clusters and avoiding overfitting by using too many clusters. The cluster corresponding to the dual-stage load-limiting devices was then used to enhance the predictions of the first stage of load limiting.

Then, a new method was developed to characterise the load-limiting devices that were classified into the cluster associated with dual-stage load-limiting devices. The requirements to identify the load-limiting stages were obtained by applying a genetic optimisation algorithm. The goal of the algorithm was to maximise the F1-score in the identification of the load-limiting characteristics. The algorithm found the optimal set of values (requirements) for a set of performance characteristics of the load-limiting device that achieved this goal. The genetic algorithm required 150 generations with a population of 30 combinations of values to find the following combination that maximised the F1-score:

- The PTP force between the group points must be smaller than 0.79 kN.
- The group points must be obtained from a continuous time series, i.e. all points from the simplified force-time history curve between the minimum and maximum time of the group must belong to the group.
- The average force of the group points must be higher than 1.36 kN.
- The duration of the group must be longer than 5.45 ms.

With the previous variables, it was noted that the algorithm resulted in false positives for some of the load-limiting devices in the sample. These were associated with two specific cases: first, cases in which the algorithm identified the section of the curve right after the firing of the pre-tensioner as a load-limiting stage; and second, cases in which a stable seat-belt force during the rebound of the occupant was identified again as a load-limiting stage. The first issue was solved by discarding potential load-limiting stages starting earlier than 5 ms after the pre-tensioner TTF. The second one was avoided by filtering out those cases in which the second load-limiting stage started later than t=75 ms.

The above method is summarised in Figures 2.4, 2.5 and 2.6, which shows examples of the identification of load-limiting characteristics. Figure 2.4 shows an example of a force-time history curve with no load-limiting device. The first figure shows the filtered force-time history curve, followed by the result of the simplification algorithm. No load-limiting device was found for this example, as no group fulfilled the abovementioned requirements. Figure 2.5 shows an example of a force-time history curve with a 1-stage load-limiting device. The simplified force-time history curve shows a constant force of around 3.5 kN. The algorithm identified this load-limiting force, which was represented in the red-shaded area. Figure 2.6 shows an example of a force-time history curve with a two-stage load-limiting device. The simplified force-time history curve shows two stages of the load-limiting device. The algorithm identified the first and second stages, shown as the red-shaded area, with a constant force of around 5.1 kN and 3.3 kN, respectively.

The performance of the load-limiter identification algorithm was also measured by using recall, precision, accuracy, and F1-score.

2.2.2 Description of the evolution of pre-tensioning and load-limiting devices over the last 40 years

A descriptive analysis was conducted to examine how the characteristics of pre-tensioner and load-limiting devices in vehicles used in NCAP full-barrier crash tests have evolved over the

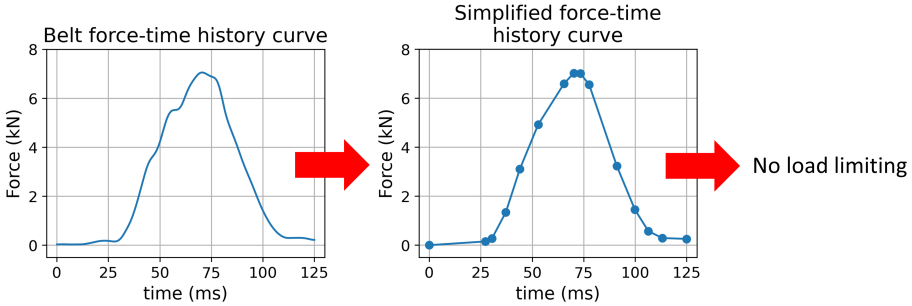


Figure 2.4: Example of applying the load-limiting characteristics identification algorithm to a case with no load-limiting device.

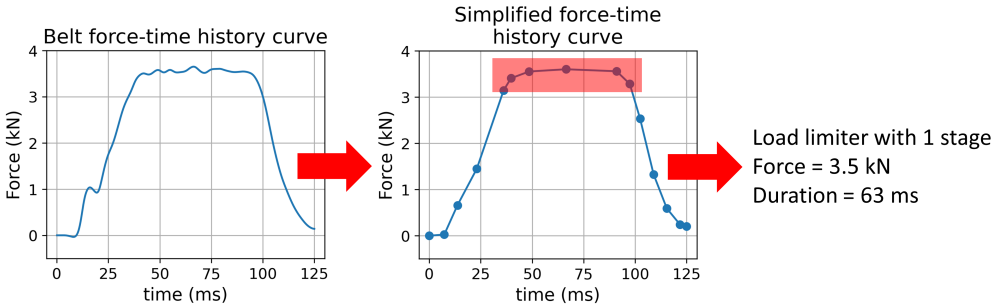


Figure 2.5: Example of applying the load-limiting characteristics identification algorithm to a case with a single-stage load-limiting device.

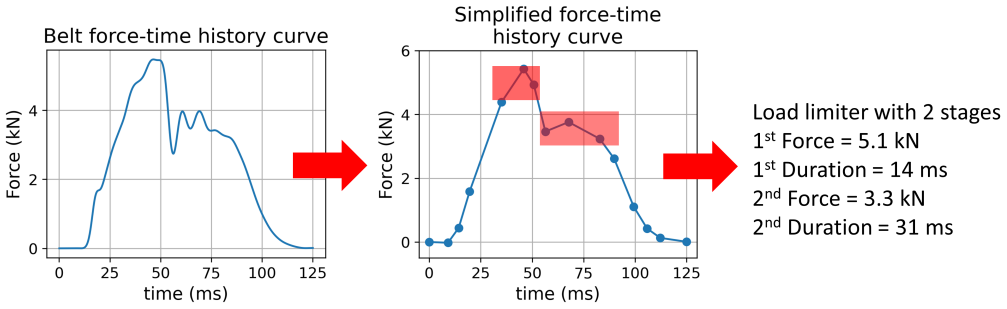


Figure 2.6: Example of applying the load-limiting characteristics identification algorithm to a case with a dual-stage load-limiting device.

past 40 years. The analysis focused on the adoption of these devices for the driver and front-seat passenger positions. A year-by-year ratio was calculated to compare the number of vehicles with pre-tensioning/load-limiter devices to those without, based on the vehicle tests from the specific year. The analysis looked at the adoption of pre-tensioning devices for each occupant position and additionally discriminated devices that pull in the shoulder- and lap-belt sections. Changes in restraint system characteristics were also analysed, including the force and TTF of the pre-tensioning device and the force of the load-limiter device. Furthermore, the evolution of pre-tensioning force and TTF was assessed regarding each belt section, excluding the occupant position. The analysis also considered load-limiter devices with two stages, examining the adoption of such devices and the correlation between the forces used for the first and second stages of the load-limiter devices.

2.2.3 Influence of restraint system characteristics on AIS 2+ and AIS 3+ real-world injuries

Data source

Data were compiled from the NASS CDS for the years 2000–2015 (NHTSA, 2024). NASS CDS data are selected based on national-level sampling to represent the frequency of crashes occurring within each of the states of the US NASS CDS includes police-reported crashes in which at least one involved vehicle is towed away due to damage, and therefore, it is skewed towards collecting more severe crashes. Then, national representativity is achieved by assigning weight factors to each of the approximately 5,000 crashes collected yearly. For each of the collected cases, the database includes information about the vehicle/s involved in the crash (including the vehicle identification number (VIN) that can be used to obtain the maker ID, the model ID, the body type of the vehicle, and the year of manufacturing), a crash reconstruction, the contributing factors from the scene of the collision as well as detailed injury data from the victims. Regarding injuries, beginning in 2010, these were coded in the NASS CDS using both the 1998 and 2008 versions of AIS. For consistency, the analyses included in this manuscript were performed using the AIS 1998 codes.

As for the detailed information regarding the characteristics of restraint systems, this information was included using the characteristics obtained with the methodology from the previous subchapter, referred to as the NCAP-testing database forward. NCAP tests between 1990 and 2015 were used to aggregate the NASS CDS data. The aggregation process used information about the vehicle and occupant position. The information about the vehicle was retrieved from the VIN decoder provided by NHTSA (NHTSA VIN DECODER, 2024) using the VIN in the NCAP-testing and NASS CDS databases. Four parameters from the vehicle were used: vehicle maker, model, body type, and year of manufacturing. Vehicles that were not present in both databases were excluded from the analysis.

The following steps were performed to aggregate each vehicle's restraint characteristics into the NASS CDS database. First, the NCAP-testing database was filtered to obtain those vehicles with the same maker ID, vehicle ID, and body type present in the NASS CDS database. Second, this list of cases from the NCAP-testing database was sorted by year of manufacturing. Third, the year of manufacturing of the vehicle involved in the crash (NASS CDS database) was used to obtain the restraint system characteristics. In the cases in which the manufacturing year of the vehicle did not match the two datasets, the nearest previous test from NCAP was used. Some cases (6 % of all possible combinations) resulted in more than one vehicle tested with the same identifiers. In this case, it was assumed that the vehicle was equipped with the most advanced restraint system found among the possible options (for

example, if only one of the coincident vehicles had a pre-tensioner, it was considered that all of them had it). Lastly, the occupant position (i.e., whether the occupant was the driver or front-seat passenger) was also used to retrieve the restraint system characteristics.

Inclusion criteria

NASS CDS data were retrieved for belted drivers and front-seat passengers who were at least 16 years old at the time of the crash. Vehicles included in the analysis were restricted to those built-in or after 1990. Only frontal collisions were considered (defined as those with Principal Direction of Force (PDOF) between 11 and 1 o'clock), excluding rollover. From the cases meeting the inclusion criteria above, several variables were chosen to be used as covariables in the analysis: occupant's sex, height, weight, and age, vehicle position of the occupant (driver/front seat passenger), vehicle wheelbase, vehicle age, model year, longitudinal delta-v, PDOF and whether the frontal airbags had deployed. The severity of head and thoracic injuries was also obtained and classified as a dichotomous variable, indicating whether the injury severity was AIS 2+ or AIS 3+ for each body region.

As for the information regarding the restraint systems, the tests included in the NCAP-testing database were used to retrieve whether the restraint system of a specific vehicle incorporated a pre-tensioner (dichotomous variable) and/or load-limiting features. Force limiting (LL) was coded as a categorical variable with three levels: no load limiter, low load-limiter force (≤ 4.5 kN), and high load-limiter force (4.5 kN $< LL \leq 7$ kN). The threshold to classify the load-limiter force as low and high was calculated using the database's 50th percentile of the load-limiting forces.

It should be noted that the pre-tensioner and load-limiting forces were obtained using the transducers placed at the lap and shoulder belt in a crash test. Therefore, the actual force at the pre-tensioner and load-limiting device may differ from the values used in this study.

Analyses

Descriptive statistics, along with univariate and multivariate logistic regression models, were developed to estimate the effects of independent risk factors on the likelihood of sustaining AIS 2+ and AIS 3+ injuries in the whole body, as well as specifically in the head-face-neck (HFN) region and the thorax (THO). The effects of independent risk factors on the likelihood of a fatal outcome were also calculated. In the univariate analysis, variables that demonstrated a significant level of $p < 0.1$ were selected as candidates for inclusion in the multivariate models. These candidate variables were identified in the previous subsection.

After performing the multivariate regressions, the included variables were checked for multicollinearity using the Variance Inflation Factor (VIF). Only variables with a $VIF < 2$ were retained in the final models, though models with variables having $VIF < 1.5$ were also tested. After eliminating collinear variables, multiple models were calculated, with those showing the lowest Akaike Information Criterion (AIC) values being selected as the best predictors of the outcomes.

In total, seven outcomes were evaluated using multivariate logistic regression models: the likelihood of sustaining a fatal injury, as well as the risk of AIS 2+ and AIS 3+ injuries for the overall body, the head-face-neck region, and the thorax.

2.3 Results

2.3.1 Characterisation of pre-tensioning and load-limiting devices in the restraint system

Data source

The information from 1,318 analysed crash tests was downloaded from the NHTSA webpage. The resulting number of force-time history curves examined was 5,138, obtained after discarding those curves in which measurement errors were identified. The retained curves corresponded to 1,300 and 1,317 force vs. time curves measured at the shoulder and lap belt of the driver position, and 1,237 and 1,284 force vs. time curves measured at the shoulder and lap belt of the front-seat passenger.

Out of the total 5,138 curves, 3,081 force vs time curves were manually labelled by the mentioned group of 9 individuals to describe the pre-tensioner and load-limiting devices. These curves were distributed evenly across the initial categories found in the data: driver seat-belt force-time history curves (815 and 740 measured at the shoulder and lap belt, respectively) and front-seat passenger seat-belt force-time history curves (805 and 721 measured at the shoulder and lap belt, respectively). The manually labelled curves were used as the validation subset for the results predicted by the algorithms.

Pre-tensioner identification: time-to-fire and force

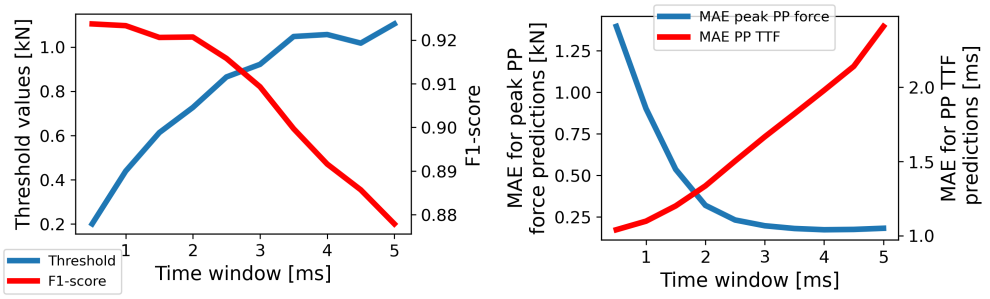
An algorithm was developed to identify three seat-belt force-time history curve characteristics: whether a pre-tensioner was incorporated in the restraint system, and if that is the case, the pre-tensioning force and TTF.

Three different sliding window widths were selected to predict the three characteristics under study. Figure 2.7 shows the F1-scores and mean absolute error (MAE) used to choose the optimal window widths. Subfigure (a) shows the threshold force and the resulting F1-score for predicting a pre-tensioner presence as a function of the window width. The maximum performance was obtained for a window width of 1 ms and a threshold of 0.51 kN (corresponding to the maximum F1-score). This method obtained a recall, precision, accuracy and F1-score of 95.0 %, 89.7 %, 95.5 % and 92.3 %, respectively. Subfigure (b) shows the MAE obtained in the identification of the pre-tensioner (PT) force and time-to-fire (TTF). The minimum MAE for the PT force was obtained using 4 ms as sliding window width, which resulted in an error of 0.17 kN. The minimum MAE for the PT TTF was obtained using 0.5 ms as sliding window width, which resulted in an MAE of 1.09 ms.

Load-limiter identification: force level

Figure 2.8 shows the results obtained after applying the clustering algorithm to the original data. These clusters were obtained with 20 % of the data. The elbow method was used to select the number of clusters and resulted in a total of seven, which are shown in Figure 2.8.

The first cluster showed curves with a sustained force over time, suggesting that load-limiting devices with one stage were grouped in this cluster. The second and third clusters exhibited a similar shape but did not clearly show the use or not use of a load-limiting device. The fourth cluster showed a first peak followed by a plateau, resembling a load-limiting device with two stages. The fifth and sixth clusters showed data loss due to measurement errors, while the seventh cluster exhibited a peak force of around 110 ms. Therefore, the analyses did not include the force-time history curves in the last three clusters. It should be noted that



(a) (b)

Figure 2.7: Performance of the algorithm to characterise the pre-tensioner. a) F1-score, in red, for different sliding window widths and their respective force thresholds, in blue. b) MAE by applying different sliding window widths for the identification of the PT force in blue and PT TTF in red.

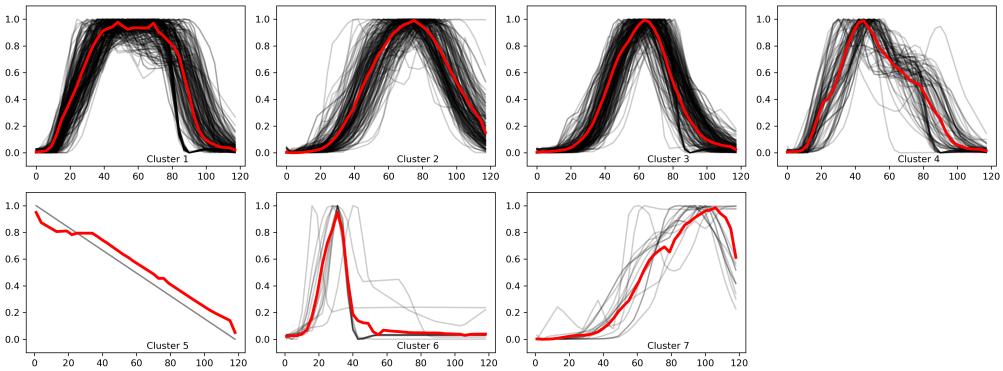


Figure 2.8: Time-series cluster centres that were obtained by applying the k-means clustering algorithm in red, and some of the force-time history curves that were identified as part of such clusters in grey.

the force-time history curves that were classified within the fourth cluster were identified as potential restraint systems featuring a dual-stage load-limiting device and used to enhance the algorithm's prediction as described in the section 2.2.1, above.

Applying the classifying algorithm to the original data resulted in three categories: restraint systems without load-limiting devices, those incorporating single-stage load-limiting devices, and those incorporating dual-stage load-limiting devices. Table 2.1 shows the results of the accuracy, F1-score, precision, and recall obtained after comparing the predictions given by the algorithm and the data manually identified in the validation set.

Table 2.1: Detection performance of load-limiting devices.

Load-limiting device type	Accuracy	F1-score	Precision	Recall
Presence of a load-limiting feature	0.891	0.898	0.941	0.859
Double-stage load-limiting feature	0.933	0.765	0.872	0.681

The algorithms used to identify and quantify the force-limiting characteristics performed slightly worse than those used in analysing the pre-tensioning devices. It should be noted that 75 % of the miscategorised cases belonged to vehicles manufactured between 1980 and 2005. Furthermore, 57 % of these miscategorised cases corresponded to vehicles manufactured before 1995.

The MAE of the prediction of the load-limiter force was assessed when the use of a load-limiting device was correctly detected. If a dual-stage load-limiting device was detected, only the first stage of the load-limiting device was used for the assessment. The MAE for this prediction was 0.23 kN. If a dual-stage load-limiting device was detected, the MAE was calculated using the force of the second stage. In this case, the MAE was 0.16 kN.

2.3.2 Description of the evolution of pre-tensioning and load-limiting devices over the last 40 years

Force-time history curves were analysed from vehicles over the last 40 years, which were tested in the NCAP full-barrier crash tests. Specifically, a total of 5,138 force-time history curves were studied from 1,355 vehicles to identify the characteristics of pre-tensioner and load-limiting devices.

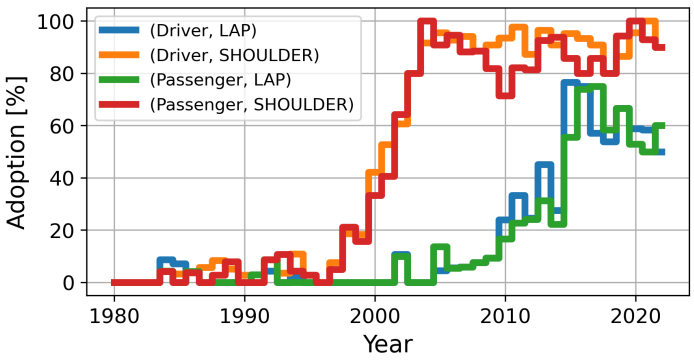
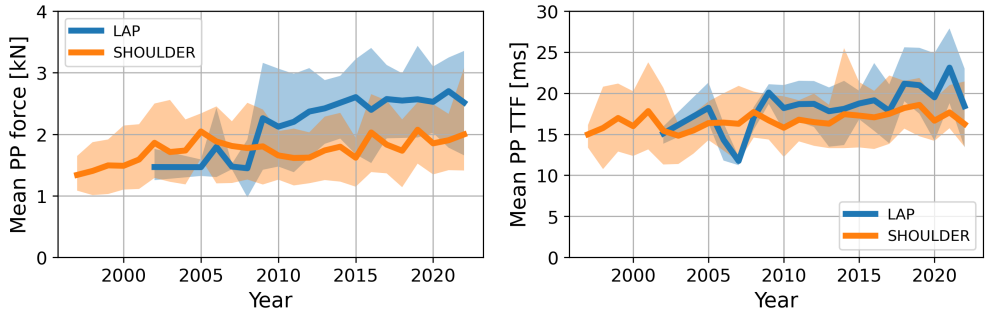


Figure 2.9: Adoption of pre-tensioners for different vehicle locations.

Figure 2.9 shows the adoption of pre-tensioner devices for restraint systems used for the driver and passenger. Furthermore, this analysis considered whether pre-tension was ob-

served for the lap or shoulder belt. The red and orange lines show the adoption of these devices in the shoulder belt for the driver and passenger, respectively.

The adoption of these devices was lower than 20 % for vehicles tested before 1995 and gradually incremented afterwards. Pre-tensioners for the shoulder belt were used in more than 80 % of the vehicles tested after 2004, and this level has been maintained up to the present. The adoption of lap-belt pre-tensioners has been slower, and began after 2002. Adoption levels between 50 % and 80 % have been found for 2015 onwards. Similar adoption levels have been found for these devices for the driver and front-seat passenger positions.



(a) (b)

Figure 2.10: Mean pre-tensioner (PP) force (subfigure a) and TTF (subfigure b) at the lap and shoulder belt. The shadowed area shows the range between the 10th and 90th percentiles of the data observed each year.

Figure 2.10 shows the mean forces and TTFs used in the pre-tensioner devices. Subfigure (a) shows the mean pre-tensioner force at the lap and shoulder belt in vehicles from 1996 onwards. The forces used in pre-tensioner devices for the shoulder belt have maintained similar values in the time analysed, ranging from 1.5 to 2.0 kN approximately. In contrast, forces used in pre-tensioners for the lap belt have increased from 1.0 to 2.5 kN over the last 20 years. Subfigure (b) shows the mean pre-tensioner TTF at the lap and shoulder belt. Similar values have been identified for these devices in the analysed period, ranging from 15 to 20 ms. However, a small difference in the mean TTF was observed for the lap- and shoulder-belt pre-tensioners from 2010 to 2022. Specifically, the TTF for the lap-belt pre-tensioner was slightly later than that for the shoulder-belt pre-tensioner, with a difference of approximately 2 ms.

Figure 2.11 shows the adoption of load-limiting devices for restraint systems used for the driver and passenger. The red and orange lines show the adoption of these devices for the driver and passenger, respectively. The adoption of these devices was below 20 % for vehicles tested before 1998 and gradually incremented afterwards. Load-limiting devices were present in over 80% of tested vehicles after 2003. Similar adoption levels have been observed for driver and passenger restraint systems.

Figure 2.12 shows the mean forces used in the load-limiting devices in vehicles from 1996 onwards. Forces used in load-limiting devices have decreased in the time analysed, starting at 5 kN and finishing at the range between 3 and 4 kN. This trend was observed for devices used at both driver and front-seat passenger positions. A difference in the mean force was observed for load-limiting devices used for the driver and passenger positions after 2010. A mean 0.6 kN higher force was observed for load-limiting devices at the driver position than for devices at the passenger position.

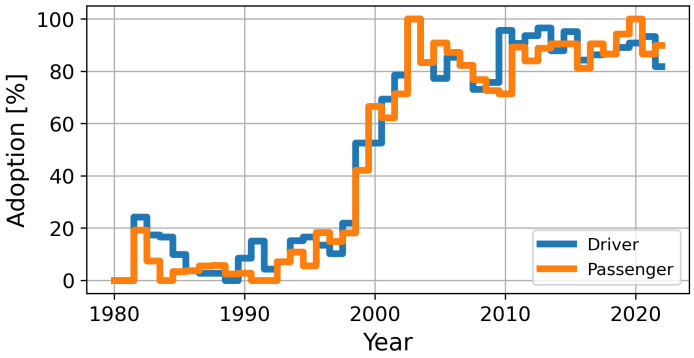


Figure 2.11: Adoption of load-limiter devices for the driver and front-seat passenger.

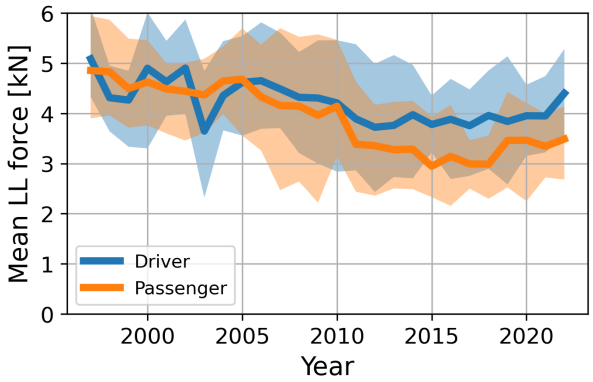


Figure 2.12: Mean load-limiter force for devices used at the driver and front-seat passenger positions. The shadowed area shows the range between the 10^{th} and 90^{th} percentiles of the data observed by year.

Figure 2.13 shows the adoption of dual-stage load-limiting devices and the relation between the forces in the first and second stages of the device. This figure shows that the adoption of this variation of the load limiter started when the load-limiting devices were present in 80 % of the vehicles tested, between 2000 and 2005. Although a similar adoption was observed for both front-seat occupants, this device was identified in more cases for the driver than for the front-seat passenger between 2010 and 2020. Figure 2.14 shows the relation between the first and second-stage forces observed for the dual-stage load-limiter devices. A linear regression was performed to fit the relation between the forces used in the first and second stages of the load-limiting devices, shown in the figure, resulting in an R^2 equal to 0.495. A quantile regression for the 25th and 75th quantiles was also performed, which is shown using the shadowed area. The mean width of this area was 0.66 kN.

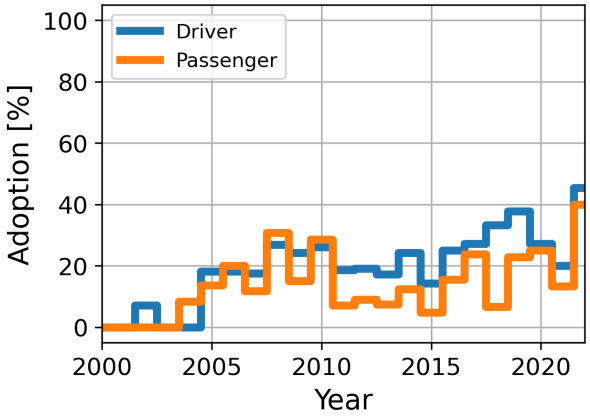


Figure 2.13: Adoption of dual-stage load-limiter devices for the driver and front-seat passenger.

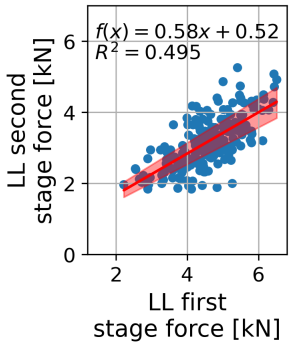


Figure 2.14: Relation between the first and second-stage forces observed for the dual-stage load-limiter devices. The red line shows the result of a linear regression, and the shadowed area shows the 25th and 75th quantiles.

2.3.3 Influence of restraint system characteristics on AIS 2+ and AIS 3+ real-world injuries

Descriptive analysis

Table 2.2 shows the distribution of the covariables considered in the study. Categorical variables are described using the number of cases ("N") where this variable was true and its per-

centage regarding the whole sample. Continuous variables are described by providing their mean and standard deviation (SD).

A total of 14,848 occupants (12,172 drivers and 2,676 front-seat passengers) were included. The age distribution within the sample was 39.1 ± 18.2 years-old (mean ± SD), with 48.3 % of the sample being male occupants. The mean age of the vehicles involved was 5.1 years-old, although the standard deviation was 3.6 years old.

Among these occupants, a lap-belt pre-tensioning device was present in 700 cases (4.7 %), and a shoulder-belt pre-tensioning device was included in 5703 cases (38.4 %). The mean value of the pre-tensioner force was similar regardless of the location of the pre-tensioner (1.8 ± 0.6 kN and 1.6 ± 0.5 kN, at the lap belt and shoulder belt, respectively). Two groups were created to analyse the effect of load-limiting devices on the injury outcomes: a group in which the load-limiting force was equal to or smaller than 4.5 kN (LL_low) and a second group in which the load-limiting force was greater than 4.5 kN but lower than 7 kN (LL_high). Cases which included other load-limiting forces were discarded. A total of 7,020 cases were observed where the occupant used a restraint system with a load-limiting device, in which 3,343 and 3,677 cases had a low and high load-limiting force, respectively. The frontal airbag was deployed in approximately 66.1 % of the cases included.

As for the injuries, only 1.2 % of the occupants considered in the database had a fatal outcome. Almost 17.4 % of the occupants sustained at least an AIS 2+ injury, and 8.1 % sustained at least an AIS 3+ injury considering the whole body. As for thoracic and head injuries, 5.2 % of occupants sustained AIS 2+ injuries in the head and in the thoracic regions, and 3.3 % and 2.0 % sustained AIS 3+ injuries in the thorax and head, respectively.

Multivariate regression analysis

The covariables that were finally retained in the model to estimate the likelihood of fatal, AIS 2+ and AIS 3+ injuries are shown in Table 2.3.

Fatal injury

While the deployment of the airbag alone did not show a significant effect in preventing fatal injuries in a crash, its interaction with delta-v was found to be a significant factor in reducing fatality risk. For each additional km/h, there was a reduction of 2 % in the probability of being killed in the collision. Increasing age, weight, and delta-v were also significant predictors of an increased risk of death. Being involved in an oblique collision (11 o'clock) was also significantly related to increased risk. The presence of a load-limiting device in the restraint system was found to be significant. While both lower and higher force load-limiting devices were associated with a reduction in fatal injury risk by almost 70 % and 60 %, respectively, the interaction between delta-v and a higher force load-limiting device diminished its effectiveness at higher delta-v values for this specific configuration.

Whole body

The airbag deployment was found to increase the risk of sustaining AIS 2+ and AIS 3+ injuries overall, but this effect needs to be understood with caution as the variable showing the interaction between the airbag and the delta-v was also significant. This interaction points towards a protective effect of the airbag, especially relevant at higher delta-v's. Increasing age, weight, and delta-v were also significantly related to the likelihood of injury. Impact directions that added a slightly lateral direction to the PDOF resulted in increased risk for the occupant for the two injury levels considered. Women were exposed to a greater risk of injury than men for the conditions considered in this study.

When the variables related to the pre-tensioning and load-limiting devices were included in the regression model, the model year of the vehicle was not significant anymore and, therefore, was dropped from the regression model. With respect to restraint systems with load-limiting devices, the inclusion of a load-limiting device resulted in a decreased risk of AIS 2+ and AIS 3+ injuries, but the effect was only observed when interacting with the delta-v. When the load-limiting device was in the low force category, the restraint system contributed to preventing AIS 3+ injuries, although the significance of the variable was marginal ($p < 0.1$). The model also found that older vehicles were more likely to result in occupants sustaining AIS 2+ and AIS 3+ injuries.

Thoracic region

As for the thoracic region, the airbag deployment was also significantly related to an increased risk of AIS 2+ and AIS 3+ injuries, but there was an interaction between this variable and the delta-v that changed this effect for delta-v above 32.5 km/h and 34.3 km/h for AIS 2+ and AIS 3+ injuries respectively. The effect of age, weight, and delta-v was similar to the one found when the whole body was considered. As also found for the whole-body area, impact directions different from the pure frontal direction contributed to an increased risk of thoracic injuries. It should be noted that the 11 o'clock direction, which would be a small overlap or a nearside oblique impact for the driver, resulted consistently in a higher risk of injury than a pure frontal or a far-side oblique impact for the driver. When the focus was on thoracic injuries, being a female occupant was associated with a higher risk of AIS 2+ and AIS 3+ injuries but only with marginal significance ($p < 0.1$). The driver position resulted in a significantly lower risk of injury to the thoracic area.

No other significant effects, including the presence of pre-tensioning and/or load-limiting devices in the restraint systems, were found significant in the analysis of the risk factors contributing to thoracic injuries.

Head-face-neck region

Regarding the head-face-neck body region, the airbag deployment was also significantly correlated to the existence of AIS 2+ and AIS 3+ injuries, but again, the significant interaction term between airbag deployment and delta-v would revert this relationship above a certain delta-v. Age and delta-v were again significantly related to increasing the likelihood of injuries also in the HFN area. As for the direction of force, it was found consistently that the 11 o'clock impact direction was significantly related to the occurrence of head injuries. In the case of HFN, the analysis did not find any significant difference in the injury likelihood between men and women.

Table 2.2: Distribution of the variables analyzed for the augmented NASS CDS database.

			NASS CDS			
			14,848 cases			
			N	%	Mean	SD
Crash	delta-v	km/h	-	-	23.3	12.9
	PDOF	11	3,131	21.1	-	-
		12	8,790	59.2	-	-
		1	2,927	19.7	-	-
Vehicle	Curb weight	kg	-	-	1,573.0	365.4
	Wheelbase	kg	-	-	279.7	27.4
	Model year	-	-	-	2001	5.4
	Vehicle age	years	-	-	5.1	3.6
Restraints	Frontal airbag	deployed	9,811	66.1	-	-
	Pre-tensioners:					
	Lap belt	present	700	4.7	-	-
	Shoulder belt	present	5,703	38.4	-	-
	Lap belt TTF	ms	-	-	16.3	4.3
	Shoulder belt TTF	ms	-	-	16.2	3.4
	Lap belt force	kN	-	-	1.8	0.6
	Shoulder belt force	kN	-	-	1.6	0.5
	Load limiter:					
	Present	yes	7,020	47.3	-	-
	Low force	yes	3,343	22.5	-	-
	High force	yes	3,677	24.8	-	-
Occupant	Age	years	-	-	39.1	18.2
	Height	cm	-	-	170.6	10.6
	BMI	kg/m^2	-	-	26.5	6.5
	Sex	male	7,174	48.3	-	-
	Driver	yes	12,172	82.0	-	-
Injuries	Faltal		174	1.2	-	-
	Head	AIS 2+	779	5.2	-	-
		AIS 3+	294	2.0	-	-
	Thorax	AIS 2+	767	5.2	-	-
		AIS 3+	491	3.3	-	-
	Overall	AIS 2+	2,588	17.4	-	-
		AIS 3+	1,204	8.1	-	-

Table 2.3: Odds ratio (OR) and significance (p) of the independent variables included in the multivariate models to estimate the likelihood of AIS2+ and AIS3+ injuries.

		Fatal injuries		Whole body injuries				Thoracic injuries				Head-face-neck injuries			
				AIS2+		AIS3+		AIS2+		AIS3+		AIS2+		AIS3+	
Variable	Reference	OR	p	OR	p	OR	p	OR	p	OR	p	OR	p	OR	p
Crash															
delta-v	-	1.09	***	1.1	***	1.11	***	1.11	***	1.11	***	1.08	***	1.09	***
PDOF	12 o'clock														
1 o'clock		–	–	1.28	***	1.28	**	1.51	***	1.5	**	1.22	#	1.4	#
11 o'clock		2.04	***	1.35	***	1.53	***	1.69	***	1.88	***	1.5	***	1.93	***
Vehicle															
Weight	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Wheelbase	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Model year	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Age	-	-	-	1.03	***	-	-	-	-	-	-	-	-	-	-
Occupant															
Age	-	1.04	***	1.02	***	1.03	***	1.05	***	1.04	***	1.01	***	1.03	***
Height	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Weight	-	1.02	***	1.01	***	1.01	***	1.00	*	1.01	**	-	-	-	-
Sex	female	-	-	0.62	***	0.75	***	0.85	#	0.82	#	-	-	-	-
Driver	passenger	-	-	-	-	0.79	**	0.77	*	0.71	**	-	-	-	-
Restraint															
Airbag depl.	not deployed	-	-	2.95	***	3.71	***	2.74	***	3.11	***	2.11	***	2.10	*
Pre-tensioners															
Lap belt	not present	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Shoulder belt	not present	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Load limiter (LL)															
Low force	not present	0.31	*	-	-	0.70	#	-	-	-	-	-	-	-	-
High force	not present	0.42	*	-	-	-	-	-	-	-	-	-	-	-	-
Combined															
delta-v * Airbag	not deployed	0.98	*	0.98	***	0.97	***	0.97	***	0.97	***	0.97	***	0.97	***
delta-v * LL															
Low force	not present	-	-	0.99	**	-	-	-	-	-	-	-	-	-	-
High force	not present	1.02	*	0.98	*	-	-	-	-	-	-	-	-	-	-

*** $p<0.01$; ** $0.01 \leq p<0.05$; * $0.05 \leq p<0.1$; ∧ $p>0.1$

2.4 Discussion

2.4.1 Characterisation of pre-tensioning and load-limiting devices in the restraint system

An algorithm was developed to analyse the force-time history curves measured in the seat-belt to obtain the restraint system characteristics used in the analysed vehicle. The algorithms were trained by comparing their predictions with a subset of 3,081 curves from which the values of the peak force and TTF of the pre-tensioner and the stage/s of load limiting had been previously identified by a group of nine people specifically trained to identify these features. To increase the robustness of the method, the authors of the study confirmed the results from the inspection of the curves. Some particular curves that were more difficult to assess were analysed with the help of a restraint system specialist.

The performance of the algorithm used to identify the presence of a pre-tensioning device on the restraint system was high according to the criteria of recall, precision, accuracy and F1-score (95.0 %, 89.7 %, 95.5 % and 92.3 %, respectively) used to assess the quality of the predictions. The mean value of the pre-tensioner peak force identified by the algorithm was within 10 % of the mean force identified by the trained people who assessed the validation subset. The prediction of the presence of a load-limiting device was more challenging, especially in restraint systems incorporating a dual-stage load-limiting device with a lower F1-score of 0.765.

The restraint system characteristics obtained from the analysis of the NCAP-testing database matched the information available in the published literature. Mercedes-Benz introduced pre-tensioners in cars in 1983 (Mitzkus et al., 1984), as it has also been identified in our analysis. Renault started to install restraint systems with load-limiting devices in 1995. Experts in the field consider load-limiting devices as the most important improvement of the three-point seat belt (Håland, 2006). It is interesting to observe that the adoption rate of load limiters has been much faster than in the case of pre-tensioners.

Another relevant consideration is how the peak force of lap-belt pre-tensioners has been increasing steadily over the last few years while the shoulder-belt pre-tensioner peak force has remained at about the same value since its introduction. The data also show that introducing pre-tensioners in the front seat follows the same trend for the driver and the passenger positions.

It is interesting to realise that the pre-tensioner introduction rate started to increase in the late 1990s, reaching a wide adoption in the mid-2000s. Despite the later development of load limiters compared to pre-tensioners, they were broadly adopted at about the same time. This increasing adoption rate could be linked to three events in the late 1990s. First, NHTSA changed from reporting test results in a technical (numerical) format to an easy-to-understand five-star rating system (1994) and began the crash test program for side impacts in NCAP (1997) (Hershman, 2001). Secondly, the Insurance Institute for Highway Safety (IIHS) began evaluating frontal crash tests in 1995. Lastly, Euro NCAP released its first results in 1997. Therefore, the desire of car manufacturers to improve the safety ratings of their vehicles could be linked to the following:

- Including more conditions in which the vehicles were tested could push for restraint systems with good performance across different crash scenarios.
- The beginning of consumer rating programs to induce improvements in vehicle safety (Samaha et al., 2011).

- Introducing a consumer-friendly presentation of the results could allow the consumer to compare the safety performance of different vehicles quickly and easily.

Regarding restraint system characteristics, load-limiting forces for drivers and front-seat passengers ranged between 3.5 and 6 kN, with small differences (<0.5 kN) in the mean values during this period. However, after 2010, a notable divergence occurred as the mean load-limiting force for front-seat passengers decreased from 4 kN to 3 kN. This trend may be associated with changes in crash test protocols, specifically the adoption of the Hybrid III 5th percentile female dummy as a surrogate for front-seat passengers in 2010.

2.4.2 Influence of restraint system characteristics on AIS 2+ and AIS 3+ real-world injuries

As shown, the fitment of advanced restraint systems into the American fleet started to grow linearly from < 10 % of the vehicles incorporating these features in 1996 to around 80 % of the fleet, including pre-tensioners and load limiters (Kent et al., 2007). This growth paralleled the increase in front-end stiffness of passenger cars during the same period, as reported in a review of NCAP tests (Swanson et al., 2003). Since then, the amount of specialised literature showing the benefits of these advanced restraint features in laboratory settings and computational studies has been substantial (Forman et al., 2009; 2008; Kent et al., 2001; 2007; Michaelson et al., 2008; Walz, 2004). Some early field studies found that force limiters could have been effective in the prevention of severe thoracic trauma in some selected real-world crashes (Foret-Bruno et al., 1978). But, to our knowledge, this is the first study using multivariate regression models and including a large sample of real-world cases that assesses the effect of incorporating restraint systems with pre-tensioning and load-limiting devices in the likelihood of injuries in the field. When the analysis is focused solely on fatal injuries, the inclusion of a load-limiting device in the restraint system, regardless of its specific force level, has been shown to significantly reduce the risk of death. This finding aligns with the results reported by Kahane (2013a), 2015 based on data from the Fatality Analysis Reporting System (FARS).

In this context, it is important to note that the two variables indicating the presence of pre-tensioning devices, either at the shoulder or lap belt sections, were not significant in predicting the likelihood of injuries across most logistic regression models, particularly when focusing on specific body regions. This is probably related to the strong association between the presence of pre-tensioning and load-limiting devices in modern vehicles. Despite this, we consistently retained the variable related to the load-limiting feature of the restraint system in the models, even when its Variance Inflation Factor (VIF) exceeded 1.5 (though it was always below 2). This decision was made because the load-limiting characteristic was considered especially crucial for understanding the impact of advanced restraints on occupants in real-world scenarios. However, this may introduce confounding effects, as the inclusion of a load-limiting device could reflect not only its specific impact but also incorporate the influence of other vehicle advancements, such as improvements in structural integrity.

It was observed that devices with both lower and higher load-limiting forces were associated with a reduced risk of fatal outcomes, and AIS 3+ injuries in the whole body at high delta-v's. It is also worth noting that advanced restraint systems are likely optimised for the delta-v conditions used in NCAP tests (35 mph), whereas the average delta-v in the NASS CDS sample analysed in this study was considerably lower (23.3 ±12.9km/h). Consequently, it is plausible that the impact of advanced restraint systems might be more pronounced in high-speed collisions.

The analysis performed here was limited only to frontal impacts, defined as those in which the PDOF ranged between 11 o'clock and 1 o'clock. The results suggest that even if the oblique directionality of the cases included here was minimal, this factor was always significantly associated with an increased risk of AIS 2+ and AIS 3+ injuries globally and also when focusing exclusively on the THO and HFN areas. These results agree with those presented by Meyer et al. (2015), although the authors of this study analysed only a convenience sample of real-world crashes in which the additional belt spool-out could have been associated with impacts of the occupants against the interior of the vehicles. Again, even though previous studies in the laboratory had shown that pre-tensioners could also be effective in improving how an occupant is restrained in oblique impacts (López-Valdés et al., 2016; Piqueras et al., 2022), this outcome was not found in the current work.

Several recent studies have found that newer vehicles were associated with a reduced risk of injury (Forman et al., 2019; Klinich et al., 2016; Ryb et al., 2011; Thomas, 2013). Forman et al. (2019) found that newer model vehicles (2009 and later) carried less risk of AIS 2+ and AIS 3+ injuries than older model vehicles. This finding was particularly significant for the lower extremity region and the skull fracture risk. However, their analysis did not find significant reductions in the change of AIS 3+ rib fractures despite all the technological advances implemented in the restraint systems of newer vehicles. These findings are interestingly related to the results presented here. In particular, our study did not find the significance of the vehicle model year when the variables related to the presence of pre-tensioning and load-limiting devices were included in the models.

Kent et al. (2007) showed a clear correlation between advanced restraint systems and newer vehicles in the NASS CDS sample. In fact, model year was one of the variables that produced a high VIF score in the collinearity checks performed to develop the multivariate models. Second, it results intriguing that the lack of significance of the reduction of AIS 3+ rib fractures found in Forman et al. (2019) is somehow paralleled here by finding a marginally significant relationship ($p < 0.1$) between AIS 3+ thoracic injuries and the presence of load-limiting and pre-tensioning devices. It is true that the effect of age on the risk of thoracic injuries was found to be the largest in the thorax. Therefore, it could be that the negative effect of age might still outperform the benefit of using advanced restraint systems. In any case, the question about why, if these advanced restraint systems have been proven to be so effective in the laboratory setting, their presence is not a significant factor when real-world data are analysed still needs to be fully answered.

Forman et al. (2019) also pointed out that females continued to be at greater risk than males of AIS 2+ and AIS 3+ for most of the injury types considered in their study, as it had already been found by Bose et al. (2011). The latter found that the odds of a belt-restrained female driver sustaining AIS 3+ injuries were 47 % larger than those of a male driver in comparable crash circumstances. Our study showed similar effects when controlling for age, height, weight, and the amount of load-limiting characteristics that are applied to the occupant, but only when the whole body and thorax were considered. This significance was not observed when the analysis focused on the HFN area. This result suggests that there might be benefits of including pre-tensioning and load-limiting devices in the restraint system for other body regions different from the ones analysed in this study.

The regression model included interaction terms between delta-v and airbag deployment to isolate the specific impact of airbag deployment on injury risk, acknowledging that this effect may vary across different delta-v levels. The inclusion of this interaction term proved statistically significant, highlighting that the airbag's protective effect reached a net protective effect when the crash delta-v exceeded a certain threshold. This threshold, calculated at ap-

proximately 31 km/h, marks the point where the airbag's role shifts from potentially increasing the risk of thoracic injuries to offering a protective effect. Figure 2.15 shows the relation of the distribution of airbag deployment with delta-v, suggesting the need to check for the interaction between these two variables.

These findings align with Segui-Gomez (2000), which investigated the effect of airbags on driver injuries at different severities. In their study, Segui-Gomez (2000) identified that airbag deployment began to exert a net protective effect at delta-vs exceeding 32.8 km/h for MAIS 3+ injuries across all drivers. In particular, they observed variation in this threshold by sex, with the protective effect initiating at different delta-vs for men (12.9 km/h) and women (52.0 km/h). However, the study shown in this chapter did not incorporate an interaction term involving airbag deployment, delta-v, and sex, leaving room for further exploration into how these variables might jointly influence injury outcomes.

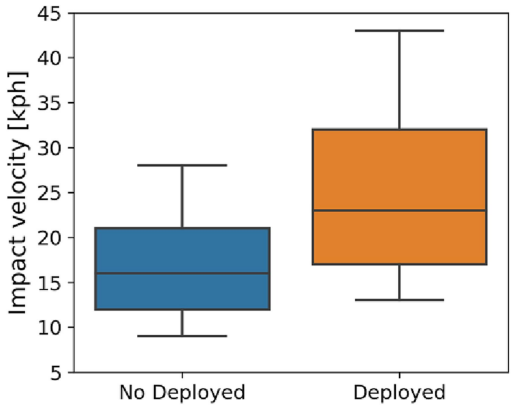


Figure 2.15: Box plot showing the distribution of the 10%–90%ile delta-v values as a function of airbag deployment

2.5 Limitations

The study has some limitations, such as the description of complex systems and the limited dataset size. The main limitation of this study is that the seat-belt characteristics are described using only the force-time history curves of the seat-belt forces, without considering other relevant variables that can affect these curves, such as slack and seat belt geometry, among others. This limitation applies to the validation subset, which was manually inspected, and the algorithm's outcome.

The identified characteristics of the pre-tensioning and load-limiting devices will present some differences from the actual configuration of these devices used in the vehicles. Although the guidelines provided for the manual inspection of the curves (included in the Appendix B) aimed to restrict the impact of this limitation, the actual configuration of the restraint system will show differences from the values presented in this study, as shown in the example case at the appendix. Factors such as the delay to fire the pre-tensioner device (i.e., the difference between the firing command and the actual firing of the device), belt slack, belt geometry, transducer position, and interactions with the airbag and seat affect the measurement of the tension force at the seat-belt. The influence of these factors cannot be quantified using the applied methodology, and therefore, they were ignored when identifying restraint system characteristics. Furthermore, this study focused on identifying if a pre-tensioning device was

present and in which sections of the belt it has influence, but the actual location of the pre-tensioning device in the lap belt could not be identified (i.e., retractor, buckle anchor point, or end-bracket anchor point).

Regarding the multivariate analysis, it should be noted that the effect of the NASS CDS weighting factors was not taken into account. However, by limiting the sample size to the raw data in NASS CDS, we were able to obtain a large enough sample to achieve statistical power and detect the significance of several independent variables. While using weighting factors could have expanded the results to a larger sample, the car model is not included in the calculation of these weighting factors. Therefore, augmenting the data in the NASS CDS database to include the load-limiters and pre-tensioner information obtained from the NCAP tests would have been questionable. The exclusion of weighting factors from the regression also influenced the generalizability of the results. Specifically, the effects of different variables on injuries are only valid within the analysed sample and cannot be extrapolated to the broader population. It's important to note that our study may be missing cases with minor injury severity. Thus, the effects presented here are likely a conservative estimation of the real-world effect of advanced restraint systems.

It is also relevant to note that the WinSmash algorithm used to calculate the delta-v in NASS CDS was updated in 2008 and later. The effects of this update were not uniform but resulted in an average increase of 8 % in the estimated delta-v. As indicated by Forman et al. (2019), in a multivariate analysis, the influence of this change would not have been relevant for the estimations of the effect of the other variables in the model.

In the case of restraint systems characteristics, the use of occupant classification systems (OCS), which affect the response of the restraint systems depending on the size of the occupants, was not considered in this study. Furthermore, the characteristics of the restraint system were identified from NCAP tests performed with a unique dummy size in each occupant position at a specific velocity. This can limit the applicability of such characteristics to other crash configurations and occupant anthropometries.

In the case of restraint systems with two-stage load-limiting devices, only the first stage was used in the multivariate analysis. This constraint was accepted to avoid having only a few cases in some of the load-limiting categories and as a sensible approach to combine single and dual-stage seat belts into one category. It should be noted that only 4 % of the analysed cases were found to have two-stage load limiters.

An interaction term between airbag deployment and delta-v was included to isolate the effects of these variables on injury risk. However, this term does not account for cases where the airbag fails to deploy due to low crash severity or an atypical crash pulse. In these instances, the absence of airbag inflation can introduce confounding effects.

2.6 Public repository of identified restraint system characteristics

The restraint system characteristics identified in this study were updated to a public repository for easy and free access by the research community, which can be accessed through the following link (Seatbelt-characteristics-repo).

2.7 Related publications

- Valdano, M., Jiménez-Octavio, J.R., López-Valdés, F.J., The effect of seatbelt pre-tensioners and load limiters in the reduction of MAIS 2+, MAIS 3+, and fatal injuries in real-world frontal crashes. Accident Analysis & Prevention. Vol. 190, pp. 107180-1 - 10718-6, September 2023.
- Valdano, M., Jiménez-Octavio, J.R., Pipkorn, B., López-Valdés, F.J., Characterisation of the features of seat-belt systems based on the analysis of large crash databases, International Research Council on Biomechanics of Injury - IRCOBI Europe 2023, Cambridge (United Kingdom). 13-15 September 2023.

3 Evaluating the effectiveness of deterministic and probabilistic approaches for predicting injury outcomes using multibody models in real-world conditions

This chapter explores the application of computer models, specifically multibody modelling, to predict injury outcomes under real-world conditions. To achieve this, the model's predictions were compared to actual injuries and injury patterns documented in real-world crashes. A multibody human body model (MB HBM), integrated with a vehicle interior, was employed to calculate the likelihood of sustaining injuries based on selected injury criteria under various impact conditions derived from real-world crash data. Additionally, a metamodel was developed to predict the likelihood of serious injuries using the computer model, which was used as a surrogate model for the computer model. Injury outcomes from real-world data, presented in Chapter 2, were compared to the estimates derived from both deterministic and probabilistic approaches. This benchmarking process evaluated the number of predicted injured occupants, the injury patterns sustained, and the conditions under which these injuries were observed.

3.1 Introduction

The analysis of restraint system impacts using real-world crash samples can present several limitations. As discussed in the previous chapter, the effects of restraint systems may be statistically insignificant within the studied sample, making it difficult to draw definitive conclusions. This chapter addresses these challenges by developing a method to evaluate the influence of restraint system characteristics on injury outcomes using computer models, specifically multibody (MB) models.

The method was designed to estimate the impact of restraint systems on injury outcomes under real-world conditions. To ensure the method's accuracy, injury outcomes predicted by the method were estimated using the deterministic and probabilistic approaches, and these outcomes were benchmarked against the injury data presented in Chapter 2.

Since it is challenging to precisely anticipate the number of crash configurations required for such a benchmark, a metamodeling approach was employed. This approach creates a surrogate model that approximates the behaviour of the computer model, allowing for efficient and accurate predictions across a wide range of crash scenarios. Various regression methods were tested to identify the best metamodel capable of accurately representing the computer model predictions.

3.2 Methods

Figure 3.1 shows a brief summary of the methodologies applied in this chapter to benchmark both deterministic and probabilistic approaches in estimating injury outcomes under real-world conditions using an MB model.

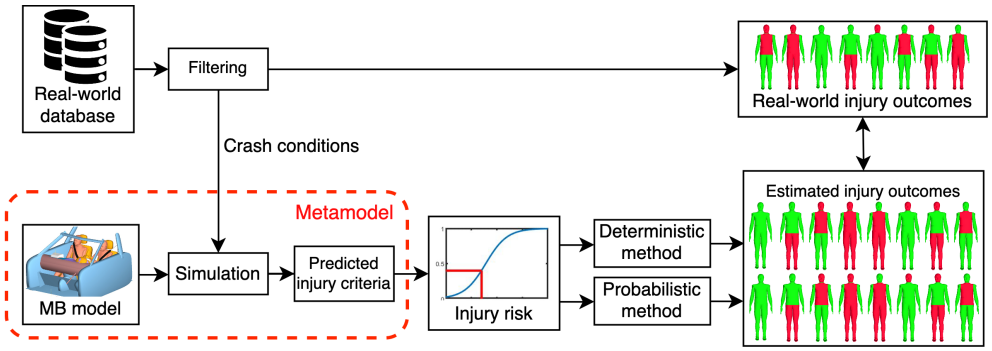


Figure 3.1: Overview of the method used to predict injury outcomes using the MB model.

3.2.1 Real-world data

Passenger vehicle crashes obtained in Chapter 2 were used in this study. These cases were extracted from the National Automotive Sampling System Crashworthiness Data System (NASS CDS) from 2000 to 2015. Furthermore, this dataset was augmented with characteristics of the restraint system used in the vehicle as explained in Subchapter 2.2.3.

The inclusion criteria used for this study were: belted drivers who were at least 16 years old at the time of the crash. Vehicles included in the analysis were restricted to those built-in or after 1990. Only frontal collisions were considered (defined as those with Principal Direction of Force (PDOF) between 11 and 1 o'clock), excluding rollover. The delta-vs observed were filtered to retain cases between 15 and 85 km/h. Additionally, only cases with intrusion below 3 cm in the instrument panel and floor were considered in this study. Although this criterion excluded 13 % of the dataset, two extra variables are excluded from the simulations, therefore decreasing the sample size necessary for the study.

AIS code from 1998 (AAAM, 1998) was used to describe the injuries sustained by the occupants in the following body regions: Head-Face-Neck (HFN), Thorax (THO), and Knee-Thigh-Hip (KTH).

3.2.2 Computational model

A multibody (MB) human body model (HBM) was used to predict the risk of an occupant sustaining an injury using Simcenter Madymo^{TM} (Helmond, The Netherlands). The HBM was positioned as a driver in the front row of a vehicle interior model, as shown in Figure 3.2. This HBM was the active Madymo human model v3.2 (Meijer et al., 2013; Simcenter Madymo, 2020), which represents a 50th percentile male occupant. The kinematic and kinetic validation of the HBM can be observed in Simcenter Madymo (2021), which includes both frontal and lateral blunt impact tests as well as sled tests. While plots of the signals used for comparing the model's predictions are presented in the document, the report does not provide a quantitative measurement of the correlation. The vehicle interior model was modelled based on a 2010 edition of the Toyota Yaris sedan model. Both models were available in Simcenter Madymo^{TM} v2020.1. A retractor and end-bracket pre-tensioner were enabled or disabled

based on the reference data. According to the reference data, either one or two-stage load-limiting seat belts were used in the computer model. Lastly, the crash acceleration pulse was calculated using the formula developed by Iraeus et al. (2015) and then applied to the vehicle model. The acceleration pulse was defined using the mean acceleration pulse, the first three eigenvectors, and the pulse duration. Delta-v, vehicle weight, and crash opponent (front-to-front, front-to-side, front-to-back, narrow object, and wide object) were used to calculate the first three eigenvectors and the pulse duration. The delta-v reported using Winsmash was scaled up with the same method used in Iraeus et al. (2016) to define the acceleration pulse. A null standard deviation was assumed when calculating the eigenvalues and pulse duration.

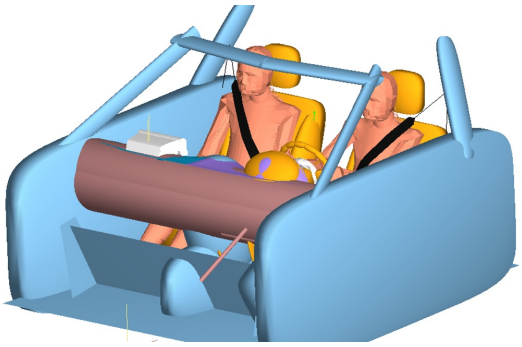


Figure 3.2: MB model used to represent real-world crashes.

Variables from the reference data were categorized into three groups: crash variables, restraint system variables, and occupant variables. Crash variables were used to calculate the acceleration pulse applied to the vehicle, restraint system variables modified the restraint system configuration, and occupant variables were not used in the simulation as the MB HBM only represents a 50th percentile male. Crash variables included delta-v (DVTOTAL), principal direction of force (PDOF, calculated using DVLONG and DVLAT), vehicle weight, and crash opponent. Restraint system variables included the deployment of the airbag, the force of the load limiter if it was included, the force of the second stage of the load limiter if it was included, the inclusion of a pre-tensioner in the end-bracket and the activation time of the airbag and pre-tensioners. Occupant variables included age, sex, height, and weight. Figures showing the distribution of these variables can be found in Appendix C.

Table 3.1 and Table 3.2 show the list of the continuous and categorical variables used as input in the simulations. Some continuous variables in the original dataset were transformed into categorical ones, such as the force of the second stage of the load limiter and the activation time of the airbag and pre-tensioners. The activation time was not available in the reference data from Chapter 2, and therefore, it was based on the distribution used in Iraeus et al. (2016). The force of the second stage of the load limiter was described in relation to the force in the first stage and categorized into 0.8, 1.3, 1.8, and 2.3 kN less than the first stage.

The occupants in the reference data had an age range from 16 to 97 years, with the median age being 36 years. The heights of the occupants ranged from 155 cm (5th percentile) to 188 cm (95th percentile), with the median height being 170 cm. The weights of the occupants ranged from 52 kg (5th percentile) to 113 kg (95th percentile), with the median weight being 76 kg.

Table 3.1: List of continuous parameters used to assemble the computer model and their distribution in the reference data.

Continuous Parameters	Reference data distribution		
	5^{th} %tile	50^{th} %tile	95^{th} %tile
Impact velocity	16 km/h	26 km/h	51 km/h
Impact angle	-32 °	0.0 °	43 °
Vehicle weight	1080 kg	1480 kg	2240 kg
Load limiter (single stage)	3.13 kN	4.52 kN	6.01 kN
Load-limiter first stage	3.38 kN	5.03 kN	6.10 kN
Load-limiter second stage	2.34 kN	3.65 kN	4.41 kN

Computer model configurations and sampling method

Different computational model configurations were generated by sampling the reference data distribution using the Sobol sequence (Bratley et al., 1988). This method was chosen for its ability to sequentially sample the distributions of the reference data, applying a quasi-random approach. Each configuration of the computational model involved using a point from the sequence within the range of [0, 1), along with the inverse cumulative distribution of the input variable from the reference data.

While most of the input parameters were sampled independently, the correlation between delta-v-to-PDOF and delta-v-to-airbag deployment was studied. Figure 3.3 shows a box plot of the distribution of PDOF using 10 km/h bins, where the range of PDOF values observed in the reference data decreased as the delta-v increased. Thus, the interaction between the delta-v and PDOF was modelled when sampling the distribution using Equation 3.1, where $\mu_{(DVTOTAL)}$ was the mean PDOF observed by a given delta-v, $\sigma_{(DVTOTAL)}$ was the standard deviation of the PDOF observed for a given delta-v, and $\Phi^{-1}(p)$ was an inverse normal cumulative function that apply a p value given by the quasi-random generator. Both the mean and standard deviation of the observed PDOF values were calculated within each 10 km/h bin. The results of applying this interaction in the generation of impact configurations can be observed in Figure 3.4. In this figure, a kernel density estimate (KDE) plot was used to show the distribution of the mentioned variables; the reference data was represented in blue, and the generated data was filled in red.

The correlation between the occupant sustaining an injury and the deployment of the airbag at high delta-vs was significant in the previous chapter. Thus, the correlation between the deployment of the airbag and the delta-v was taken into account when sampling the distributions. A similar approach as the one used with the delta-v and PDOF was also applied to model this interaction. Bins of 10 km/h were used to obtain the ratio of the airbag deployed/not deployed and later used to generate the configuration of the computational model. Figure C.6 shows the distribution of the airbags being deployed at different velocities.

$$PDOF = \mu_{(DVTOTAL)} + \sigma_{(DVTOTAL)} \cdot \Phi^{-1}(p), \qquad p \in (0, 1) \tag{3.1}$$

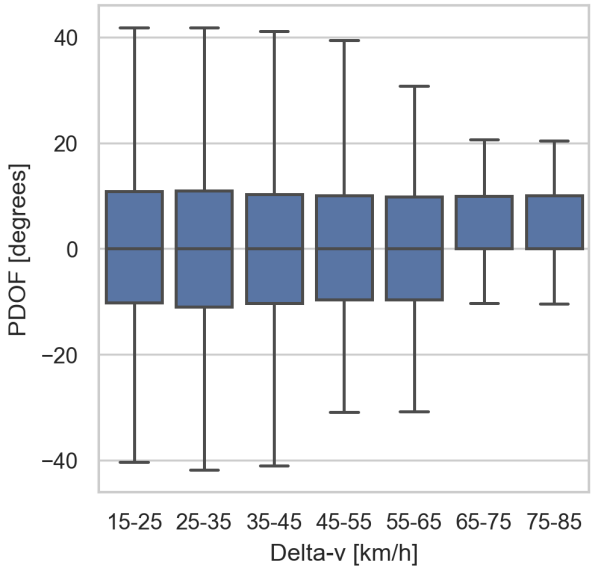


Figure 3.3: Box plot of the PDOF distribution using 10 km/h bins.

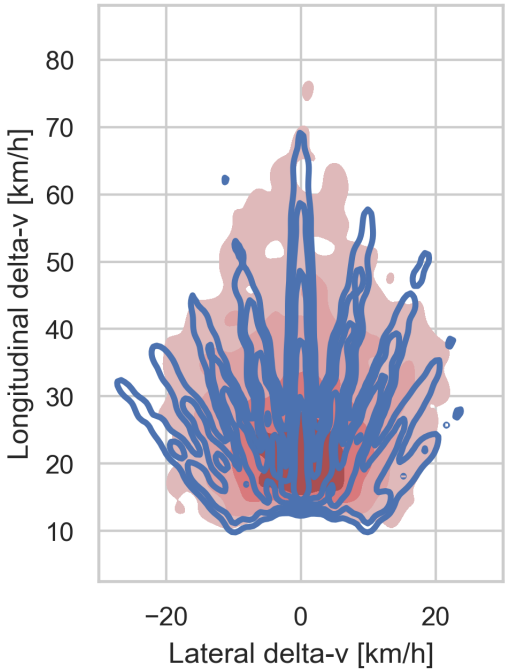


Figure 3.4: Distribution of longitudinal and lateral delta-vs for reference data, in blue, and generated data, in filled red. Distributions were obtained using a kernel density estimate (KDE) plot.

Table 3.2: List of categorical parameters used to assemble the computer model and their distribution in the reference data.

Categorical Parameters	Reference data distribution	
Crash opponent	Front-to-front	32.2 %
	Front-to-side	37.2 %
	Front-to-back	20.6 %
	Narrow object	5.5 %
	Wide object	4.5 %
Airbag deployment	Deployed	68.1 %
	Not deployed	31.9 %
LL included	Not included	51.7 %
	Single stage	43.6 %
	Double stage	4.6 %
LL second stage force	0.5 to 1.0 kN less than first stage	11.1 % [a]
	1.0 to 1.5 kN less than first stage	51.3 % [a]
	1.5 to 2.0 kN less than first stage	29.5 % [a]
	2.0 to 2.5 kN less than first stage	8.1 % [a]
Pre-tensioner included	Not included	50.9 %
	Retractor	43.9 %
	Retractor and end-bracket	5.2 %
Activation time of restraint system	10 ms	30.2 % [b]
	18 ms	30.8 % [b]
	28 ms	19.2 % [b]
	40 ms	19.8 % [b]

[a] Relative to cases with a double-stage load-limiting device.
[b] Relative to cases with a pre-tensioner or airbag fired.

Injury criteria

The injury criteria were developed to calculate the risk of sustaining injuries of a certain severity based on physical measurements, such as acceleration, force, or strain, using an injury risk function. This study focused on injuries sustained in the Head-Face-Neck (HFN), Thorax (THO), and Knee-Thigh-Hip (KTH). From each region, some injury criteria were selected to predict the injuries in such regions. The region and the injury criteria selected are shown in Table 3.3. From the reference data, relevant injuries related to the selected injury criteria were identified using the AIS code (AAAM, 1998). Body region, type of anatomic structure and specific anatomic structure or nature were used to perform the identification process.

Two injury criteria were used to predict the risk of head injuries. Skeletal injuries were linked to the HIC_{15} criteria (Eppinger et al., 1999; Hertz, 1993; Versace, 1971). Injuries in the soft tissue of the head, considering injuries in the internal organs and unconsciousness, were linked to the BrIC criterion (Takhounts et al., 2013). Neck injuries were linked to the Nij criterion (Prasad et al., 1984), taking into account injuries in the neck and cervical spine. Similarly, injuries in the thorax were linked to the maximum compression of the chest (Neathery et al., 1975), considering thoracic and thoracic spine injuries. Femur skeletal injuries were considered for injuries in the KTH body region, which was linked to the maximum compression force in the femur (Kuppa et al., 2001). Only femur injuries were taken into consideration in

Table 3.3: List of body region and injury criteria used for injury prediction.

Body region	Body structure	Injury criteria	Numerical Injury Identifier[a]
Head	Skeletal	HIC_{15}	1-5-04
	Internal organs	BrIC	1-4-any
	Unconsciousness		1-6-any
Neck	Neck	N_{ij}	3-any-any
	Cervical spine		6-any-02
Thorax	Thorax	Max. chest deflection	4-any-any
	Thoracic spine		6-any-04
Thigh (femur)	Skeletal	Max. femur compression force	8-5-18

[a] Desribed as REGION90-STRUCTURE-STRUSPEC.

this body region, as the injury risk function developed by Kuppa et al. (2001) is widely accepted for predicting distal femur injuries (Yoganandan et al., 2014). A binary variable was created to describe if the occupant sustained an AIS 3+ injury in these body regions. Furthermore, the maximum AIS (MAIS) for each body region was calculated and categorized as MAIS 3+ and MAIS 0-2. Table 3.4 shows the different injury risk functions used to predict the injury risk of sustaining an AIS 3+ injury using the MB HBM.

Table 3.4: List of injury risk functions used with the MB HBM.

Body region	Injury criteria	Injury risk function	Reference
HFN	HIC15	$P_{AIS\ 3+}(HIC_{15}) = \Phi\left(\frac{ln(HIC_{15}) - \mu}{\sigma}\right)$ where Φ = cumulative normal distribution $\mu = 7.45231$ and $\sigma = 0.73998$	Eppinger et al. (1999)
	BrIC	$P_{AIS\ 3+}(BrIC) = 1 - exp^{-\left(\frac{BrIC}{0.987}\right)^{2.84}}$	Takhounts et al. (2013)
	N_{ij}	$P_{AIS\ 3+}(N_{ij}) = \frac{1}{1 + exp^{3.2269 - 1.9688 N_{ij}}}$	Eppinger et al. (1999)
THO	Max. Chest Deflection [mm]	$P_{AIS\ 3+}(D) = \frac{1}{1 + exp^{10.5456 - 1.568 (D/1.5)^{0.4612}}}$	Hu et al. (2015) Laituri et al. (2005)
KTH	Max. Femur Compression Force [kN]	$P_{AIS\ 3+}(F_C) = \frac{1}{1 + exp^{4.9795 - 0.326 F_C}}$	Kuppa et al. (2001)

3.2.3 Development of the metamodel

The outcome of the simulations performed with the computer model mentioned above was used to develop a metamodel for each injury criterion listed in Table 3.3. This process was similar to the one described by Perez-Rapela et al. (2020). It initially involved two sets of 50 simulations: the first was used to train an initial metamodel (training sample), and the second was

used to assess the prediction capability of the metamodel (testing sample). The number of simulations in the training sample was increased in an iterative fashion until the metamodel predictions had enough quality, which was assessed by calculating the error of the metamodel predictions compared to the actual output of the simulations in the testing sample. Specifically, the two criteria used to assess the quality were:

1. First, the Mean Absolute Error (MAE) between the prediction given by metamodels and the output of the testing sample was smaller than a set limit (this criterion will be referred to as the magnitude of the prediction error).

2. Second, the variation of the magnitude of the prediction error between two subsequent iterations was smaller than a set limit (this criterion will be referred to as the stability of the prediction error).

These two criteria will be further explained in a later subsection (subsection 3.2.3). Figure 3.5 summarizes the method used to calculate the metamodels. This process used the distribution of the reference data to obtain the two sets of samples already mentioned: the training and testing samples. The metamodels were developed using the training sample output, and their prediction error was measured using the testing sample output. The final metamodel was obtained if the quality criteria were fulfilled, and the process ended there. Otherwise, the testing sample was merged with the training sample to train a new metamodel, which was then used in the next quality assessment. Concurrently, a new set of 50 testing samples was obtained to assemble and run new simulations for the new quality assessment.

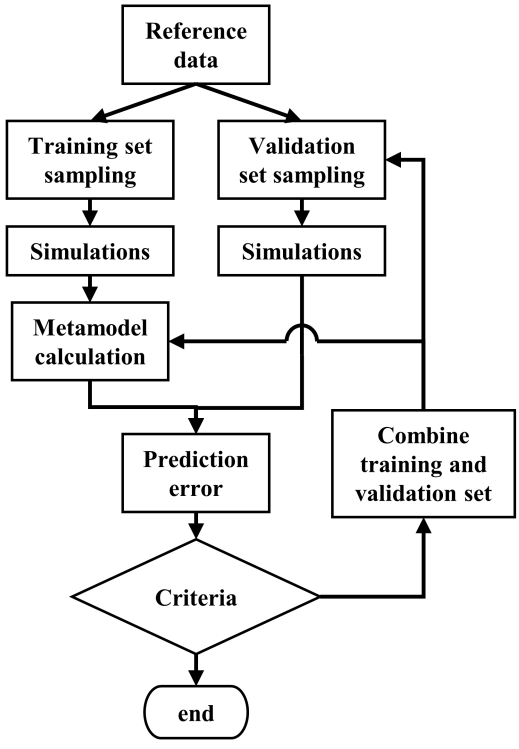


Figure 3.5: Overview of the method used to calculate the metamodels.

All failed simulations were discarded in this development and replaced with successful ones. Once all the required computer simulations were finished successfully, four different

methods were used to obtain the metamodels. These methods were selected as the ones exhibiting the lowest prediction error in the work published in Joodaki et al. (2020a) and Larsson et al. (2024): the least absolute shrinkage and selection operator (LASSO), neural networks (NN), support vector regressions (SVR) and Gaussian process regression (GPR). Table D.1 shows the set of hyper-parameters used in these metamodels to avoid over- and under-fitting the training data. This process used a K-fold cross-validation method (James et al., 2021) with K equal to five. To improve the performance of the regressions, all input variables were scaled before fitting the metamodels to a zero-to-one scale. Furthermore, output parameters were also scaled to zero-to-one for the training using NN.

Calculation of the quality criteria to assess the metamodels

The magnitude of the prediction error was calculated using the mean absolute error (MAE) for the testing sample. In this process, the MAE was calculated by taking the average absolute difference between the metamodel prediction and the simulation output under the same impact conditions using each sample in the testing set. Three error levels were defined by the authors: "unacceptable", "acceptable", and "good" magnitude of the prediction error. These levels were determined based on a percentage of the reference values in Table 3.5, referenced as threshold values in the table. An "acceptable" prediction level was reached when the MAE was under 15 % of the reference value, and a "good" prediction was reached when this value was under 10 % of the reference value; otherwise, an "unacceptable" prediction was obtained. The stability of the prediction error also used the MAE and the reference value for its definition. In this case, two subsequent metamodels were used to calculate the change in the prediction error; the MAE difference between those metamodels must be under 3 % of the reference value to obtain a stable prediction error.

In order to achieve these objectives, the metamodel needs to accurately account for the impact of different input variables. Therefore, the sample size must be sufficiently large for the metamodel to capture these variations. This size will be dependent on the complexity of the output that the metamodel aims to predict. However, as indicated in Table 3.2, more than 50 % of the cases did not involve the use of a load-limiter device, and approximately 32 % of the samples did not have the airbag deployed. This could result in one-third of the simulations having variations solely due to differences in impact conditions. This imbalance in the training dataset could lead to issues in the converge process (Alejo et al., 2007; Hensman et al., 2015; Pulgar et al., 2017). Therefore, some distributions of the categorical parameters were adjusted just for the process of training the metamodel. This was done to create more variations in the samples, with the goal of reducing the total number of samples needed to meet the quality criteria. As a result, the number of simulations without a load-limiter device was adjusted to 10 %, and within this subset, the number of cases with a dual-stage load-limiter was increased to 40 %. The same changes were applied to the pre-tensioner devices. The cases with no pre-tensioner device were modified to 10 %, and the number of cases with pre-tensioners in the retractor and end-bracket increased to 40 %.

3.2.4 Injury analysis

Simulation outputs (a physical magnitude such as acceleration, displacement, or force) were transformed into the probability of sustaining an injury of a specific AIS for each considered body region. The conversion between physical parameters and injury risk was achieved using injury risk functions. In particular, three body regions and their respective injury risk curves were considered in the analysis: Head-Face-Neck (HFN), Thorax (THO), and Knee-

Thigh-Hip (KTH), as shown in Table 3.5. The criteria used for the KTH body region were calculated for the left and right regions.

Table 3.5: List of injury criteria and threshold values used for injury prediction with the determinitic method.

Body region	Injury criteria	Threshold value
HFN	HIC_{15}	700
	BrIC	1
	N_{ij}	0.5
THO	Max. Chest Deflection	$60mm$
KTH	Max. Femur Compression Force	$7.56kN$

After obtaining the probability of sustaining an injury of a certain AIS severity using the injury risk curves, this probability value needs to be converted to a binary value, i.e., whether the injury was sustained or not. Injury patterns were calculated using this output from the metamodels and compared with those observed in the reference data.

Two methods were used to convert the probability value (% likelihood of a serious injury) into a binary value (serious/no-serious injury): the deterministic and the probabilistic methods. Figure 3.6 shows the methodology used to predict if the occupant was severely injured using the abovementioned methods.

The deterministic method calculated whether there was or was not an injury based on a set of threshold values, which are listed in Table 3.5. These threshold values were based on the lower limit for the five-star-rating limit of the Euro NCAP Assessment Protocol for Adult Occupant Protection (EuroNCAP, 2021). The deterministic method predicted a serious injury ($AIS \geq 3$) if the physical magnitude obtained in the simulation was higher than the corresponding threshold value; otherwise, a non-serious injury ($AIS < 3$) was predicted.

The probabilistic method evaluated the injury severity using the risk of a serious injury and a random evaluation based on a binomial distribution as proposed by Kim et al. (2020). The evaluation procedure took the following steps for each injury criterion:

1. The physical magnitude is obtained from the metamodel (e.g., chest deflection = 70 mm).
2. The risk of sustaining a serious injury associated with the physical magnitude is computed using the corresponding injury risk curve (e.g., the predicted risk of AIS 3+ thoracic injury for a chest deflection = 70 mm is 64 %).
3. The probability of sustaining a non-serious injury is calculated (e.g., the probability of non-sustaining an AIS3+ injury is 36 % = 100 % − 64 %).
4. A random value between 0 and 1 is generated.
5. The random value and the non-serious injury probability are then used to evaluate if a serious injury is predicted. If the random value is lower than the probability value of sustaining a non-serious injury, the method predicts a non-serious injury; otherwise, a serious injury is predicted.

This method emulates the process of conducting many simulations and selecting one of them randomly. All of these simulations would have predicted the same amount of chest compression (70 mm), and therefore 64 % of them would have resulted in a serious injury to the thoracic region. Thus, the random selection of one of these simulations would result in a 64 % chance of selecting an occupant with a serious thoracic injury.

Figure 3.6 summarizes the two methodologies that are compared in this study, showing how a chest deflection of 70 mm is transformed into a serious or non-serious injury. The functions used to compute the risk of sustaining a serious injury are shown in Table 3.4. Once the injuries were predicted using either of the classification methods, the injury distribution predicted by the metamodels was calculated, taking into account the three considered body regions, and then compared to the pattern observed in the reference data. Each injury pattern was defined as the combination of non-serious injuries or serious injuries at each body region, such as occupants with serious injuries in the thorax and the head or occupants with only serious injuries in the head.

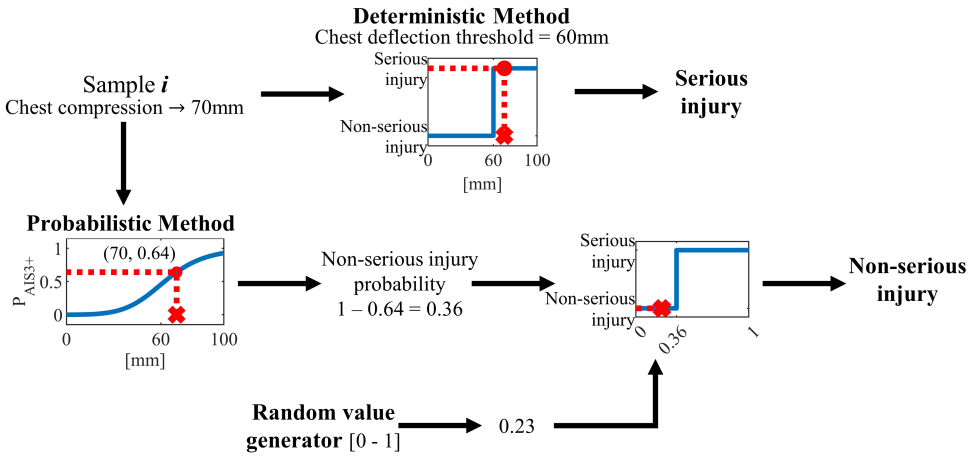


Figure 3.6: Example of applying the deterministic and probabilistic methods to the chest deflection criteria.

3.2.5 Prediction of seriously-injured occupants in a simulated population

The deterministic and probabilistic methods were applied to the outcome of the metamodels to predict the distribution of severely injured occupants. The sample, in which both methods were applied, was generated using the Sobol sequence and following the distribution of the reference data. The metamodels were used to compute the response of the human surrogate in those conditions.

A sensitivity analysis was carried out to determine a minimum sample size to ensure that the predicted distribution was obtained robustly, i.e., that the predicted distribution of injuries was the same for the given sample and a larger one. In particular, the criterion chosen to assess the convergence of the model was based on the comparison of the predicted seriously-injured occupants between two samples of different sizes: one being 1.5 larger than the other one. When the difference in the percentage of predicted seriously injured occupants between these two samples was lower than 0.5 % in each body region, it was considered that the model converged and that the sample size was the appropriate one.

Then, after reaching the acceptable sample size, the predictions given by the two methods were compared according to the following three metrics:

1. Percentage of predicted injured occupants: This metric evaluates the percentage of injured occupants for each injury and body region.

2. Predicted injury patterns: This metric evaluates the correlation between the predicted injuries in different body regions, considering if the model can predict simultaneous injuries.
3. The distribution of predicted injuries as a function of the delta-v. This metric was chosen to assess whether the methods predicted injuries under similar impact conditions as those found in the real-world data. The comparison was limited only to the thoracic region, which had the highest number of serious injuries among the body regions in the reference data.

3.3 Results

3.3.1 Real-world data

After applying the inclusion criteria to the dataset from Chapter 2, the size of the dataset was reduced to 9,957 crashes. A total of 219 occupants sustained at least one serious injury (AIS 3+) in any of the studied body regions. As shown in Table 3.7, 72, 157, and 70 occupants sustained serious injuries in the Head-Face-Neck (HFN), the Thoracic (THO), and the Knee-Thigh-Hip (KTH) body regions, respectively (note that a single individual could have received AIS 3+ injuries in two or even in the three body regions). Table 3.7 also displays these serious injuries' anatomical structure classification if considered for the injury criteria.

The top subplot in Figure 3.7 illustrates the injury patterns observed in the reference data. Each subplot comprises three columns, each representing one of the three body regions considered in this study. The total height of these columns indicates the total number of samples, with each column subdivided to display the ratio between cases with an injury severity lower than AIS 3 (i.e., AIS 0-2) and those with injuries of AIS 3+ severity. The injury severity is presented for body regions, with the maximum AIS (MAIS) depicted for each region, categorizing them into MAIS 0-2 and MAIS 3+.

The coloured horizontal areas are related to the different injury patterns found in the reference data, and the width of these areas represents the relative frequency of the injuries. For instance, the blue area indicates that an injury pattern consists of MAIS 0-2 for the HFN, MAIS 3+ for the THO, and MAIS 0-2 for the KTH. This injury pattern was found in 1.3 % of the cases in the reference data (see the value within the first row of the table included in the figure, within the column with the blue rectangle in the heading). The larger area (brown area) corresponds to the cases in which the injuries to the three body regions were of severity lower than AIS 3+, and this injury pattern was found in 97.4% of the cases. In total, eight different injury patterns were identified in the reference data, and they correspond to each of the colours listed in the legend of Figure 3.7.

This representation also allows for obtaining the number of occupants with serious injuries in each of the body regions. For instance, the number of MAIS 3+ injuries in the THO body region is the sum of the blue area, red area, cyan area, and grey area.

The most usual pattern found in the reference data was an occupant with no AIS 3+ injuries in any of the three considered body regions (97.4 % of the cases), followed by occupants with only MAIS 3+ thoracic injuries (blue area, 1.3 % of the cases), MAIS 3+ injuries only in the KTH region (yellow area, 0.6% of the cases), and then MAIS 3+ injuries only in the HFN (orange area, 0.5% of the cases). All other injury patterns accounted individually for less than 0.5 % of cases.

Figure 3.8 provides a clearer view of the injury patterns for occupants who sustained serious injuries in the reference data. While Figure 3.7 is dominated by cases with no AIS 3+

injuries, this second figure focuses exclusively on the 2.6 % of occupants who experienced AIS 3+ injuries, thereby discarding cases without severe injuries. The table at the bottom of the figure shows the frequency of each injury pattern relative to the number of occupants that sustained at least one AIS 3+ injury in the studied body regions.

3.3.2 Metamodel training process

The application of the quality criteria resulted in using 1,200 samples to train the metamodels. Table 3.6 shows the magnitude of the prediction error using the percentage of the reference value for each injury criterion in the last iteration of the training process. As seen from the table, all the injury criteria reached an MAE below 10 % of the corresponding reference value. The metamodels obtained using GPR resulted in the lowest prediction error for most injury criteria (with a mean value of 1.0 %), followed by those obtained with NN (a mean value of 1.5 %).

Figure D.1 shows the evolution of the prediction error with the increase of the sample size in the training process of each metamodel. The prediction error in the training process followed two trends: a prediction error oscillating around a constant value (only observed for the HIC criterion) and a decreasing prediction error. Both of these trends were also reported in Perez-Rapela et al. (2020).

The BrIC, N_{ij}, and chest deflection criteria showed that the GPR metamodel converged more rapidly than other metamodels. However, there was no clear difference when focusing on HIC. Regarding the evaluation of the femur compression force criteria, the prediction errors relative to the limit were the lowest. However, this was because the forces predicted by the model were consistently below 0.8 kN, showing few interactions between the lower extremities and the knee bolster. Figure D.2 shows the actual and the predicted values of the simulation and the GPR metamodel, respectively. The N_{ij} criterion showed the lowest R^2 value among the injury criteria, as it showed the largest dispersion from the 1:1 ratio between actual and predicted values.

Table 3.6: Magnitude of the prediction error for each injury criteria and regression method.

Injury criteria	MAE (% of reference value)			
	LASSO	SVR	NN	GPR
HIC_{15}	1.7	2.5	1.9	2.0
BrIC	2.0	3.3	2.3	1.2
N_{ij}	1.8	2.2	1.8	1.2
Max. Chest Deflection	3.4	5.0	2.6	1.3
Max. Femur Compression Force (Left)	0.2	0.2	0.2	0.1
Max. Femur Compression Force (Right)	0.1	0.1	0.1	0.2

3.3.3 Prediction of seriously injured occupants in a simulated population

Once the metamodels were trained, they were used to predict the serious injuries in a simulated population that included a significant variation in the impact conditions representing the reference data. Figure D.3 shows the percentage of seriously-injured occupants for each sample size. Although the converse criterion was fulfilled within the first 5,000 samples, a

larger dataset was generated using the metamodels. A total of 10,000 samples were finally generated, as the change in the percentage of seriously-injured occupants settled below 0.1 %, as shown in Figure D.4. This figure illustrates how the ratio between seriously-injured/non-injured occupants changed according to the proposed converge criteria in Subsection 3.2.5.

Table 3.7: Seriously injured occupants observed in the reference data grouped based on the body region.

Body region	n	% relative to total	% relative to body region
HFN injuries	72	0.8	
Head Injuries	53		
Skeletal injuries	11	0.1	15.3
Internal organ injury or unconsciousness	49	0.5	68.1
Neck injuries	25	0.3	34.7
THO injuries	157	1.6	
KTH injuries	70	0.7	

Table 3.8: Body region and source of the injuries predicted using the deterministic and probabilistic methods.

	Deterministic method		Probabilistic method	
Body region and Injury criteria	% relative to total	% relative to body region	% relative to total	% relative to body region
HFN injuries	2.1	-	19.2	-
HIC_{15}	0.6	28.4	0.0	0.0
BrIC	1.8	83.4	14.7	81.2
N_{ij}	0.0	0.0	5.4	28.0
THO injuries	26.8	-	7.1	-
KTH injuries[a]	0.0	-	1.5	-

[a] Takes into account the left and right femurs simultaneously.

Table 3.8 shows the proportion of MAIS 3+ injuries predicted for each body region by the different injury prediction methods. For example, the deterministic method predicted that 2.1 % of the population would sustain an AIS 3+ to the HFN body region. The table also displays which injury criteria predicted an AIS 3+ injury within each body region. For instance, within the HFN region using the deterministic method, 28.4 % of head injuries were attributed to the HIC criterion, while 84.4 % were attributed to the BrIC criterion. Note that the sum of the percentages in a given body region may exceed 100 %, as occupants could have sustained multiple injuries within the same region.

Three comparisons were conducted between the reference data and the predicted injuries using both deterministic and probabilistic methods. First, the number of injuries was compared in the datasets relative to the total number of cases, focusing on body regions and injury criteria. Second, the ratio of injuries observed in different anatomical structures was exam-

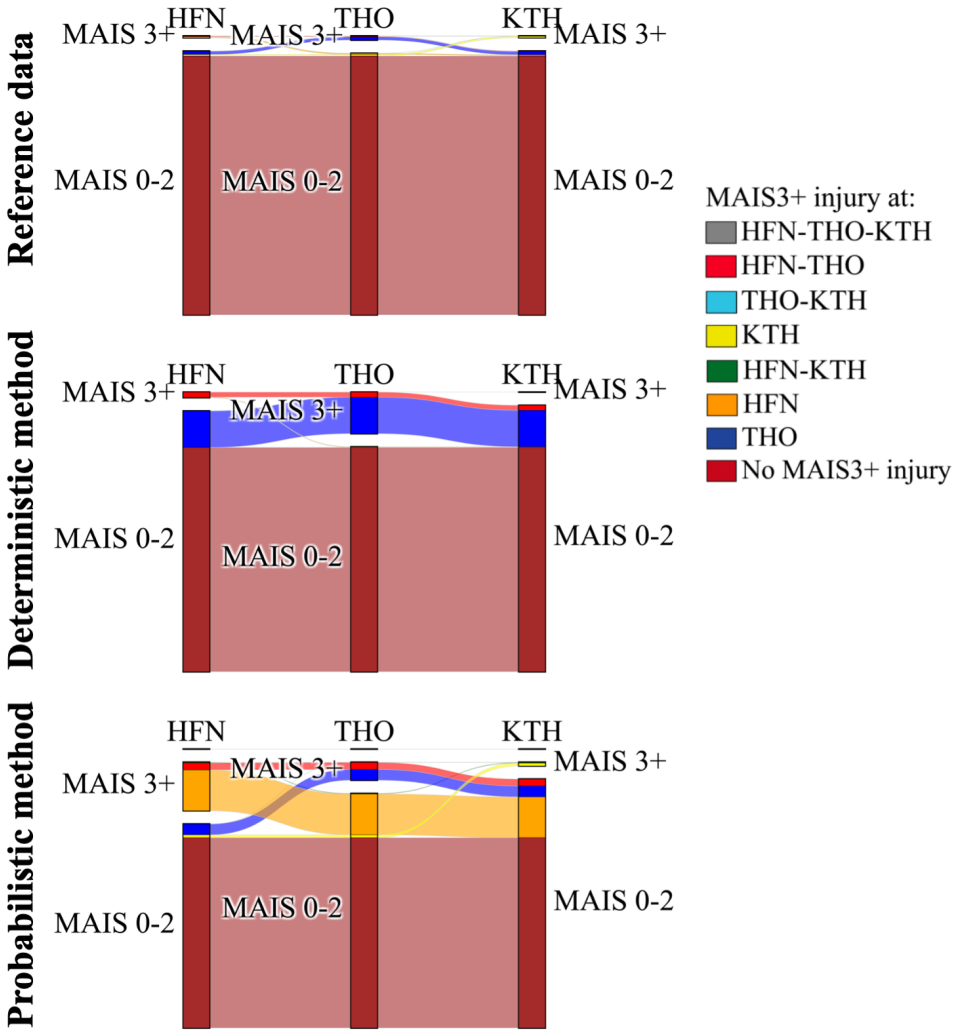
ined for the HFN region relative to the total number of injuries in that region. Lastly, the different patterns were analyzed for the predicted injuries using both methods.

Focusing on the first two comparisons, the following categories can be observed below: *i*) Head skeletal injuries against the injuries predicted with HIC_{15}; *ii*) Head internal organ injury and unconsciousness against the injuries predicted with BrIC; *iii*) Neck injuries against the injuries predicted with N_{ij}; *iv*) Thoracic injuries against the injuries predicted with the maximum chest compression; *v*) Femur skeletal injuries against the injuries predicted with the maximum femur compression force in the right and left side. From these comparisons, it can be stated that:

1. The number of MAIS 3+ injured occupants in the HFN body region was over-predicted by 2.6 times using the deterministic method.
    - The number of occupants with head fractures was 5 times over-predicted.
    - The number of occupants with head internal organ injury and unconsciousness was over-predicted by 3.6 times.
    - The number of occupants with neck injuries was underpredicted, as no simulation exceeded the threshold for the N_{ij} criterion.
2. The number of MAIS 3+ injured occupants in the HFN region was over-predicted 24 times with the probabilistic method.
    - The number of occupants with head fractures was underpredicted, as no AIS 3+ injuries were predicted for this criterion.
    - The number of occupants with head internal organ injury and unconsciousness was over-predicted by 29.4 times.
    - The number of occupants with neck injuries was 18 times over-predicted.
3. The deterministic method over-predicted the number of occupants with MAIS 3+ THO injuries by 16.7 times. The probabilistic method over-predicted the number of occupants with MAIS 3+ THO injuries by 4.4 times.
4. The number of occupants with femur fractures was underpredicted using the deterministic method, as no AIS 3+ injuries were predicted for this criterion. The probabilistic method over-predicted the number of occupants with femur fractures by 2.1 times.

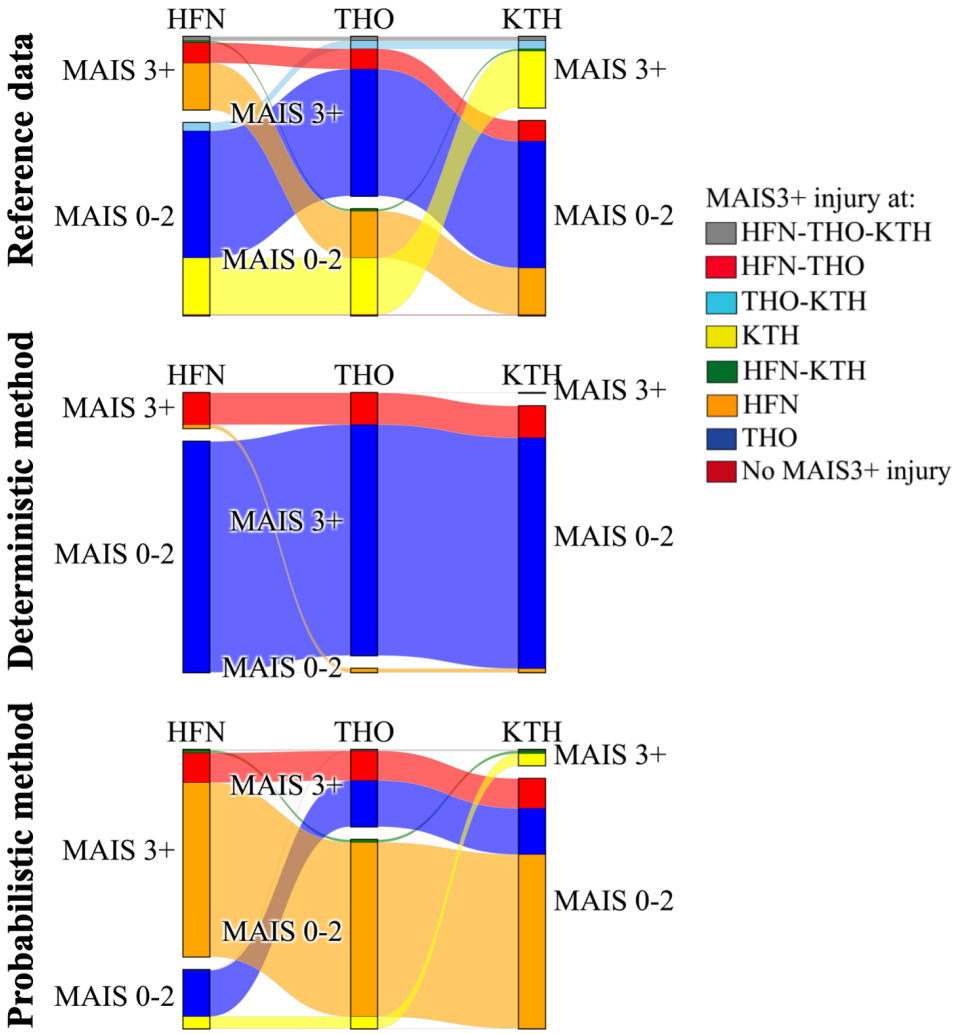
Figure 3.7 shows the injury patterns observed in the reference data and those obtained after applying the deterministic and probabilistic methods to the results from the metamodels. As shown in Figure 3.7, the deterministic method (mid-plot) over-predicted the number of seriously injured occupants in the HFN and THO body regions in relation to the reference data (top-plot). In addition, this method failed to capture all the injury patterns observed in the reference data, more clearly seen in Figure 3.8 with only occupants that sustained at least one AIS 3+ injury in the studied body regions. In particular, the following injury patterns that were present in the reference data were not captured by the deterministic methods: only HFN AIS3+ injuries, HFN and THO AIS3+ injuries, and only THO AIS3+ injuries. These figures also show that the probabilistic method (bottom plot in both figures) over-predicted the serious injuries for all regions. However, this method was capable of capturing more injury patterns observed in the reference data. In addition to the visualization of the injury patterns, Figures 3.7 and 3.8 also include a table showing the percentages of each injury pattern in the reference data and the predictions given by each method. Furthermore, the table in Figure 3.8 shows the frequency of the injury patterns relative to the number of injured occupants with at least one AIS 3+ injury.

For the analysis of the conditions predicting thoracic injuries, both injurious and non-injurious, as a function of delta-v, Figure 3.9 presents a box plot including results from the ref-



Injury Pattern (%)								
Reference data	0.0	0.2	0.1	0.6	0.0	0.5	1.3	97.4
Deterministic method	0.0	1.9	0.0	0.0	0	0.2	13.7	84.2
Probabilistic method	0.0	2.8	0.0	1.1	0.2	16.2	4.3	75.5

Figure 3.7: Injury patterns observed in the reference data and applying the deterministic and probabilistic methods.



Injury Pattern (%) relative to injured occupants								
Reference data	1.5	7.6	3.1	21.4	0.8	17.6	47.7	-
Deterministic method	0.0	11.9	0.0	0.0	0.0	1.4	86.6	-
Probabilistic method	0.2	11.2	0.2	4.6	1.0	65.5	17.4	-

Figure 3.8: Injury patterns observed in the reference data and applying the deterministic and probabilistic methods, without considering cases where no AIS 3+ injury was reported for the studied body regions.

erence data and predictions from the deterministic and probabilistic methods. The whiskers represent the 5th and 95th percentiles for each category. This comparison revealed that in the reference data, the first serious injury occurred at 23 km/h and the last non-serious injury at 53 km/h, as defined by the 5th and 95th percentiles, respectively. The probabilistic method produced similar results, with the first serious injuries at 23 km/h and the last non-serious injuries at 49 km/h. However, the deterministic method showed a different distribution, with these percentiles at 36 km/h and 39 km/h, respectively. The large number of outliers observed for both prediction methods resulted from using the 5th and 95th percentiles as whiskers, leading to 1,000 samples being classified as outliers.

Figure 3.9 presents a box plot comparing results from the reference data with predictions from the deterministic and probabilistic methods for the analysis of conditions predicting thoracic injuries. Injurious and non-injurious outcomes were analyzed as a function of delta-v. The whiskers represent the 5th and 95th percentiles for each category. This comparison revealed that, according to the reference data, the first serious injury occurred at 23 km/h, and the last non-serious injury occurred at 53 km/h, based on the 5th and 95th percentiles, respectively. The probabilistic method produced similar results, with the first serious injuries also occurring at 23 km/h and the last non-serious injuries at 49 km/h. In contrast, the deterministic method displayed a different distribution, with the 5th and 95th percentiles at 36 km/h and 39 km/h, respectively. The large number of outliers observed for both prediction methods is attributed to the use of the 5th and 95th percentiles as whiskers, which led to 1,000 samples being classified as outliers.

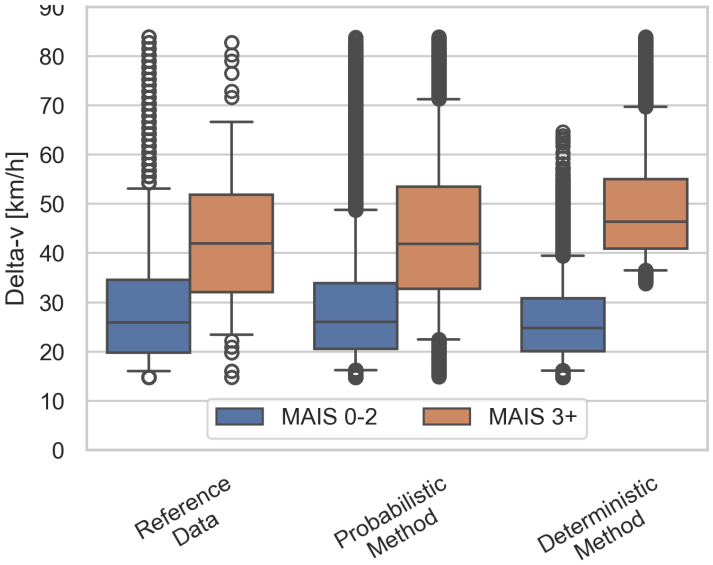


Figure 3.9: Box plot of MAIS 0-2 and MAIS 3+ thoracic injuries observed in the reference data and applying the deterministic and probabilistic methods. Whiskers in the boxplot represent the 5th and 95th percentile of the observed data.

3.4 Discussion

The comparison between real-world injury data and data from computer model predictions is not straightforward due to the difficulties associated with the estimation of injury outcomes

using computer models. Computer models can predict physical magnitudes, from which the likelihood of sustaining an injury of a certain AIS level can be calculated using an injury risk function. However, the number and the severity of injuries in real-world crashes are obtained directly from the investigation of the collisions. Thus, two approaches can be applied to ensure the comparability of the two data sources: either injury observed in the real world can be used to obtain the injury risk that correlates the likelihood of certain injury severity with the impact conditions of the crash, as in Iraeus et al. (2016) and Takhounts et al. (2019), or a method can be employed to assign whether the computer model predicts an actual injury or not in a given impact condition as in Perez-Rapela et al. (2020), Kim et al. (2020), Bance et al. (2021), and Wang et al. (2021). To our knowledge, this is the first work that has studied the performance of both approaches in evaluating injury outcomes using a comparable methodology.

Neither of the two methods used in this study to evaluate serious injuries has resulted in the same injury distribution observed in the real-world reference crashes, regardless of considering only specific injuries or injuries grouped by body region. It was observed that each individual method could have a better performance for some body regions but worse for others, which opens the door to assessing whether a combination of the methods could obtain a better overall performance.

3.4.1 Computer model and metamodels

The convergence of the metamodels showed similar trends as those observed in Perez-Rapela et al. (2020). The metamodels with the lowest prediction error were obtained using the GPR method, which was not employed in the referenced study. Although the regression with NN showed a higher prediction error, it can be trained to predict several outputs with the same network, which was not applied in this study.

The criteria to evaluate the converge performance was based on the thresholds used for the deterministic method, as 10 % and 15 % of the values. Although the metamodels showed prediction errors below these thresholds for almost all iterations, a variation in the impact conditions and/or threshold values can modify the converging process considering the prediction error and its stability as more samples are used in the training process. Thus, further research should explore how these variations influence the metamodels' ability to consistently and accurately predict the injury criteria when different crash conditions or HBM models are used.

3.4.2 Prediction of injury outcomes and assessment of the deterministic and probabilistic approaches

The metamodels allowed the prediction of injury outcomes for a population of 10,000 individuals, which is approximately ten times greater than the number of simulations carried out to obtain the metamodels. This reduction in the computational cost using metamodels saved weeks of computing, and it will take even higher importance if the method is applied to more complex HBMs, such as finite element (FE) models.

Deterministic method

Differences were found between the predictions based on the deterministic method and the reference data: serious injuries were over-predicted for some body regions, and no injuries were predicted for others, some of these over-predicted serious injuries were associated with

an injury criterion that was not related to the actual injury, and the predictions did not capture some injury patterns observed in the reference data.

The observed over-prediction of thoracic injuries aligns with previous studies, such as Hu et al. (2015). In that study, which used the same HBM, an over-prediction of chest deflection was noted when compared to predictions using THUMS v4.0. To address this issue, the authors adjusted the injury risk function based on the chest deflection of the MB HBM to predict an injury risk similar to the one observed in a NASS CDS dataset. This modified injury risk function was applied in the current study (see Table 3.4), resulting in a smaller discrepancy between the probabilistic method and the reference data in relation to the observation using the deterministic method. Although other studies have proposed scaling methods for HBM predictions (Forman et al., 2022), it is crucial that adjustments to the injury criteria be validated with multiple sources of reference data. This ensures that the adjustments do not obscure the model's lack of bio-fidelity by over-fitting a particular dataset.

No serious injury was predicted for the KTH region with the deterministic method, as no simulation overcame the threshold used for the maximum compression force of the femurs. Although this may be linked to the validation of the vehicle interior model and the contact forces between the occupant and the knee bolster, the probabilistic method over-predicted the mentioned injuries. Lastly, the deterministic method over-predicted the percentage of HFN injuries by 2.6 times. However, this correlation does not ensure similar outcomes for each injury criterion separately. While the BrIC criterion over-predicted soft tissue head injuries by 3.6 times, the HIC_{15} criterion over-predicted the number of occupants with skull fractures by 5 times, and there were no neck injuries. These results suggest that the benchmarking of the method must consider each injury criterion separately to guarantee similar predictions by the model as those observed in the reference data.

It is important to recognize that the accuracy of injury prediction using the deterministic method relies on the threshold values selected. In this study, they were chosen from the lower boundary of the five-star-rating Euro NCAP protocol (EuroNCAP, 2021). However, alternative thresholds could have led to different results. On the one hand, the method demonstrated an over-prediction of the number of injuries observed in the reference data and failed to capture certain injury patterns observed in real-world scenarios. The selection of a higher threshold could have been used to avoid this over-prediction. On the other hand, this approach failed to capture the large range of velocities in which serious thoracic injuries were observed in the reference data. Thus, increasing the threshold to solve the over-prediction of injuries would have negatively impacted this metric, as it would have resulted even in higher delta-vs at which serious injuries were predicted and even more different distributions. Figure 3.10 shows the mentioned impact of increasing the threshold to decrease the total number of AIS 3+ injuries in the THO region by applying a 10 % and 20 % increase in the threshold. This modification resulted in increasing the velocity at which the AIS 3+ injuries were predicted. Furthermore, this modified threshold was not able to predict a similar overlapping range of delta-v in which cases with and without AIS 3+ injuries were observed in the reference data.

Although Bance et al. (2021) and Wang et al. (2021) also employed the deterministic method, as they used other metrics to compare their results with the reference data, a direct comparison with the results of the current study is not possible.

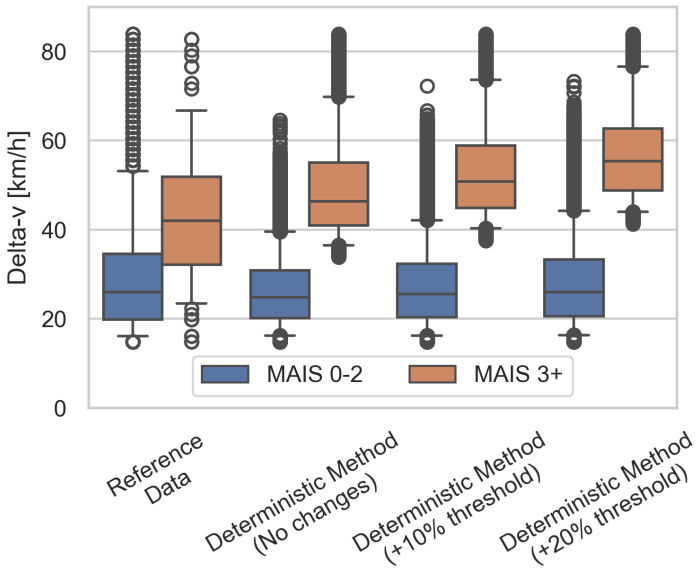


Figure 3.10: Box plot of MAIS 0-2 and MAIS 3+ thoracic injuries observed in the reference data and applying the deterministic methods with two extra variations. The first and second variations used a 10 % and 20 % higher threshold than the reference case. Whiskers in the boxplot represent the 5th and 95th percentile of the observed data.

Probabilistic method

Contrary to the deterministic method, the probabilistic method was able to predict all injury patterns observed in the reference data. However, the number of cases within each injury pattern (the width of the line) was different from the reference data, with large over-predictions of HFN injuries.

The main difference between the injuries predicted by each injury criterion using the probabilistic method and the reference data was observed with the BrIC criterion. Takhounts et al. (2019) found that this criterion is highly dependent on the delta-v and PDOF of the crash. In their study, a weighted average of the injury risk predicted using the BrIC criterion was used to compare predictions with NASS CDS data, accounting for the frequency of each crash configuration. This approach showed comparable results for injury risk prediction. Although this method was not applied in the current study, the frequency of crash configurations was considered when obtaining the input dataset for the metamodels. However, this led to an over-prediction of the number of injured occupants.

The probabilistic method also over-predicted the number of THO AIS 3+ injuries. However, it accurately approximated the 5th and 95th percentiles of delta-v for both seriously and non-seriously thoracic-injured occupants, with a difference of 5 kph or less. As previously mentioned, the injury risk function used for this criterion was scaled according to Hu et al. (2015). Therefore, a greater over-prediction of THO AIS 3+ injuries would be expected if the original injury risk function had been applied. Additional scaling could be also implemented to match the reference data in this region. However, this was not applied as the weights of the NASS CDS samples were not used in the reference data.

3.5 Limitations

The first limitation of the study was associated with the reference data. As the reference data was based on the results from Chapter 2, the weights of the NASS CDS samples were not used. Consequently, the reference dataset was biased towards more injurious outcomes. However, including these weights could have introduced bias in the additional variables augmented in the dataset (i.e., restraint system variables). Furthermore, while the width of the injury patterns would have changed with the inclusion of weights, the same patterns would still be present in both datasets.

The second limitation of the study was associated with the computer model used to calculate the metamodels. Although MB models are computationally less expensive, implementing an FE model could be necessary to capture more complex behaviours related to injury causation in the real world, as explored in the next chapter. FE models allow the use of strain-based injury criteria, which could be applied to all crash directions. As the applied injury criteria used in this study are based on those developed for ATDs, they cannot be considered omnidirectional, and the discussion of how appropriate they are for non-pure frontal impacts is relevant but out of the scope of this work. Furthermore, the model used in the study only represents the 50th percentile male, which does not represent the whole population observed in the reference data. While a previous Madymo human body model was developed to be scalable to different anthropometries (Rodarius et al., 2007), the latest version of the Madymo human model was used here to carry out the simulations. Regarding the vehicle interior and restraint system, this study focused on modifying the seatbelt system. This resulted in using the same airbag and vehicle interior properties for all simulations.

The third limitation of the study was associated with the prediction of injuries. On the one hand, not all the injury criteria have publically available validation data, such as comparing the injury prediction capabilities of the model with injuries reported by PMHS tests and field data. Although the quality report of the model mainly focuses on kinematic and kinetic validation of the model (Simcenter Madymo, 2021), previous studies have used the applied injury criteria (Bance et al., 2021; Hu et al., 2015; Kim et al., 2020). On the other hand, Hu et al. (2015) and Forman et al. (2022) refined injury risk functions for thoracic injuries in MB and FE models, respectively, to better align computer model predictions with real-world data. A similar strategy could be applied to this method to enhance the predictive accuracy of metamodels. However, there are inherent approximations in the computer model, such as the use of a single male anthropometry and variations in the vehicle. Modifying injury thresholds or injury risk functions to improve data alignment may obscure these model limitations and result in an ad-hoc solution specific to this study. Furthermore, other injury criteria could be applied to the HBM to predict occupant injuries. A possible replacement could be implemented with the DAMAGE criteria (Gabler et al., 2019), which Östh et al. (2023) found over-performing the BrIC in a set of vehicle crash tests using the THOR 50th percentile male and the WorldSID 50th percentile male ATDs. However, the solver version used in this study did not support the use of such criteria.

3.6 Related publications

- Valdano, M., Jiménez-Octavio, J.R., Pipkorn, B., Otero-Peinador, A., Sánchez-Merchante, L.F., López-Valdés, F.J., Evaluation of AIS3+ car occupant injuries using deterministic and probabilistic methods in frontal crashes. Computer Methods in Biome-

chanics and Biomedical Engineering. Vol. 27, nº. 12, pp. 1714 - 1730, September 2024.

4 Evaluating the effectiveness of deterministic and probabilistic approaches for predicting injury outcomes using finite element models in real-world conditions

This chapter examines the use of finite element (FE) models to predict injury outcomes in real-world conditions. This modelling approach enables the detailed representation of occupant anatomy, applying morphing techniques to reflect a diverse population, and using more advanced injury criteria. To validate the model's predictive capabilities, its predictions were compared with actual injuries and injury patterns documented in real-world crashes. An FE human body model (FE HBM), combined with a generic vehicle interior model, was used to estimate the likelihood of injuries based on selected criteria across various impact scenarios derived from real-world crash data. Additionally, a metamodel was developed to predict the likelihood of moderate injuries, serving as a surrogate model for the full computer model. As presented in Chapter 3, the injury outcomes from real-world data (obtained in Chapter 2) were compared to estimates from deterministic and probabilistic approaches. This benchmarking assessed the predicted number of injured occupants, the injury patterns observed, and the conditions under which these injuries were observed.

4.1 Introduction

In Chapter 3, a method was introduced to evaluate the influence of restraint system characteristics on injury outcomes, utilising a multibody (MB) model to complement the study present in Chapter 2. Although MB models offer a simplified representation of occupants and allow for computational efficiency, the model in this study was constrained by its reliance on a single occupant anthropometry to represent the entire population. To overcome this limitation, more advanced models could be employed, such as finite element (FE) models which are capable of being morphed to represent different anthropometries. In this study, a morphing technique was applied to the FE model to simulate a wider range of population anthropometries, moving beyond the typical reliance on a 50^{th} percentile male occupant, thereby providing a more inclusive prediction of injury outcomes. Furthermore, the FE model facilitated the application of advanced injury criteria to evaluate potential injuries in specific anatomical structures, a factor that was integrated into the analysis.

As in the previous chapter, the estimations of the method, using both deterministic and probabilistic approaches, were benchmarked against the injury data presented in Chapter 2. In this instance, the evaluation incorporated variations in impact conditions, restraint system configurations, and occupant characteristics.

Nearly 90 % of the computational cost required for the crash configuration estimations used in the benchmark in Chapter 3 was reduced by employing a metamodel. These computational savings are particularly significant when working with FE models, as the reduced cost can prevent the need for simulations that would otherwise take weeks or even months, depending on the available computational resources. Additionally, as more variables are introduced in this study, the number of crash configurations is expected to increase, further driving up computational demand. To address this, various regression methods were tested to determine the most suitable metamodel capable of accurately capturing the computer model's behaviour.

4.2 Methods

Figure 3.1 shows a summary of the methodologies applied in this chapter to benchmark both deterministic and probabilistic approaches in estimating injury outcomes under real-world conditions using an FE model.

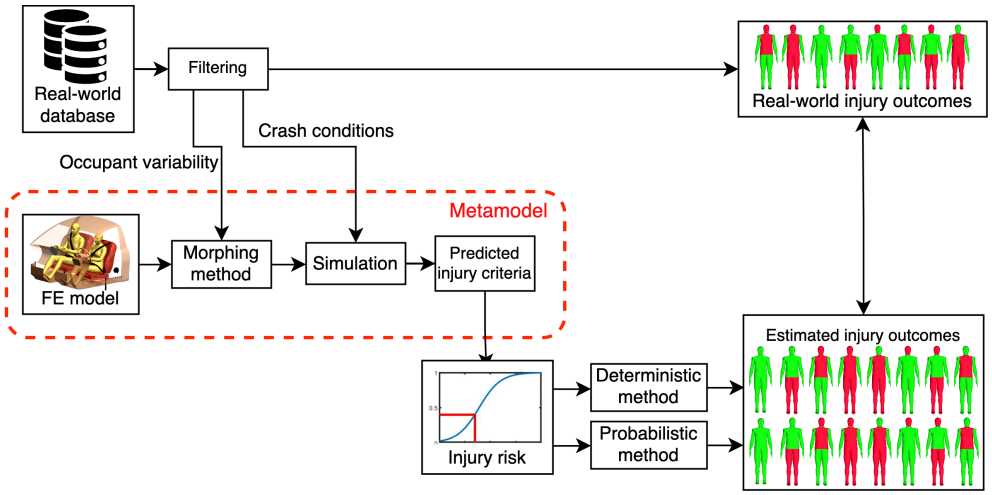


Figure 4.1: Overview of the method used to predict injury outcomes using the FE model.

4.2.1 Real-world data

Passenger vehicle crashes obtained in Chapter 2 and filtered in Chapter 3 were used in this study. Occupant variables (such as height, weight, age and sex) were included in this chapter for the analysis, as the human body model (HBM) was morphed to consider different anthropometries.

AIS code from 1998 (AAAM, 1998) was used to describe the injuries sustained by the occupants in the following body regions: Head-Face-Neck (HFN), Thorax (THO) and Knee-Thigh-Hip (KTH).

4.2.2 Computational model

A finite element (FE) model was used to predict the likelihood of occupant injuries using LS-Dyna (16 cores, v12.1 mpp single precision, Livermore Software Technology, Livermore, CA). The FE model was assembled using a generic vehicle interior (GVI) model and a HBM.

A frontal impact generic vehicle interior sled model (referred to as the GVI model hereafter) (Iraeus et al., 2016) was modified, and its response was validated to simulate frontal crashes, with the HBM positioned as the driver. The modifications included replacing the original seat model with an open-source model (*FE_models / Open Access Front Seat Model · GitLab* 2024) and updating the belt system, except for the buckle and the buckle cable restraint. The seatbelt model was routed using Oasys Primer (v.17.0, Oasys Ltd., London, UK) without additional slack. Bending beams (Dahlgren et al., 2020) were implemented in the lap and shoulder belt webbing to prevent excessive belt bending. Appendix E provides a detailed description of the GVI model, the specific modifications performed, and their validation.

As detailed in Chapter 3, a retractor and end-bracket pre-tensioner were enabled or disabled based on the reference data. Additionally, a one- or two-stage load-limiting device was activated or deactivated accordingly. The crash acceleration pulse was calculated using the formula developed by Iraeus et al. (2015).

A baseline HBM was used to generate different anthropometries to represent all occupants observed in the reference data. The SAFER HBM (SHBM) v11.0.2 (Pipkorn et al., 2023) was used as the baseline model. Although validation with the used version of the model has not been released, validations with previous versions of the model included Pipkorn et al. (2019), Iraeus et al. (2019), Larsson et al. (2021b), Larsson et al. (2024), Kleiven (2007), and Fahlstedt et al. (2022). Once the HBM was morphed, it was seated according to the sex, height, and weight of the occupant based on the methodology described by Park et al. (2016), utilising the marionette method. The seat position was also calculated following Park et al. (2016), with an additional 25 mm upward displacement to account for some anthropometries' lack of seat cushion deformation. The strain and stresses of the pelvis flesh from the HBM were imported from the positioning simulation to the crash simulation.

To ensure the stability of the models, each morphed HBM was used to run four simulations. These simulations included pure frontal crashes at speeds of 30, 50, and 70 km/h and a 45 km/h far-side crash at 30°. A total of 200 simulation results were meticulously inspected to identify and rule out any contact intersections, poor-quality elements or numerical instabilities.

Table 4.1 shows all the parameters used to obtain the computer model.

Table 4.1: Parameters used to obtain the computer model.

Acceleration pulse	GVI model	Occupant model
• Delta-v	• First stage LL	• Occupant height
• PDOF	• Second stage LL	• Occupant weight
• Vehicle weight	• Airbag deployment	• Occupant sex
• Crash opponent	• Retractor PP	
◦ F2F	• Anchor PP	
◦ F2S	• Activation time delay	
◦ NO	• Knee bolster stiffness	
◦ WO	• Steering column force	
	• Airbag radius	
	• Airbag pressure	

Generation of Morphed Human Body Models

The distribution of occupant variables is detailed in Appendix C.3. The reference dataset showed nearly equal numbers of male and female occupants, with male occupants comprising 50.4 % of the total. Figure C.14 shows the population distribution of male and female occupants by age, revealing no significant differences below 80 years old (i.e., a difference below 5 % in each 10-year age group). Figure C.15 illustrates the age distribution of the reference data, regardless of sex, indicating that 79 % of the occupants were younger than 55 years old and 89 % were below 65 years old.

Regarding the weight and height of the reference data, Figure C.16 shows a univariate and bivariate distribution of these variables, grouping the data by occupant's sex. Based on Larsson et al. (2024), 50 combinations of weight and height (25 combinations for each sex) were chosen to represent the reference population. A kernel density estimate (KDE) was calculated using the occupant's weight and height observed in the reference data. Figure 4.2 shows the probability distribution of female and male occupants using the KDE. The probability levels in the KDE plot, shown with the different coloured areas, were adjusted to ensure that 90 % of the occupants of each sex were contained within the outermost contoured area in each plot. Black points show the combinations of weight and height used to morph the HBMs. Furthermore, Table F.1 and F.2 show the combination of parameters to morph female and male occupants, respectively.

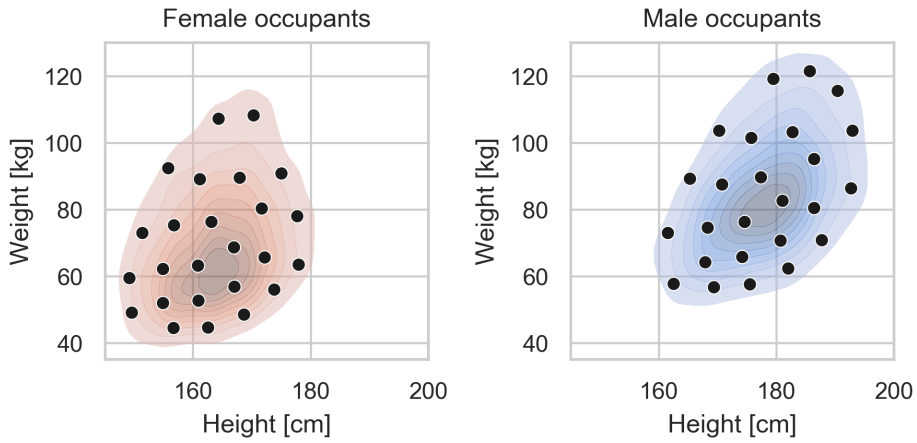
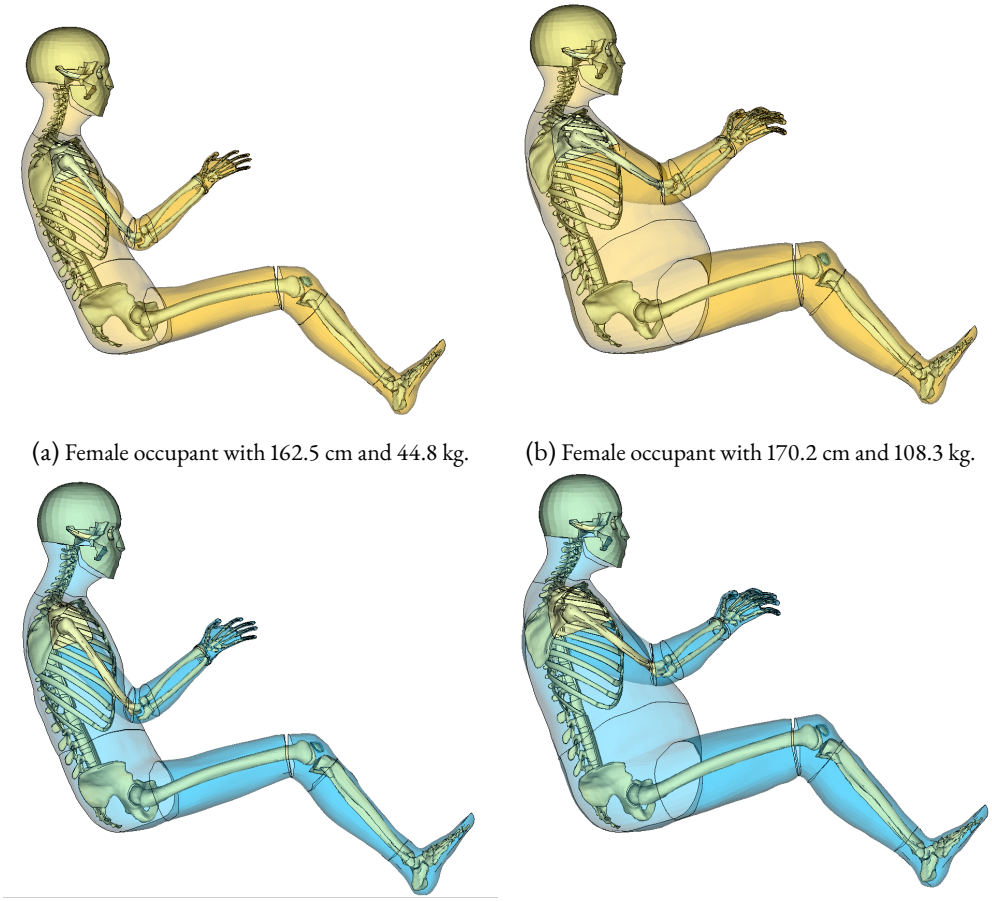


Figure 4.2: Distribution of occupants' height and weight in the reference dataset, illustrated using kernel density estimation (KDE). The shadowed areas in each plot represent 95 % of the population for female and male occupants in the reference data. The black points indicate the 50 combinations of weight and height used to morph the HBM, with female occupants shown in the left-hand subfigure and male occupants in the right-hand subfigure.

The HBM morphing method (Hwang et al., 2016a) was employed with the SHBM to obtain HBMs with anthropometric targets from Table F.1 and Table F.2 in Appendix F. This method combines a statistical skeleton geometry model with a statistical human shape model to determine the geometry of the ribcage, femur, pelvis, tibia, and body surface. A set of landmarks in both the statistical models and the SHBM was then used to modify the mesh of the SHBM accordingly. The models' element quality and contact interference were observed and modified if necessary to avoid numerical errors in the simulations. This method focused on modifying the geometry of the occupant, and material properties were not altered in the process.



(a) Female occupant with 162.5 cm and 44.8 kg.

(b) Female occupant with 170.2 cm and 108.3 kg.

(c) Male occupant with 175.5 cm and 57.6 kg.

(d) Male occupant with 185.6 cm and 121.3 kg.

Figure 4.3: Examples of morphed HBMs before performing positioning simulation.

Age and seating height ratio must also be defined for the HBM morphing method. All occupant anthropometries used an age and seating height ratio of 45 years and 0.52, respectively. The baseline SHBM material properties were not modified for this study. This age was also applied to the method used to calculate the occupant position using Park et al. (2016).

Figure 4.3 illustrates four examples of morphed HBMs, showing one short and underweight and one tall and overweight subject for each sex. The height and weight of each subject can be observed in the caption of the figure.

Injury criteria

As in Chapter 3, this study focused on injuries in the following regions: Head-Face-Neck (HFN), Thorax (THO), and Knee-Thigh-Hip (KTH). However, the injury criteria and severity used in this study differ from those applied in the previous chapter, as the FE model and applicable injury criteria differ from the multibody (MB) model used previously. For each region, specific injury criteria were selected to predict injuries, as outlined in Table 4.2. Injuries in the reference data were identified using the AIS code (AAAM, 1998), with the identification process considering the body region, the type of anatomical structure involved, and the specific structure or nature of the injury.

Table 4.2: List of body region and injury criteria used for injury prediction with the FE HBM.

Body region	Body structure	Injury criteria	Numerical Injury Identifier[a]
Head	Skeletal	HIC_{15}	1-5-04
	Internal organs	DAMAGE	1-4-any
	Unconsciousness		1-6-any
Thorax	Thorax	Fractured ribs	4-5-02 Excluded: 450214.3
Thigh (femur)	Skeletal	Max. femur compression force	8-5-18

[a] Desribed as REGION90-STRUCTURE-STRUSPEC.

Two injury criteria were used to predict the risk of head injuries. Skeletal injuries were linked to the HIC_{15} criteria (Eppinger et al., 1999; Hertz, 1993; Versace, 1971). Injuries in the soft tissue of the head, considering injuries in the internal organs and unconsciousness, were linked to the DAMAGE criterion (Gabler et al., 2019). Neck injuries were not considered in this study. The number of fractured ribs (NFR) (Forman et al., 2012) was considered for thoracic injuries, using the risk of rib fracture given by the maximum strain of each rib. Femur skeletal injuries were considered for injuries in the KTH body region, which was linked to the maximum compression force in the femur (Kuppa et al., 2001).

The DAMAGE criterion (Gabler et al., 2019), developed at UVa, was used to predict head internal organ injuries and unconsciousness. This criterion has been adopted by Euro NCAP (2023a) since 2023. Furthermore, Östh et al. (2023) found that the DAMAGE criterion outperformed the BrIC criterion in predicting the maximum principal strain (MPS) in a set of vehicle crash tests using THOR-50M and WorldSID-50M. The technical bulletin for brain injury calculation published by Euro NCAP (2023b) applies an injury risk function for AIS 2+ injuries, which was applied in this study. Consequently, the AIS level for skeletal injuries was modified to include AIS 2, ensuring consistent injury severity criteria across the same body region.

Iraeus et al. (2016) estimated the risk of rib cage injuries, specifically classified as two or more fractured ribs (NFR 2+) and three or more fractured ribs (NFR 3+), corresponding to AIS 2+ and 3+, respectively, using an HBM. These estimates were then compared to injury risk data obtained from real-world crashes reported in NASS CDS. However, some AIS codes using AAAM (1998) have uncertainty when describing NFR 2+ and 3+. Therefore, this study just focused on NFR 2+ injuries, where there is no such uncertainty.

The NFR2+ risk was calculated using a probabilistic age- and strain-based rib fracture risk method (Forman et al., 2012). This method employed the maximum first principal strain observed in each rib during simulation, applied in a strain-based and age-adjusted injury risk function (Larsson et al., 2021a). Each rib's fracture risk was then combined to obtain the risk of NFR2+ and the expected number of broken ribs.

The femur skeletal injuries included AIS 2+ injuries. This ensured the uniformity of injury severity across all body regions in comparison with the reference data.

A binary variable was created to describe if the occupant sustained an AIS 2+ injury in these body regions. Furthermore, MAIS 2+ for each body region was calculated. Table 4.3

shows the different injury risk functions used to predict the injury risk of sustaining an AIS 2+ injury using the FE HBM.

Table 4.3: List of injury risk functions used with the FE HBM.

Body region	Injury criteria	Injury risk function	Reference
HFN	HIC15	$P_{AIS\,3+}(HIC_{15})$[a]$= \Phi\left(\frac{ln(HIC_{15})-\mu}{\sigma}\right)$	Eppinger et al. (1999)
	DAMAGE	$P_{AIS\,2+}(DAMAGE) = 1 - exp^{-\left[\frac{0.017+0.957DAMAGE}{0.459}\right]^{3.875}}$	Euro NCAP (2023b)
THO	Rib strain (per rib)	$P_{fracture\_rib}(strain,age)$[b]$= 0.5 + 0.5erf\left[\frac{ln(strain)-(\beta_0+\beta_1 age)}{\sqrt{2}\alpha}\right]$	Larsson et al. (2021a)
KTH	Max. Femur Compression Force [kN]	$P_{AIS\,2+}(F_C) = \frac{1}{1+exp^{5.7949-0.5196F_C}}$	Kuppa et al. (2001)

[a] Where Φ = cumulative normal distribution, $\mu = 6.96352$ and $\sigma = 0.84664$.
[b] Where erf = error function, $\beta_0 = -2.9866$, $\beta_1 = -0.0130$ and $\alpha = 0.3026$.

4.2.3 Development of the metamodel

Table 4.2 lists the injury criteria measured from the simulations using the computer model, which were utilised to develop the metamodel. This process followed the methodology described in Section 3.2.3. While the criteria for the HFN and KTH regions followed the same procedure, the THO criterion was handled differently. The number of fractured ribs was calculated by measuring 24 rib strains in each simulation and evaluating the risk of fracture within each rib. The rib fracture risk was calculated using the maximum strain within the rib and the occupant's age. The aggregation of the risks for all ribs provided the likelihood of sustaining NFR 2+ injuries. Since the model's anthropometry was not age-dependent, the metamodel was designed to predict maximum rib strains alone, which were subsequently combined with age when evaluating fracture risk to predict the injury outcome.

The iterative process used to develop the metamodels, as explained in Subsection 3.2.3, involved 50 simulations per iteration, with one simulation for each morphed HBM. Since 200 simulations were initially run to verify the stability of the morphed HBMs on the GVI model, these simulations were leveraged to create the initial metamodel and start the subsequent iterations.

All non-normally terminated simulations were discarded. Following the methodology in Chapter 3, once all required computer simulations were completed, the metamodels were developed using the least absolute shrinkage and selection operator (LASSO), neural networks (NN), support vector regressions (SVR), and Gaussian process regression (GPR). Table G.1 lists the set of hyperparameters used for these metamodels to prevent overfitting and underfitting the training data. This process employed a K-fold cross-validation method (James et al., 2021) with K equal to five. To improve regression performance, all input variables were scaled to a zero-to-one range before fitting the metamodels. Additionally, output parameters were also scaled to zero-to-one for the regressions using NN. SVR and GPR regression methods were the only ones applied to the maximum strain per rib simulation output.

Both the magnitude of the prediction error and the stability of the prediction error, as defined in Section 3.2.3, were applied in this study. These levels were determined based on a percentage of the reference values in Table 4.4, referenced as threshold values in the table. Regarding those criteria modified in this study, the threshold for the DAMAGE criteria was

fixed to 0.42, equivalent to 50 % risk of sustaining an AIS 2+ injury, and the threshold for the number of fractured ribs criteria was fixed to 2 expected fractured ribs. The expected number of fractured ribs was calculated using Equation 4.1 from Iraeus et al. (2021). Where P_i was calculated for each rib using the probability rib fracture function ($P_{fracture\_rib}(strain, age)$), from Table 4.3, with an age equal to 45 years.

$$\text{Expected number of fractured ribs} = \sum_{i=1}^{24} P_i \tag{4.1}$$

Table 4.4: List of injury criteria and threshold values used for injury prediction with the FE HBM.

Body region	Injury criteria	Threshold value
HFN	HIC_{15}	700
	DAMAGE	0.42
THO	Number fractured ribs	2 Expected fractured ribs (45 year-old)
KTH	Max. Femur Compression Force	$7.56kN$

As defined in Section 3.2.3, the convergence of the metamodels was achieved when the prediction error was below 10 % of the threshold values, and the stability of the prediction error was below 3 % of the threshold values for the variation of prediction error between iterations.

The training samples were generated with some modifications to the reference data distributions. As stated in Section 3.2.3, the restraint system configuration distribution was adjusted to improve the convergence speed of the metamodel. Specifically, the distributions of the inclusion of load-limiting and pre-tensioning devices were modified. For detailed information on these adjustments, refer to Section 3.2.3.

The computer model used in this study also included two groups of variables, which were not present in the model used in Chapter 3. The first group comprised occupant variables, specifically sex, height, and weight, as the occupant's age was not considered in the morphing process. As mentioned above, each loop iteration of the metamodel obtention process involved one simulation with each of the 50 morphed HBMs.

The second group of variables contemplated the configuration of the GVI model, which included the airbag size and pressure, steering wheel compression force, and knee bolster stiffness. Two values were used for each of these GVI configuration variables, one with a low value and another with a high value. These values were chosen using the normal distributions described in Iraeus et al. (2016) by taking the mean value of the distribution and adding or subtracting the standard deviation to obtain the high and low values, respectively. All possible combinations were generated using a full factorial design, and the configuration of each simulation was created by selecting from those combinations according to a uniform distribution. Appendix E contains the exact values and the validation of the GVI model.

4.2.4 Injury analysis and prediction of AIS 2+ occupant injuries in a wide spectrum of impact conditions

The deterministic and probabilistic methods described in Chapter 3 were employed, utilising the developed metamodels, to predict whether occupants sustained injuries in the studied body regions. The functions used to compute the risk of sustaining AIS 2+ injuries are shown

in Table 4.3. Table 4.4 shows the threshold values used with the deterministic method to calculate whether there was or was not an AIS 2+ injury.

The sample used to compute the distribution of injuries was generated using the Sobol sequence (Bratley et al., 1988). The initial parameters for starting the sequence were the same as those used in Chapter 3. Consequently, the impact conditions for both samples were identical, with the addition of occupant variables for this study. The metamodels for each injury criterion used this sample to obtain the response of the human surrogate, and the deterministic and probabilistic methods were then employed to compute the distribution of injuries based on these results.

The dataset used to compute the distribution of injuries was expanded until a minimum sample size was achieved, ensuring robust computation of the injury distribution. This minimum sample size was determined by applying the criteria outlined in Subsection 3.2.5, ensuring that repeated computations would result in consistent results.

The predictions generated by the two methods were compared based on the percentage of predicted injured occupants, injury patterns, and the distribution of predicted injuries as a function of the delta-v in the collision.

4.3 Results

4.3.1 Real-world data

The dataset of real-world crashes used in Chapter 3 was also utilised in this study. Injuries for the Head-Face-Neck (HFN), Thoracic (THO), and Knee-Thigh-Hip (KTH) body regions were recalculated for AIS 2+ injury severity across 9,957 crashes. Table 4.6 details the MAIS 2+ injuries sustained by occupants in the HFN, THO, and KTH body regions. The HFN region was the most frequently injured, with 281 MAIS 2+ cases, focusing solely on head injuries, followed by the THO region with 173 MAIS 2+ cases. The KTH region had the fewest injuries, with 70 MAIS 2+ cases. In the HFN region, head skeletal injuries were observed in less than 5 % of the cases. In contrast, internal organ injuries or unconsciousness were noted in 98.2 % of the cases with MAIS 2+ in this region.

The top subplot in Figure 4.7 displays the injury patterns observed in the reference data. Each subplot is divided into three columns, representing the three body regions analysed in this study. The total height of each column corresponds to the total number of samples, with subdivisions indicating the proportion of cases with injury severities lower than AIS 2 (i.e., AIS 0-1) and those with injuries of AIS 2 or greater (AIS 2+). Injury severity is categorised by body region, with the maximum AIS (MAIS) shown for each, and classified into MAIS 0-1 and MAIS 2+. The coloured horizontal areas within each column correspond to the different injury patterns identified in the reference data, with the width of these areas reflecting the relative frequency of each injury pattern. In total, eight distinct injury patterns were identified in the reference data, each represented by a colour indicated in the legend of Figure 4.7.

The most usual pattern found in the reference data was an occupant with no AIS 2+ injuries in any of the three considered body regions (97.4 % of the cases), followed by occupants with only MAIS 2+ HFN injuries (orange area, 2.4 % of the cases), MAIS 2+ injuries only in the THO region (blue area, 1.3% of the cases), and then MAIS 2+ injuries only in the KTH (yellow area, 0.5% of the cases). All other injury patterns accounted individually for less than 0.5 % of cases.

Figure 4.8 provides a clearer view of the injury patterns for occupants who sustained moderate injuries in the reference data. While Figure 4.7 is dominated by cases with no AIS 2+

injuries, this second figure focuses exclusively on the 4.7 % of occupants who experienced at least one AIS 2+ injury in the studied body regions.

4.3.2 Development of the metamodel

A total of 1,400 samples were used to train the metamodels, including the 200 samples initially used to check the stability of the morphed HBMs. Table 4.5 presents the magnitude of the prediction error as a percentage of the reference value for each injury criterion in the final iteration of the training process. As observed in the table, not all the injury criteria reached the convergence criteria for an acceptable or good magnitude of the prediction error defined in Subsection 3.2.3.

Table 4.5: Magnitude of the prediction error for each injury criteria and regression method.

Injury criteria	MAE (% of reference value)			
	LASSO	SVR	NN	GPR
HIC_{15}	34.8	35.9	33.3	24.9
DAMAGE	14.6	14.3	14.4	8.8
Expected number of broken ribs	24.5	-	-	11.0
Max. Femur Compression Force (Left)	2.1	3.3	2.2	1.1
Max. Femur Compression Force (Right)	2.1	3.4	2.1	1.1

The metamodels obtained using GPR resulted in the lowest prediction error for all injury criteria (with a mean value of 9.4 %), followed by those obtained with NN (a mean value of 13.0 %). Figure G.1 shows the evolution of the prediction error with the increase of the sample size in the training process of each metamodel. The HIC criterion was the metamodel with the worst performance and did not fulfil the criteria from Subsection 3.2.3 to achieve convergence. Although this criterion did not converge, the iterative training process was stopped. This decision was taken after observing all other criteria converge to an acceptable or good magnitude of the prediction error and analysing the converging process for the HIC criterion.

The FE model predictions for the HIC criterion ranged from 1 to 10,000. The distribution of these results showed 75 % of the predictions below a value of 220 and 90 % below 1,500. In addition to monitoring the evolution of the prediction error using the testing sample, the training error was also analysed in this case. Figure 4.4 illustrates the evolution of the MAE measured using the training sample. While a decrease in error was observed for the first 1,000 samples, no significant reduction occurred in the subsequent 400 samples, with the error stabilising above the 15 % threshold of "acceptable" prediction error. The source of this error in the training sample is more clearly demonstrated in Figure 4.5. Subfigures (a) and (b) display the same data, with subfigure (b) presenting a zoomed-in view of the range from 0 to 1,000. The black dashed line represents a one-to-one correlation between the metamodel's actual and predicted values using the metamodel. As seen in both subfigures, there was noticeable dispersion in the values at both small and large values, resulting in a poor fit of the training data.

For the other injury criteria, a sample size between 500 and 600 crashes was sufficient to achieve an acceptable prediction error. As observed in the previous chapter, the femur compression force criterion had the lowest prediction errors relative to the limit. Figure G.2 compares the actual simulation values with the predicted values from the GPR metamodel.

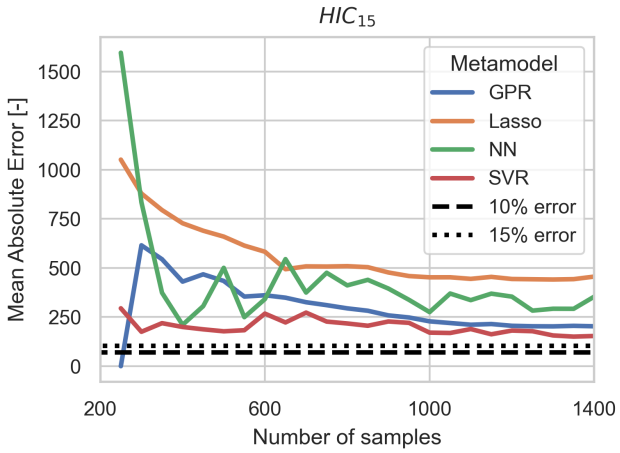
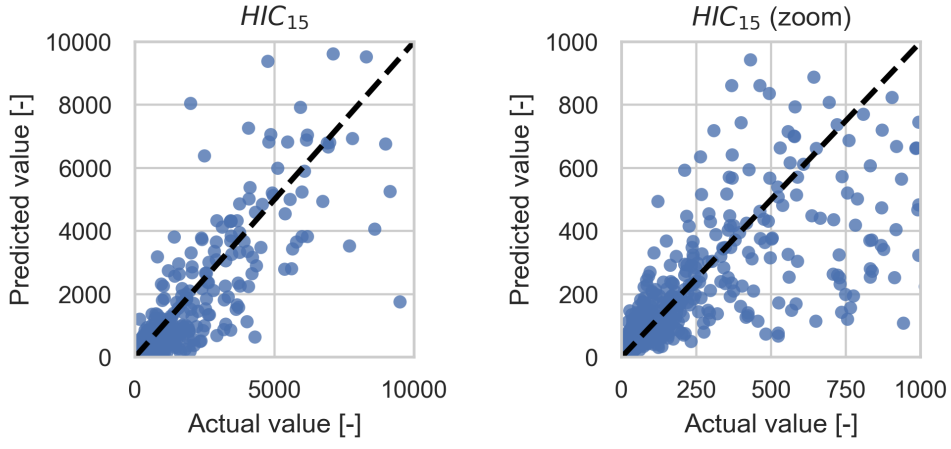


Figure 4.4: Mean absolute error (MAE) obtained for the HIC criterion metamodel using LASS, SVR, NN, and GPR. The horizontal axis shows the number of samples used to train the metamodel, and the vertical shows the MAE of the predictions using the metamodel. The dashed black line displays the MAE values that have an error of 10 % of the threshold criteria value ("good" magnitude of prediction error), and the dotted black line displays the MAE values that have an error of 15 % of the threshold criteria value ("acceptable" magnitude of prediction error).



(a) (b)

Figure 4.5: Actual vs. predicted values of the 1,400 simulations used to develop the GPR model for the HIC criterion. Subfigure "a" displays values in the range of 0 to 10,000, while Subfigure "b" focuses on the range of 0 to 1,000.

Among the injury criteria, the HIC criterion exhibited the lowest R^2 value, followed by the DAMAGE criterion.

An additional analysis was performed to evaluate the ribs criterion, which employed two different metrics for the deterministic and probabilistic approaches. The first metric assessed the expected number of broken ribs, ranging from 0 to 24. The second metric calculated the risk of NFR 2+, which ranged from 0 to 100 %. However, the NFR 2+ risk tends to saturate when two or more ribs are expected to be broken, and the convergence criterion

was based only on the expected number of broken ribs. To provide a more comprehensive assessment, the risk of NFR 2+ was calculated for 1,400 samples using both the simulation and metamodel outputs.

Figure 4.6 displays the risk of NFR 2+ for the evaluated samples, which clustered predominantly in two distinct regions: one with a low risk of NFR 2+ (less than 20 % risk) and another with a high risk of NFR 2+ (above 80 % risk). These regions accounted for 62.0 % and 17.2 % of the samples, respectively, while 20.8 % of the samples fell within a buffer zone between these two regions. Within this buffer zone, 50 % of the samples exhibited an absolute difference of less than 6.5 % in risk compared to the actual predictions, and 75 % had a difference of less than 15.5 %. The remaining 25 % of cases, where larger discrepancies were observed between actual and predicted values, were primarily characterised by underpredictions.

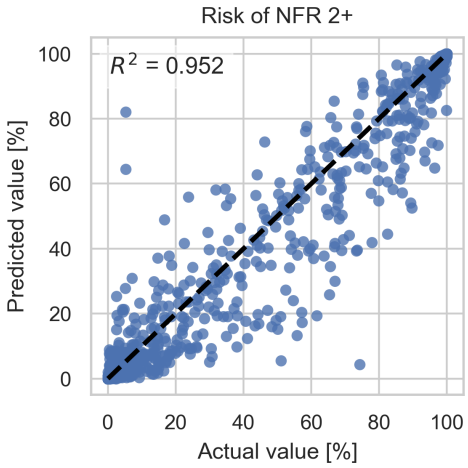


Figure 4.6: Actual vs. predicted values of the risk of NFR 2+ for the 1,400 simulations used to develop the GPR model.

4.3.3 Prediction of seriously injured occupants in a simulated population

The obtained metamodels were used to predict AIS 2+ injuries in the studied body regions based on the impact conditions from the reference data. The HIC criterion for head skeletal injuries was excluded as it did not meet the requirements during the convergence process. Consequently, HFN injuries were predicted using only the DAMAGE criterion. Although the occupant's age was not included as an input in any metamodel, it was used in combination with rib strains to predict the fracture risk for each rib.

Figure G.5 displays the percentage of MAIS 2+ injuries in each body region for occupants at various sample sizes. Following the method applied in the previous chapter, a total of 25,000 samples were generated, as the change in the percentage of MAIS 2+ injuries in each body region stabilised below 0.1 %, as shown in Figure G.6. This figure illustrates how the ratio of MAIS 2+ to MAIS 0-1 predicted injured occupants evolved according to the proposed convergence criteria.

Table 4.7 presents the ratio of MAIS 2+ injuries relative to the total number of simulations predicted for each body region using both the deterministic and probabilistic methods. For example, the deterministic method estimated that 3.2 % of the population would sustain an

Table 4.6: AIS 2+ injured occupants observed in the reference data grouped based on the body region.

Body region	n	% relative to total	% relative to body region
HFN injuries	281	2.8	
Head skeletal injuries	13	0.1	4.6
Head internal organ injury or unconsciousness	276	2.8	98.2
THO injuries	173	1.7	
KTH injuries	70	0.7	

Table 4.7: Body region and source of the MAIS 2+ injuries predicted using the deterministic and probabilistic methods.

Body region	Deterministic method	Probabilistic method
HFN injuries[a]	3.2	10.8
THO injuries	19.9	25.1
KTH injuries[b]	0.0	1.2

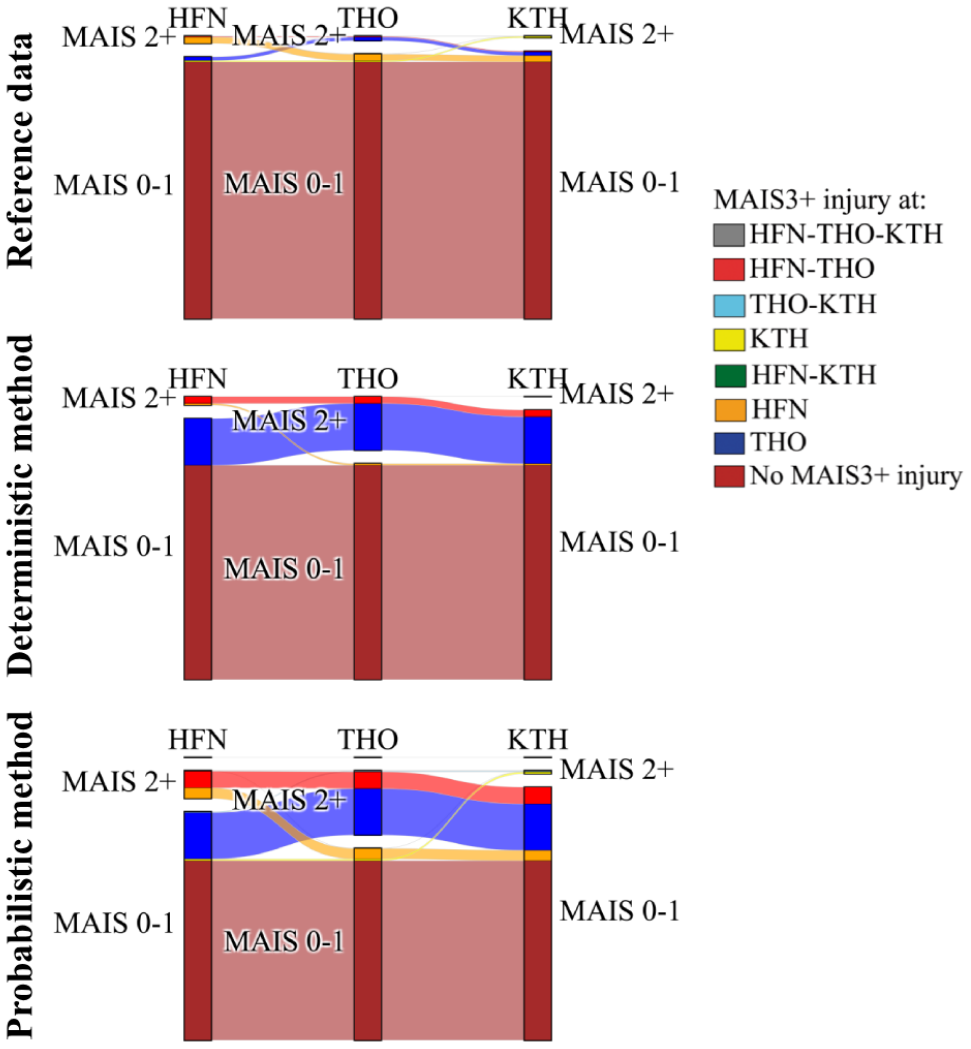
[a] Takes into account only DAMAGE criterion.
[b] Takes into account the left and right femurs simultaneously.

MAIS 2+ injury to the HFN body region. As noted in the previous chapter, no injuries were predicted for the KTH region using the deterministic approach. Overall, both methods showed a tendency to overpredict injuries compared to the reference data.

Focusing on head injuries related to internal organ damage and unconsciousness, both methods overpredicted the injury distribution compared to the reference data. The deterministic method over-predicted MAIS 2+ injuries by 1.1 times in relation to those observed in the reference data, while the probabilistic method over-predicted the reference data by a factor of 3.9. In the thoracic region, rib fractures were significantly overpredicted by both methods, with overpredictions of 6.4 times and 8.1 times the reference data for the deterministic and probabilistic methods, respectively. Lastly, AIS 2+ injuries in the KTH region were overpredicted by a factor of 1.7 using the probabilistic method.

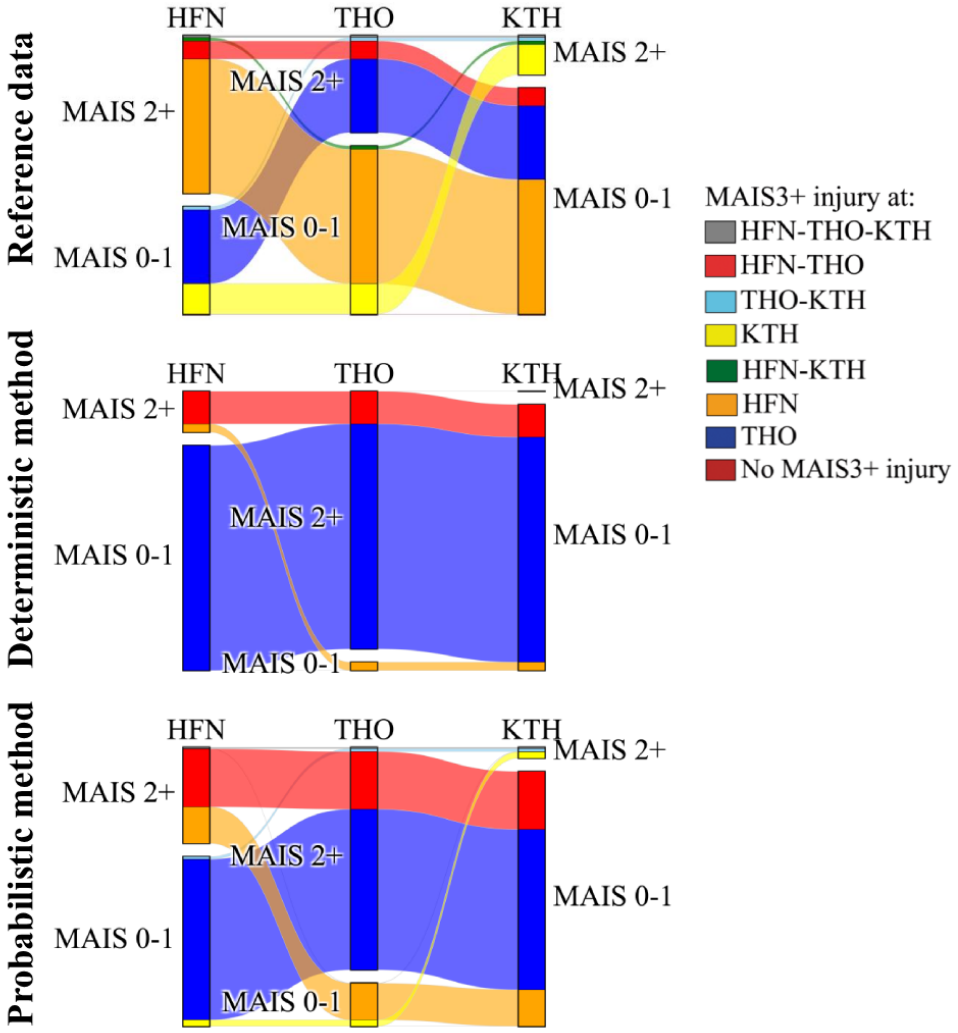
Figure 4.7 compares the injury patterns observed in the reference data with those predicted by the deterministic and probabilistic methods using the metamodel results. As shown, both the deterministic method (middle plot) and the probabilistic method (bottom plot) overpredicted the number of MAIS 2+ injuries in the HFN and THO body regions compared to the reference data (top plot). Additionally, as observed in Chapter 3, the deterministic method did not predict any injuries in the KTH region, which resulted in missing several injury patterns present in the reference data.

Although the percentage of occupants sustaining at least one MAIS 2+ injury in the body regions studied was higher than that observed in the previous chapter, 95.3 % of the occupants in the reference data sustained only MAIS 0-1 injuries. Thus, Figure 4.8 specifically highlights the injury patterns for those who sustained at least one MAIS 2+ injury. On the one hand, the probabilistic method (bottom plot in both figures) consistently overpredicted MAIS 2+ injuries across all regions but was able to capture most of the injury patterns found



Injury Pattern (%)								
Reference data	0.0	0.3	0.1	0.5	0.1	2.4	1.3	95.3
Deterministic method	0.0	1.9	0.0	0.0	0	0.6	17.4	75.9
Probabilistic method	0.0	6.5	0.3	0.7	0.0	4.1	18.1	70.0

Figure 4.7: Injury patterns observed in the reference data and applying the deterministic and probabilistic methods.



Injury Pattern (%) relative to injured occupants								
Reference data	0.8	6.8	1.3	11.4	1.3	50.5	27.7	-
Deterministic method	0.0	12.3	0.0	0.0	0.0	3.1	86.4	-
Probabilistic method	0.5	21.8	1.1	2.5	0.1	13.7	60.3	-

Figure 4.8: Injury patterns observed in the reference data and applying the deterministic and probabilistic methods, without considering cases where no AIS 2+ injury was reported for the studied body regions.

in the reference data. The most relevant injury partners in these cases were injuries to the THO, HFN, and THO-HFN combined. However, their relative frequency did not match the observed in the reference data. On the other hand, the deterministic method (middle plot) did not predict certain injury patterns. The threshold used for the femur injuries did not result in any AIS 2+ injuries for this region, even with the inclusion of a wider spectrum of occupant anthropometries.

Regarding the conditions under which injuries were predicted, Figure 4.9 presents a box plot comparing results from the reference data with predictions from the deterministic and probabilistic methods. MAIS 0-1 and MAIS 2+ injury outcomes for the THO region were analysed as a function of delta-v. The whiskers in the box plot represent the 5th and 95th percentiles for each category. As shown in the figure, 95 % of the MAIS 0-1 injured occupants in the reference data experienced delta-vs below 53 km/h. For MAIS 2+ injuries in this region, injured occupants were observed across the entire range of delta-vs, with the 5th percentile observed at 17 km/h. In the results from both the deterministic and probabilistic methods, 95 % of the MAIS 0-1 injuries were associated with delta-v below 42 km/h. Regarding MAIS 2+ injury outcomes, the 5th percentile delta-v was observed at 26 km/h for the deterministic method and 24 km/h for the probabilistic method.

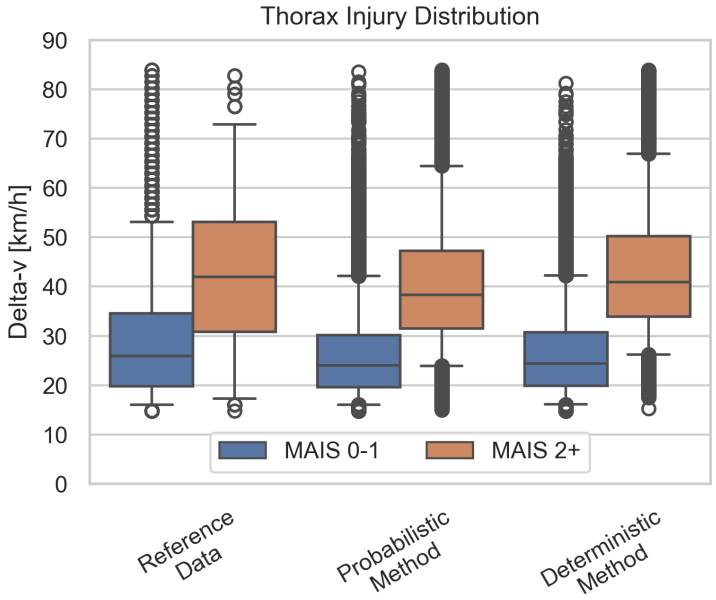


Figure 4.9: Box plot of MAIS 0-1 and MAIS 2+ thoracic injuries observed in the reference data and applying the deterministic and probabilistic methods, as a function of the delta-v in the collision. Whiskers in the boxplot represent the 5th and 95th percentile of the observed data.

4.4 Discussion

This chapter explored the proposed methodology from Chaper 3 to predict injury outcomes using FE models. Real-world injury data and data from computer model predictions were used to benchmark these predictions.

4.4.1 Computer model

The GVI model was modified and validated as described in Appendix E, using three different configurations of the load-limiting device to compare against crash test data. Although the GVI model's response fell within the reference data corridors, it did not fully capture the variation observed in the crash test results. These corridors were established by considering the 5^{th} and 95^{th} percentile measurements at each time point. Similar to Larsson et al. (2024) validation of the passenger side GVI model and Boyle et al. (2024) development of a parametric GVI model, the GVI model used in this study showed a tendency to over-predict chest deflection in the Hybrid III 50^{th} percentile male model relative to the crash tests measurements.

While this approach ensured that the model's response was aligned with the crash tests, no correlation method was employed to quantitatively assess the similarity of the measurements. Future validations of this model could benefit from employing correlation methods, such as the CORA method (Gehre et al., 2009; 2011) and ISO/TS 18571 (International Organization for Standardization, 2024), to provide a more rigorous evaluation of the model's accuracy. These methods enable a quantitative comparison between the model's output and reference data. Additionally, further exploration of specific GVI model configurations could help identify setups that align more closely with the reference data, such as larger airbags combined with a low load-limiting force. These configurations could be applied at higher frequencies when conducting large-scale simulations, potentially improving the overall predictive capability of the model.

Perez-Rapela et al. (2020) used a total of 405 morphed HBMs to create a metamodel capable of predicting outcomes in far-side impacts. Although this process was automated to seat the model and fit the seat belt before running crash simulations, these pre-simulation tasks can be as computationally demanding as the crash simulations themselves. An alternative approach was used to reduce computational cost while maintaining similar prediction accuracy. Larsson et al. (2024) studied how many anthropometries are necessary to generate morphed HBMs that can represent the population's response in crash scenarios. While their study concluded that as few as 14 anthropometries could be used for a specific set of crash conditions, this recommendation applies to a unique scenario. Considering potential non-linearities in crash outcomes, Larsson et al. (2024) recommended using at least 25 male and 25 female occupants, evenly distributed, to ensure adequate population representation.

Following this recommendation, the SHBM was morphed into 50 different anthropometries to represent the population observed in the reference data. It is important to note that this approach may be conservative, as most crashes in the reference data involved a delta-v below 50 km/h (approximately 90 %), while the impact condition in Larsson et al. (2024) was set at 50 km/h. Although the lower delta-v may lead to similar injury criteria predictions across various occupants, including all 50 anthropometries may still be necessary to capture the non-linearities in the model's response. Testing the complexity of this response, while ensuring the stability of the model, could provide valuable insights into determining the optimal number of occupants to include. This could allow for a reduction in the number of anthropometries used without compromising model accuracy, ultimately saving computational costs.

The morphing of the SHBM focused on height, weight, and sex variability to represent the occupants in the reference data. However, other factors potentially relevant to injury outcomes were not included in this process, such as material properties. While the morphing algorithm generated a mean statistical model based on height, weight, and sex, individual variations within the population sharing these characteristics may still significantly influence injury outcomes.

For instance, Larsson et al. (2022) and Larsson et al. (2023) explored the impact of local variations in the ribs on the prediction of rib fractures, highlighting factors such as rib cage shape, rib cross-section, cortical bone thickness, and material properties. Although some variability in these factors can be indirectly captured by height, weight, and sex, individual deviations from the mean statistical shape were not considered. These deviations can be important for calculating accurate injury outcomes. However, the primary focus of this study was on developing a method for such calculations, and as a result, these individual variations were not incorporated. Future studies can expand on this work by including a wider range of anatomical variations to enhance the predictive accuracy of injury outcomes.

Variations in occupant position and seatbelt routing were not considered in this study. Corrales et al. (2024) demonstrated that factors such as rib cortical material and its constitutive model are critical when predicting the response of aged occupants. Moreover, occupant posture, including thorax curvature and ribcage morphology, was shown to influence injury outcomes significantly. However, a fixed age of 45 years was used in this study to calculate occupant posture and limit the number of model variations. Other variations that could affect occupant response were also not included, such as out-of-position postures (Leledakis et al., 2021b), seat position (Leledakis et al., 2022; Ressi et al., 2022), and seatbelt routing (Leledakis et al., 2023).

The inclusion of the aforementioned individual variations in the proposed method could introduce potential limitations when combined with metamodels. While metamodels allow for simulating a subset of the total cases needed to estimate injury outcomes, adding more variables could substantially increase the dataset size required for training. Perez-Rapela et al. (2020) tackled this challenge by reducing the number of regression variables during the training process. Alternative approaches can include clustering input variables to reduce dimensionality, as suggested by Xu et al. (2015). However, individual variations might be classified as noise or less significant factors in the training process, potentially limiting their study using the metamodel. A possible alternative would involve estimating the model's response to individual changes based on the results from the mean statistical anthropometry, as proposed by Solhed (2022). This approach could balance metamodel complexity with the inclusion of individual variability, allowing the metamodel to account for these variations without drastically increasing the computational cost.

4.4.2 Development of the metamodel

Regarding the performance of the metamodels, the lowest prediction errors were achieved using the GPR method, which was consistent with the findings from the previous chapter. When examining the risk of NFR 2+, a slightly lower R^2 value of 0.95 was observed compared to the higher R^2 reported by Larsson et al. (2024) for frontal impacts during metamodel development. This reduction in R^2 could be attributed to the use of a limited selection of anthropometries to represent the entire population, as well as variations in the impact conditions not included in the referenced study. However, when focusing on the expected number of fractured ribs (NFR), as illustrated in Figure G.2, the R^2 value was 0.98, closely matching the results found in the aforementioned study.

In contrast with the findings of the previous chapter, neither the GPR method nor the other used in this study was able to reach an acceptable prediction error for the HIC criterion. A review of simulations where the difference between predicted and actual HIC values exceeded 1,000 revealed two primary characteristics that could explain these discrepancies. Alternative scaling methods were tested to determine if they could improve the results. These

included applying a log transformation to the HIC values to reduce the data's range across four orders of magnitude and using the injury risk function so that the regression does not try to overfit large HIC values. However, neither method achieved the "acceptable" level of prediction error. Thus, this criterion was not included in the subsequent analysis of predicted injuries.

The first characteristic was the late or non-activation of the airbag, leading to head impacts with the steering wheel, resulting in excessively high HIC values. The second characteristic involved the inclusion of larger occupants (i.e., individuals taller than 180 cm) in the simulation, where head contact with the roof liner occurred. This contact did not account for the deformation of the roof liner, which may have been omitted in the original model (Iraeus et al., 2016), as the occupant in the original model was not as tall as those in these cases. The absence of this consideration of contact likely contributed to inaccuracies in the predicted HIC values. While Larsson et al. (2024) also included tall occupants in their simulations, their focus on front-seated passengers did not lead to any head impacts with the steering wheel, and the positioning of the seat likely prevented contact with the roof liner. Future studies could focus on better characterising the interaction between the head and the roof liner, particularly for taller occupants, to improve the model's accuracy in such scenarios.

While these characteristics did not significantly impact the metamodel for the DAMAGE criterion, the HIC criterion was highly sensitive to such scenarios. However, this sensitivity was not accurately captured by the metamodel. None of the regression methods used in this study were able to adequately reflect the sharp peaks in the HIC criterion, exposing limitations in the current approach when non-linearities in the model lead to large, isolated differences in occupant response. An adaptive sampling method could have mitigated these large errors by increasing the number of samples around problematic conditions, focusing on simulations with similar characteristics. This approach would allow the model to better account for non-linear responses and improve the accuracy of the metamodel in regions where significant deviations occur.

Schneider et al. (2022) developed a framework for an adaptive sampling method. This framework selects new samples during the metamodel training process by identifying regions of interest based on the prediction error of the metamodel. However, there is a possibility that the newly defined region of interest may not fully align with the specific focus of the current study, potentially limiting the framework's effectiveness. In this chapter, head skeletal injuries were observed in less than 5 % of the cases in the reference data, while head internal organ injuries or unconsciousness were present in 98 % of the cases. Given this distribution, allocating significant computational resources to improve predictions for head skeletal injuries may not be justifiable, especially when the focus could instead be placed on the more prevalent head internal organ injuries and loss of consciousness. A complementary approach could involve using a smaller, more targeted input space to study head skeletal injuries, allowing for more efficient resource allocation.

An alternative method to improve the prediction accuracy of the metamodels could involve applying a different scaling technique to the data used in the training process. Amorim et al. (2023) demonstrated that the difference between the best and worst scaling techniques is statistically significant in most cases. Therefore, selecting an alternative scaling method could potentially help capture the peaks observed in the HIC criterion without compromising the overall accuracy across the entire training dataset. However, incorporating different scaling methods into the current methodology may significantly increase the computational cost of the regression process. This is because each combination of scaling methods and metamodel

hyperparameters would need to be tested, leading to a larger number of regression processes overall.

4.4.3 Prediction of injury outcomes and assessment of the deterministic and probabilistic approaches

The primary change in the reference data compared to the previous chapter was the focus on a different injury severity level. In this study, AIS 2+ injuries were calculated for the selected body regions. Another significant change was the use of an advanced injury criterion, which predicts injury outcomes for specific anatomical structures based on element strain. These modifications led to some differences in the observed injury outcomes within the reference data. Head injuries emerged as the most frequent injury type, while the thoracic region ranked second, primarily due to the focus on rib fractures. If all thoracic injuries were considered, this region would be the most frequently injured.

The metamodels were employed to predict injury outcomes for a population of 15,000 individuals. This population size was more than ten times greater than the number of simulations used to develop the metamodel and 25 times larger than the minimum number of simulations required to achieve similar accuracy. Additionally, this population was 50 % larger than the one used in the previous chapter. The increased population size could be attributed to the inclusion of variations in occupant anthropometry and age, which broadened the scope of the analysis and allowed for a more comprehensive assessment of injury outcomes across a diverse group of individuals.

Regarding the prediction of injury outcomes using both the deterministic and probabilistic approaches, neither method fully replicated the injury distribution observed in real-world reference data. This discrepancy was evident whether injuries in a body region or injury patterns were considered. Although both methods provided valuable insights, the real-world injury patterns were not entirely captured by either approach.

The results from both the deterministic and probabilistic approaches applied to the FE model correlated with the findings from the previous chapter, which used an MB model, despite accounting for different injury severity levels and injury criteria. A prominent characteristic of the estimations was the over-prediction of injury outcomes compared to the reference data. Furthermore, no injuries were predicted for the KTH region, similar to the previous chapter, even though this study included obese occupants, which previous studies have reported to be at a higher risk of injuries in this area (Carter et al., 2014; Enayatollah et al., 2020; Jakobsson et al., 2005; Joodaki et al., 2020b; Rupp et al., 2013). The absence of predicted injuries in the KTH region resulted in missing several injury patterns observed in the reference data, some of which were captured by the probabilistic method. A detailed discussion of the injury outcomes for each body region is provided in the following paragraphs.

Regarding injuries in the KTH region, injuries just in the femur were included in the study. Although other injuries could be at a higher risk compared to the femur, such as pelvis injuries (Pipkorn et al., 2020), specific injury criteria must be developed and validated for the HBM used in the study to predict those injuries. As mentioned above, the predictions with the deterministic and probabilistic methods were totally different. These differences could be attributed to three main factors or a combination of them. The first factor is the injury criteria applied to this region. More advanced injury criteria could have been applied to this region using strain or stress-based injury criteria, such as the developed by Schubert et al. (2021) for the VIVA+ HBM.

The second factor influencing the lack of predicted injuries in the KTH region is the GVI model used in the study and how the interactions between the HBM and the knee bolster were modelled. Ressi et al. (2024) studied the replication of real-world trends using simulations with THUMS and VIVA+ HBMs seated in the GVI model. Their results showed a lower injury risk to the lower extremities in relation to the reference data, and they identified the penetration-based contact between the HBM knees and the knee bolster as a potential cause. This contact allows for significant penetrations into the knee bolster surface, potentially failing to capture the full interaction between the lower extremities and the vehicle. Furthermore, the validation process in Appendix E highlighted another factor that could contribute to this lower injury risk. None of the chosen restraint system configurations were able to replicate the highest femur forces observed in the reference data. The largest compression forces recorded in crash tests ranged from 4 to 6 kN, depending on the restraint system configuration used in the vehicle, whereas the simulations produced compression forces in the range of 1 to 2.5 kN including all cases. Although these peak forces align perfectly with the mean responses observed in the crash tests when those vehicles included a load-limiting device, it was within the responses observed for vehicles without load-limiting devices but under their mean response. It is worth noticing that the stiffnesses used in the validation process are below the mean stiffness used in the development of the GVI model (Iraeus et al., 2016).

A second set of simulations was conducted to study the effect of using a stiffer knee bolster on the GVI model. A stiffness of 12 kN/100 mm was applied. Although this nearly doubled the highest value used in the initial validation process, this stiffness was 60 % of the maximum stiffness used in the development of the GVI model (Iraeus et al., 2016).

Figure 4.10 illustrates the impact of this modification on femur forces. The figure presents three distinct corridors: the blue corridor represents the forces observed in crash tests, the orange corridor corresponds to the response with validated stiffness values, and the green corridor shows the results from simulations using the increased stiffness. The results showed that the higher stiffness resulted in greater compression forces in both femurs. For vehicles equipped with a load-limiting device (LL), the increased stiffness pushed the compression forces toward the outer limits of the NCAP test corridor (i.e., the reference data). For vehicles without a load-limiting device (No LL), the compression forces matched the mean response but remained significantly below the peak compression forces of 6 kN observed in the reference crash data. To achieve these peak forces, even stiffer knee bolsters would be necessary. Therefore, when selecting knee bolster stiffness for crash simulations, it is crucial to consider the presence of load-limiting devices.

The third factor is the deterministic and probabilistic approaches used in the study. In the deterministic method, the injury criterion threshold applied was the same as that used in the previous chapter to calculate higher injury severity. Adjusting this threshold to a smaller value might predict more injurious outcomes. However, the injury risk associated with this threshold is relatively low compared to other thresholds used in this study. For instance, a 7.56 kN compression force in the femur corresponds to a 13 % AIS 2+ injury risk (Kuppa et al., 2001), while the 0.42 threshold for the DAMAGE criterion results in a 50 % AIS 2+ injury risk (Euro NCAP, 2023b), and a lower threshold for the femur force will result in a larger discrepancy. Newer injury risk functions could be implemented for this criterion, such as those proposed by Laituri et al. (2005) and Rupp et al. (2010). This disparity in risk values between thresholds raises concerns about the robustness of the deterministic method. Since each injury criterion may require thresholds that estimate very different injury risk.

Regarding the HFN region, the probabilistic method resulted in a larger over-prediction of injuries compared to the deterministic method. Despite using different criteria and models

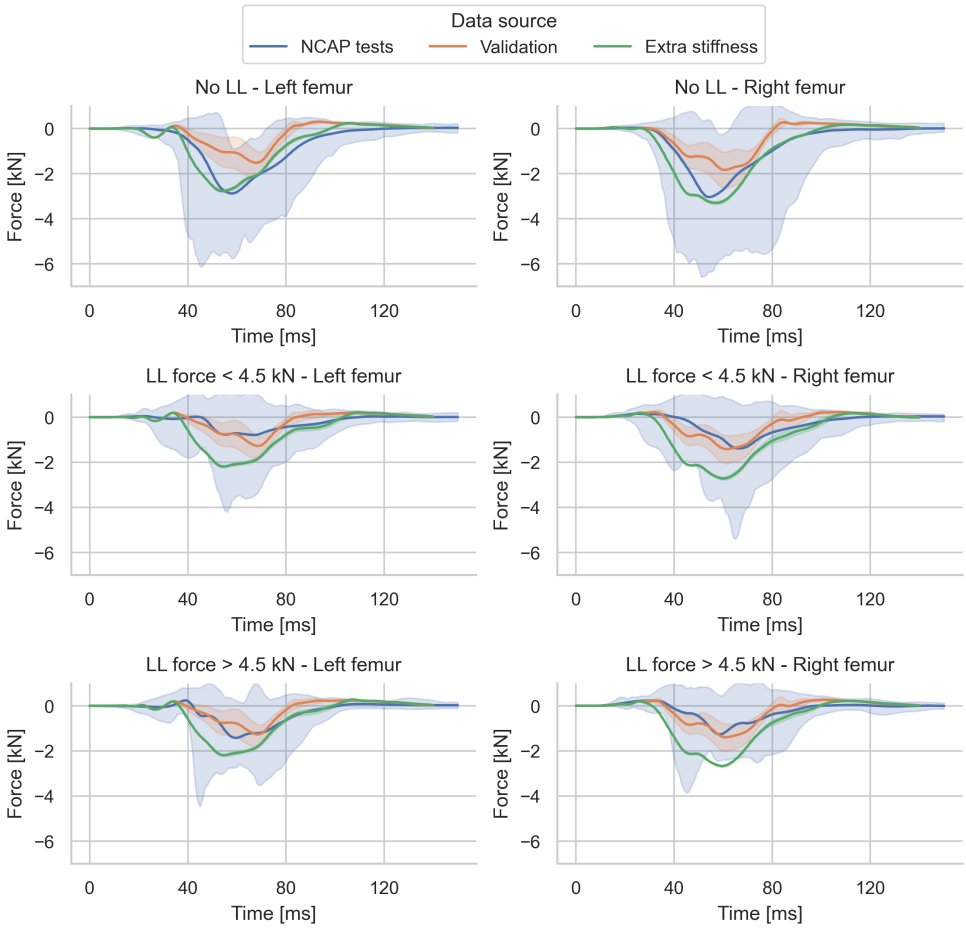


Figure 4.10: Femur compression forces observed in the NCAP tests and the GVI model using the simulations carried out for the validation process and simulations with extra stiffness in the knee bolster.

in the previous chapter, both approaches consistently demonstrated a tendency for the probabilistic method to overestimate injury risks more than the deterministic approach. Specifically, for the DAMAGE criterion, a threshold of 0.42 was identified, corresponding to a 50 % injury risk for an AIS 2+ injury (Euro NCAP, 2023b). To further understand these discrepancies, cases where the probabilistic method predicted moderate injury while the deterministic method did not were reviewed. The values of the DAMAGE criterion in these instances ranged from 0.13 to 0.419, with an associated injury risk of 2 % to 50 %.

A distinct trend emerged from the crash condition analysis: all observed cases involved oblique far-side impacts. This suggests that the DAMAGE criterion may tend to over-predict injuries in these specific scenarios. Although the criterion performs well in certain narrow contexts, it could be overestimating injury outcomes in both reconstructed and real-world crash situations as previously observed in other studies (Östh et al., 2023; Prasad et al., 2024; Ressi et al., 2024). This over-prediction may be tied to the current validation limitations of both the GVI and HBM models, as neither has incorporated head rotational data into their validation processes. The absence of this critical validation component may restrict the applicability of these models and injury criteria in the present study, particularly when evaluating injury risks in oblique impacts.

Lastly, in the THO region, a strain-based injury criterion was implemented to estimate injuries to specific anatomical structures, particularly the rib fracture. This criterion posed the greatest complexity in developing the metamodel and applying the deterministic and probabilistic approaches. The metamodel was designed to estimate the 24 maximum principal strains (one for each rib) from the HBMs, as the metrics used for the deterministic and probabilistic methods were different. If the same metric had been applied for both methods, a single metamodel could have been developed to predict either the risk of NFR 2+ or the ENFR, as was done in Larsson et al. (2024). While it would have been possible to develop two separate metamodels to predict the risk of NFR 2+ and ENFR, it was decided to create a more complex metamodel to avoid inconsistencies in the predictions caused by different performances of the fitting process. This approach ensured consistency in the prediction process across both methods, despite the added complexity in metamodel development.

The results for this injury criterion were the most closely aligned for both methods. This is particularly evident in Figure 4.9, where, unlike the THO region results with the MB model, both the deterministic and probabilistic methods predicted thoracic injuries under similar conditions (i.e., the delta-v of the collision). This convergence suggests that, despite the complexities in predicting thoracic injuries, the use of an expected outcome in the injury criterion led to similar performances with both methods. Unlike other injury criteria, the expected outcome for the THO criterion does not allow for modifications, as it is directly tied to the definition of the injury outcome. While this criterion supports the use of an expected outcome, the calculation of an estimated outcome may not be possible for other injury criteria.

The over-prediction of THO injuries using both methods was expected, as the injury criterion was developed based on an earlier version of the HBM, which already exhibited a tendency to over-predict injury risk (Larsson et al., 2021a) relative to the NASS CDS database. The authors of that study attributed these discrepancies to an under-reporting of rib fractures in the clinical outcomes due to limitations in the CT diagnostics used to identify such injuries, as reported by previous studies (Crandall et al., 2000; Lederer et al., 2004; Schulze et al., 2013).

This chapter presented an estimation of injury outcomes considering the whole population without focusing on the differences in the injury outcomes for the specific anthropometries. Previous research has shown that specific segments of the population, such as elderly, obese, and female occupants, are at a higher risk of death and serious injuries in these body regions (Abrams et al., 2020; 2022; Bose et al., 2011; Carter et al., 2014; Forman et al., 2019; Kahane, 2013b; Noh et al., 2022; Ranmal et al., 2024; Viano et al., 2008). However, it was outside the scope of this study; estimations using the metamodel can be used to explore how the variations in the restraint systems impact each one of these segments of the population.

This chapter provided an estimation of injury outcomes for the overall population without examining variations in injury outcomes associated with specific anthropometries. Previous research highlights that certain population segments, such as elderly, obese, and female occupants, are at an elevated risk of severe injuries and fatalities in particular body regions (Abrams et al., 2020; 2022; Bose et al., 2011; Carter et al., 2014; Forman et al., 2019; Kahane, 2013b; Noh et al., 2022; Ranmal et al., 2024; Viano et al., 2008). While this study did not focus on these subgroup-specific outcomes, the metamodel developed here could be applied to examine how variations in restraint systems might impact each of these vulnerable population segments.

4.5 Limitations

Limitations concerning the reference data used for benchmarking predictions were acknowledged. One key limitation was the lack of weighting applied to the sample data, which could influence the observed frequency of injury patterns. Furthermore, limitations in how the reference data were gathered, such as potential underreporting of rib fractures, may affect the total number of injured occupants, leading to discrepancies between the predicted outcomes and real-world data. Regarding the restraint system configuration, this study did not consider the use of occupant classification systems (OCS) that could modify the restraint system according to the crash conditions.

Additional limitations were associated with the computer models used to predict occupant responses in the simulated crashes. As noted earlier, the validation of the GVI model did not incorporate more advanced techniques to vary the frequency of restraint system configurations. For instance, some combinations may be observed at a higher frequency, such as low-force load-limiting devices paired with larger airbags at lower pressure and knee bolsters with reduced stiffness, among other configurations. These variations could have a significant impact on predicting injury outcomes from the model, as they may not be optimised to protect the occupant. Moreover, the current validation did not fully capture the complete range of responses observed in the reference crash tests, further limiting the model's predictive accuracy. The HBM used in this study also contributed to some of these limitations. Its validation focused primarily on displacements and deformations, while head rotations were less frequently considered. This omission may limit the applicability of certain injury criteria, especially those related to rotational kinematics and head injuries. Individual variations from the mean anthropometry, occupant positions and belt routing were also not considered to limit the number of occupants used to obtain the metamodels.

The third limitation of the study was related to the prediction of injuries. This study focused exclusively on AIS 2+ injuries, which limited the ability to directly compare the results with those of the previous chapter, where a higher injury severity was considered. Furthermore, the inclusion of different anthropometries was not fully accounted for in some injury criteria. For example, the femur compression force injury function was developed based on occupants with an average age of 63 years and a weight of 69 kg. The scaling of predicted forces based on the changes in the anthropometry, as implemented in Hu et al. (2019), was not explored in this study. Limitations regarding the deterministic and probabilistic methods, such as the scaling of thresholds or injury risk functions used in this study, were also not considered.

5 Discussion, future research and conclusions

This final chapter concludes the work presented in this dissertation. First, it provides a discussion of the results outlined in the previous chapters, aligning them with the objectives set in Chapter 1, and potential future research directions are proposed. Following this, the key conclusions drawn from the dissertation are summarised.

5.1 Discussion and future research

The overall aim of this dissertation was to develop a method for assessing the impact of restraint systems on injury outcomes, considering the variability of real-world crash conditions. To achieve this, specific objectives were established to gather essential data and create two complementary approaches to this aim. While this dissertation focuses on a subset of observed real-world crashes, the methods and assessments developed can be applied to other crash configurations. This broader application would enable a more comprehensive understanding of how restraint systems affect injury outcomes across various crash scenarios, offering potential insights for enhancing vehicle safety in a wide range of conditions.

The following subsections provide a detailed discussion of the methodologies employed to achieve the specific objectives detailed in Chapter 1, along with an exploration of the potential limitations of the results provided in each chapter.

5.1.1 Objective 1: To develop a method for identifying the presence and characteristics of restraint system features

As early as 1962, Bertil Aldman stated that "… the occupant must be decelerated in an upright posture. This requires the use of an upper restraint in combination with the lap restraint". In his work, he performed tests with different safety belt systems in which the three-point seat belt was always more protective than any other restraint tested (Aldman, 1962). The three-point seat belt was first incorporated in a Volvo vehicle in 1959, and shortly after, other vehicle manufacturers started to install these restraint systems in their vehicles as well. The introduction of pre-tensioners in the 1980s and load limiters in the mid-1990s resulted in the seat belt systems currently present in most vehicles in high-income countries.

Even if almost all restraint systems incorporate these advanced features to optimise the protection of occupants, pre-tensioners and force-limiters can be tuned depending on other vehicle characteristics, such as internal vehicle dimensions, size of the airbags, structural stiffness, etc., to maximise their effectiveness. Even if these devices have been shown to be very effective in reducing the severity of injuries in computational and laboratory-based experimental studies, a more robust assessment of advanced restraint systems needs to be done using real-world data (Segui-Gomez et al., 2009). Such evaluation requires knowing if advanced restraint system features were installed in the vehicle and their configuration. To the authors'

knowledge, this information is not available in existing real-world collision databases that are publically available.

Modern vehicles employ multiple components to effectively restrain occupants during collisions. This objective focused on identifying the pre-tensioning and load-limiting features implemented in seat belt systems, based on force-time history measurements from crash tests. Using these measurements, an automatic algorithm was developed to identify simplified characteristics to describe the behaviour of the pre-tensioning and load-limiting devices, such as the force applied by the pre-tensioning device, its activation time, and the force used by the load-limiting mechanism. While this approach does not fully capture the complexity of these devices, the selected characteristics are widely used in the literature for their description. The method resulted in an excellent performance in identifying the pre-tensioning and load-limiting features (F1-scores equal to 0.95 and 0.90, respectively), and their characteristics, although some limitations were observed, such as limited performance with dual-stage load-limiting devices (F1-score equal to 0.77).

The approach developed to identify restraint system configurations was constrained by the available data. Other components that play a critical role in effectively restraining the occupant (such as the airbag, seat, and knee bolster) could not be directly characterised based on current crash test measurements. Recorded videos from crash tests can be used to identify the presence of different airbags, their locations, and activation times and estimate their sizes using their relative ratio to the Anthropomorphic Test Device (ATD). However, these characteristics are insufficient to fully describe these complex devices or to develop a simplified model. Therefore, detailed reconstructions of crash tests could be employed as a tool to indirectly estimate the configurations of these devices, providing valuable information for future studies. For this task, a method similar to that employed by Iraeus et al. (2016) to develop the generic vehicle interior model could be followed but applied to unique vehicles.

The characteristics of the restraint system were identified under a specific set of controlled conditions, including a unique combination of delta-v, PDOF, and occupant type. This specificity may influence how the restraint system behaves under varying impact conditions or with different occupants.

Future studies could extend the methodologies developed in this dissertation to analyse the restraint system characteristics in a broader range of crash tests, such as side impact crash tests under the NCAP program and other crash test configurations conducted by the Insurance Institute for Highway Safety and the Highway Loss Data Institute (IIHS-HLDI). By identifying these features on a per-vehicle basis, it may be possible to detect shifts in the restraint system characteristics when crash conditions or occupant types differ from those studied in this dissertation. However, the identification algorithm's parameters may require adjustments to suit new crash scenarios. For instance, a lower delta-vs could shorten the plateau in the force-time history measurement, or in some cases, no force-limiting might be applied, making the identification of this feature more challenging.

Shifts in the identified restraint system characteristic may result from the implementation of occupant classification systems (OCS), which utilise additional sensors to modify the restraint system's behaviour based on the occupant's specific attributes. While a complete characterisation may not be extrapolated from their behaviour in crash tests, including more crash tests and occupant anthropometries could offer a partial understanding of the OCS's influence on restraint system performance. This approach could be particularly relevant for analysing the impact of OCS using real-world crash databases, which may include modern vehicles equipped with such systems, by augmenting these databases with the expected restraint system characteristics for given impact conditions.

5.1.2 Objective 2: To assess the effect of restraint systems on injury risk using real-world crash data

Understanding the contribution of various safety features in preventing injuries and fatalities is essential for advancing vehicle safety. Numerous studies have demonstrated the significant benefits of advanced restraint systems in controlled laboratory settings and computational models (Forman et al., 2009; 2008; Kent et al., 2001; 2007; Michaelson et al., 2008; Walz, 2004). However, the actual effectiveness of these systems can only be estimated when they are assessed using real-world injury data (Segui-Gomez et al., 2009). Recent research has increasingly shown that newer vehicles are associated with a lower risk of injury (Forman et al., 2019; Klinich et al., 2016; Ryb et al., 2011; Thomas, 2013). Early field studies suggested that features such as load limiters may have been effective in preventing severe thoracic trauma in certain real-world crash scenarios (Foret-Bruno et al., 1978). However, to the best of our knowledge, this study represents the first to use a large sample of real-world data to assess the impact of incorporating advanced restraint system features on the likelihood of injuries.

The NASS CDS (National Automotive Sampling System Crashworthiness Data System) database was augmented by using the vehicle identification number (VIN) from the vehicles involved in crashes to retrieve detailed information on the restraint system features. NASS CDS, which began in the late 1970s and concluded in 2015, provides a comprehensive dataset that includes vehicles both with and without the identified advanced restraint system features. In 2015, the Crash Investigation Sampling System (CISS) replaced NASS CDS. It provides, as it is currently actively including more cases, an improved database that better reflects the current vehicle population, incorporating newer models and more detailed injury coding. CISS also enhanced sampling methods to more accurately represent real-world crash scenarios. Future studies could explore the impact of advanced restraint systems by leveraging the CISS database or by combining data from both crash databases. However, the presence of more modern vehicles in the CISS data presents certain challenges, particularly due to the inclusion of OCS, which may affect the comparability of injury outcomes and restraint system performance across different crash conditions.

A multivariate logistic regression model was employed to evaluate the impact of restraint systems on AIS 2+, AIS 3+, and fatal injuries. During this analysis, the two variables representing the presence of pre-tensioning devices (whether pre-tensioning was observed at the shoulder or lap belt sections) were not found to be statistically significant in predicting the likelihood of injuries across most models. This lack of significance is likely attributed to the strong correlation between the presence of pre-tensioning and load-limiting devices in modern vehicles, potentially masking the individual effect of pre-tensioners on injury outcomes. Despite this, the variable related to load-limiting features was retained in the regression models, as it plays a critical role in mitigating injury by controlling the force exerted on occupants during a crash. Future analyses should consider not only the presence of pre-tensioning devices but also the specific level of pre-tensioning force applied to the occupant. This additional factor could offer further insights into the role these devices play in reducing injuries, providing a more nuanced understanding of how pre-tensioners interact with load-limiters to influence injury outcomes.

The analysis of fatal injuries revealed that the presence of a load-limiting device in the restraint system, regardless of the specific load-limiting force, significantly reduced the risk of death. This finding supports previous research by Kahane (2013a), 2015 based on data from the fatality analysis reporting system (FARS), reinforcing the effectiveness of load-limiting devices in reducing fatalities. Additionally, it was observed that devices with lower load-limiting

forces were associated with a reduced risk of AIS 3+ injuries across the entire body in frontal impacts. Furthermore, any level of load-limiting force was found to significantly reduce the risk of AIS 2+ injuries when combined with the delta-v. This suggests that load-limiting devices may have different effects depending on the delta-v. It is also important to note that advanced restraint systems are likely optimised for the impact conditions used in NCAP tests, which are conducted at 35 mph (56 km/h). In contrast, the average delta-v in the NASS CDS sample used for this study was substantially lower, at 23.3 ±12.9km/h, with around 90 % of crashes observed at delta-vs below 50 km/h. Consequently, it is plausible that the impact of advanced restraint systems might be more pronounced in high-speed collisions.

Airbag deployment was found as another restraint system feature with a significant impact, particularly when considered alongside delta-v. The airbag deployment, when considered alone, was found to significantly increase the risk of AIS 2+ and AIS 3+ injuries for the whole body and specific body regions. This initial finding suggests that airbags, while designed to mitigate injury, may introduce additional risks under certain conditions. However, when an interaction term between airbag deployment and delta-v was included in the models, statistically significant results were observed for all injury outcomes (AIS 2+, AIS 3+, and fatal injuries). This finding suggests that the influence of airbag deployment on injury risk is highly dependent on crash severity, as represented by delta-v. The odds ratios (OR) derived from the logistic regression analysis indicated that at delta-vs exceeding 31 km/h, airbag deployment had a positive effect in reducing injury risk for the thoracic region.

As mentioned above, newer vehicles have been found to be associated with a reduced risk of injury. For example, Forman et al. (2019) found that vehicles manufactured in 2009 and later were linked to a lower likelihood of AIS 2+ and AIS 3+ injuries compared to older models. However, despite the advancements in modern restraint systems, their analysis did not find significant reductions in AIS 3+ rib fractures. This aligns with the findings from the current study, where vehicle model year did not emerge as a significant factor when variables related to pre-tensioning and load-limiting devices were included in the multivariate models. Additionally, the model year produced high Variance Inflation Factor (VIF) scores during collinearity checks, indicating a strong correlation with the presence of advanced restraint systems.

One possible explanation for the findings could be the influence of other variables, such as occupant characteristics, which may outweigh the impact of restraint systems on injury risk. For instance, the increased vulnerability of older occupants might outweigh the protective benefits offered by advanced restraint systems in certain cases. The true effectiveness of these restraint systems in real-world crashes remains an important area for further investigation, as understanding the gap between their proven laboratory performance and real-world outcomes is essential for enhancing vehicle safety across diverse occupant populations. The computer modelling methods developed in Chapter 3 and 4 explore a complementing tool for assessing the impact of these features, enabling a more nuanced analysis of how restraint systems perform under varying real-world conditions.

5.1.3 Objective 3: To assess the effect of restraint systems on predicted injury outcomes based on computer modelling incorporating real-world crash conditions

As mentioned in the previous section, relying solely on observations from real-world crashes presents limitations. These limitations include a potential lack of statistical significance, resulting from factors such as small sample sizes or correlation effects between the variables

under study. To overcome these challenges, a method was developed using computer models to assess the impact of restraint system characteristics on injury outcomes through the generation of an ad-hoc population of crashes, allowing for a more controlled and effective assessment.

Chapters 3 and 4 focused on developing this method using two modelling techniques and two approaches to estimate injury outcomes: deterministic and probabilistic. While both approaches provided valuable insights, neither approach fully captured the injury distribution observed in real-world reference data.

The developed method employed a metamodelling approach using surrogate models to predict the outcomes of multibody (MB) and finite element (FE) simulations, thereby significantly reducing computational costs. This approach considerably decreased the time invested to perform the simulations, saving weeks to months in obtaining injury outcomes. Estimated savings calculated with the computational resources available during the study were approximately five months for the MB simulations and one year for the FE simulations. Although the computational resources used to perform the simulations for both models were different, the calculated computational cost using a single CPU per simulation was 125 times larger using the FE model than for the MB model.

Among the tested metamodels, Gaussian Process Regression (GPR) achieved the lowest prediction errors for most of the studied injury criteria. However, the metamodelling approach had limitations in accurately capturing the sharp peaks in the HIC, particularly when nonlinearities in the model led to large, isolated variations in occupant response.

Although the proposed method used these metamodels as a black box to minimise computational costs, the LASSO regression method can enable the understanding of the relationships between different inputs and outputs of the resulting function using the coefficients calculated by the method. However, this method is limited to representing linear and quadratic interactions between inputs. Alternative approaches could be explored to better capture the nonlinear relationships in the simulation model outcomes. These approaches can be based on understanding already trained neural network models, such as analysing the sensitivity of the different inputs (Pizarroso et al., 2022) or training a new set of explainable artificial intelligence (XAI) models (Angelov et al., 2021). Exploring such methods, though valuable, was beyond the scope of the current study.

The MB and FE models implemented in the study incorporated different levels of variability to estimate injury outcomes in real-world crashes. The injury criteria used to estimate the outcomes of both models differed, as these criteria had to be chosen based on the selected human body models (HBMs). The MB model included variability in impact conditions (such as delta-v and PDOF) and in restraint system characteristics (such as pre-tensioning and load-limiting characteristics). In contrast, the FE model also enabled the variability in occupant anthropometry and the inclusion of injury criteria that accounted for the occupant's age in the calculations. A mean statistical anthropometric model was used to morph the FE HBM and represent the variations of the occupants observed in the reference data, considering weight and height and based on Larsson et al. (2019). However, individual variations from this mean anthropometric model were not included.

Although not originally within the scope of this study, the restraint system characteristics identified in Chapter 2 were used to compare the responses of the FE Generic Vehicle Interior (GVI) model with crash tests that incorporated variations in restraint system features. The comparison revealed that the Anthropomorphic Test Device (ATD) model consistently over-predicted chest deflection, regardless of the restraint system configuration. This overestimation had been reported in previous studies as well (Boyle et al., 2024; Larsson et al., 2024),

even when different ATD models were utilised. These over-predictions may be attributed to different factors, such as the use of a combination of restraint system characteristics that were not optimised for the vehicle, a different ATD position relative to the vehicle interior, different belt routing, and different positions of the elements in the vehicle interior.

In addition to chest deflection, differences were observed in femur compression forces when benchmarking the model against reference data, particularly between vehicles with and without load-limiting features. These differences could be the result of modifications in knee bolster stiffness, restaining of the pelvis, and knee intrusion throughout the years, as lower femur forces were observed in those vehicles with load-limiting features. To further enhance the current GVI model, it would be valuable to expand the validation of the parametric GVI model developed by Boyle et al. (2024) by incorporating a larger number of reference crash tests. This would help to improve the model's correlation. Moreover, investigating which combinations of restraint system characteristics (such as knee bolster stiffness and airbag size) best correlate with specific features, like load-limiting devices in the reference vehicles, could provide further insights into optimising vehicle safety systems.

Regarding the prediction of injury outcomes, neither the deterministic nor the probabilistic approach fully captured the injury distribution reflected in real-world reference data. However, even if these methods may not provide a precise estimate of the total number of injuries under varying crash conditions, they may still offer insights into how different factors impact injury outcomes. Specifically, they can help understand how changes in crash conditions and restraint systems may influence injury outcomes and potentially guide improvements in injury reduction strategies. A detailed discussion of the findings from both methods is provided below, offering a comparative analysis of their performance and contributions to injury prediction models.

The deterministic approach, being the simplest method proposed for estimating injury outcomes, relied on setting a threshold value. If the metamodel predicted a value exceeding this threshold, an injurious outcome was estimated. Despite its simplicity, this approach introduced several limitations. One of the key issues was the sensitivity of the results to the chosen threshold. While the threshold could be adjusted based on the AIS severity of the predicted injury, in this study, the threshold values were derived from the lower limits set by the Euro NCAP assessment protocol Euro NCAP (2023a). These thresholds, along with those of similar magnitude, are frequently utilised in both governmental and non-governmental crash tests to assess occupant protection in regulatory and consumer rating programmes. However, the injury risks linked to these thresholds varied substantially across different injury risk functions, even for the same injury severity levels. For instance, the threshold for femur injuries resulted in a 7% to 13% injury risk for AIS 3+ and AIS 2+, respectively. Notably, no injurious outcomes were predicted in this case, even when overweighted occupants were considered with the FE model. In contrast, for head internal injuries and unconsciousness, the injury risk ranged from 50% to 65% for AIS 3+ and AIS 2+, respectively. However, an overestimation of the number of injurious outcomes was observed in this category. This highlights the difficulty of applying a deterministic approach to estimate injury outcomes and the potential lack of robustness when defining the thresholds. Other alternatives to define the threshold values may result in similar outcomes, such as defining a threshold based on an injury risk (e.g., 5 %, 25 % and 50 %), due to this lack of robustness.

The probabilistic method, though more complex, provided more injury patterns that were not captured by the deterministic approach. For example, those injury patterns that included a femur injury were captured by both the MB and FE models, as injuries in this body region were predicted by applying this method. Additionally, the probabilistic method offers the

flexibility to be fine-tuned by adjusting the injury risk functions, as suggested by Forman et al. (2022) and Hu et al. (2015). However, while these adjustments can improve the model's predictive capabilities, they may also mask some inherent limitations in the model's accuracy. Fine-tuning might lead to a better fit to specific datasets but could potentially obscure underlying issues, such as model oversimplifications or gaps in the representation of real-world injury mechanisms. Therefore, caution is needed when interpreting results from such adjustments, as they may be affected by limitations in the model.

It should be noted that both MB and FE models failed to predict correctly the real-world injury outcomes. The vehicle interior model plays a critical role in influencing these outcomes. For femur injury predictions, as previously mentioned, adjustments to the knee bolster stiffness may be necessary to improve correlation with real-world data. The current stiffness settings may not accurately reflect the conditions that lead to femur injuries, suggesting the need for further refinement. When it comes to head injuries, the injury criteria relied on head angular velocities, a parameter not typically incorporated into the validation processes of either HBMs or vehicle models. This reliance on angular velocity presents a challenge, as current validation efforts, particularly those involving vehicle interiors, are limited by the availability of suitable data. Most validation tests have been conducted using the Hybrid III 50^{th} percentile male dummy, which generally lacks an angular velocity sensor.

Thoracic injuries were the most frequently observed injuries in the study. Two different injury criteria were applied to the HBMs, each with varying levels of complexity. The MB HBM, using chest deflection as an injury criterion, applied a deterministic approach that resulted in a narrow set of conditions for predicting injurious or non-injurious outcomes. Specifically, this method predicted that almost all occupants would be injured if the delta-v exceeded 39 km/h, while nearly none would be injured if the velocity was below 36 km/h. This sharp threshold does not reflect the broader range where injuries are observed in real-world crash conditions. In contrast, the probabilistic approach applied to the MB model produced a more realistic range of outcomes, closely matching real-world observations where both injured and non-injured occupants were found across a wider spectrum of crash conditions.

When focusing on thoracic injuries for the FE HBM, both the deterministic and probabilistic approaches yielded similar results. In this case, a strain-based injury criterion was applied, and the deterministic approach used an expected outcome that aligned with the injury definitions employed in the study. Although the deterministic and probabilistic methods tended to over-predict the number of injured occupants, the inclusion of a broader set of variations using the FE model allowed both methods to capture a wide range of conditions where both injured and non-injured occupants were observed. This broader range of outcomes is particularly valuable, as it enables a more detailed examination of which variables contribute most significantly to these injury outcomes.

Finally, and as the overall aim of this dissertation was to develop a method for assessing the impact of restraint systems on injury outcomes, the ad-hoc population of crashes generated using modelling was used to assess the effect of load-limiting devices on injury outcomes. Both the reference (real-world) data and the population of crashes generated via modelling were analysed to determine whether the load-limiting force applied to the occupant had a statistically significant impact on injury outcomes, or if variations in force levels did not significantly influence these outcomes. An unequal variance t-test was conducted to compare injured and non-injured occupants in both the reference data and the model's estimations, focusing specifically on the load-limiting forces applied to the occupant.

Given the significant interaction between the load-limiter feature and delta-v in the multivariate logistic regression analysis from Chapter 2, two different delta-v ranges were established to assess the impact of the load-limiting feature. In addition to delta-v, several other crash conditions were analysed as follows:

- Cases with a delta-v range:
  - 20 to 30 km/h
  - 45 to 55 km/h
- Cases with a PDOF from -10º to 10º
- Cases in which the airbag was deployed
- Cases with a load-limiting device of a single-stage

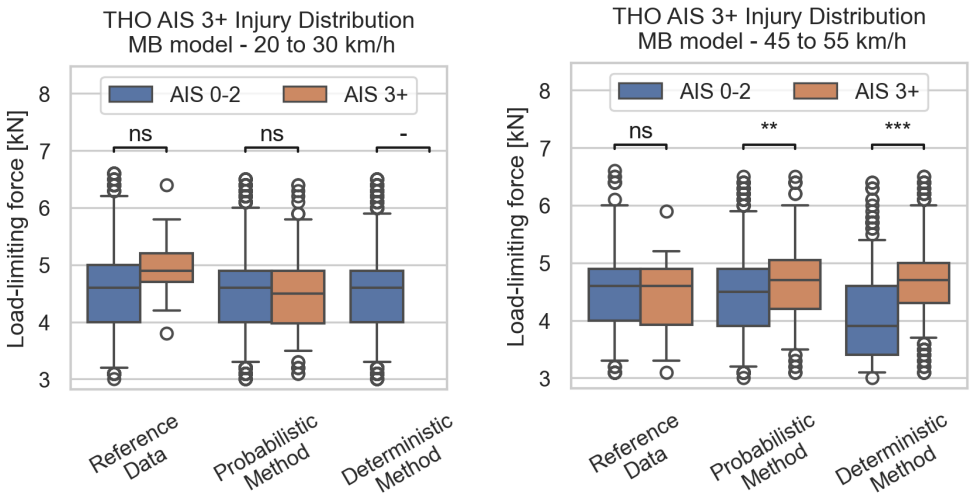
Figure 5.1 shows the results observed for the reference data and the MB model applying both the deterministic and probabilistic methods with regard to THO AIS 3+ injuries. Subfigure 5.1b and 5.1a display the results for the 20 to 30 km/h and 45 to 55 km/h delta-v ranges, respectively. In the analysis, five levels of statistical significance were employed: "ns" indicating not significant; "#" representing marginally significant; "*" indicating $p < 0.05$; "" representing $p < 0.01$; and "" indicating $p < 0.001$. When the evaluation was not possible due to a lack of data, this was represented with a hyphen ("-"). These significance levels were used to evaluate the impact of the load-limiting forces on injury outcomes across the different crash conditions.

This analysis revealed no significant variations in injury outcomes within the 20 to 30 km/h delta-v range. As previously noted, the deterministic method did not predict any injuries in this range. While the probabilistic method estimated some injuries, there was no significant difference between injured and non-injured occupants based on varying load-limiting forces. Similarly, no significant difference was observed in the reference data, although AIS 3+ injuries were primarily observed with higher load-limiting forces. However, when focusing on the 45 to 55 km/h delta-v range, significant differences were observed in the estimations only from the deterministic and probabilistic methods. In this higher delta-v range, lower load-limiting forces were significantly associated with a lower number of estimated injuries.

Figure 5.2 shows the results observed for the reference data and the FE model applying both the deterministic and probabilistic methods with regard to the number of fractured ribs (NFR). Subfigure 5.2b and 5.2a display the results for the 20 to 30 km/h and 45 to 55 km/h delta-v ranges, respectively. The same statistical significance levels applied in Figure 5.1 were used.

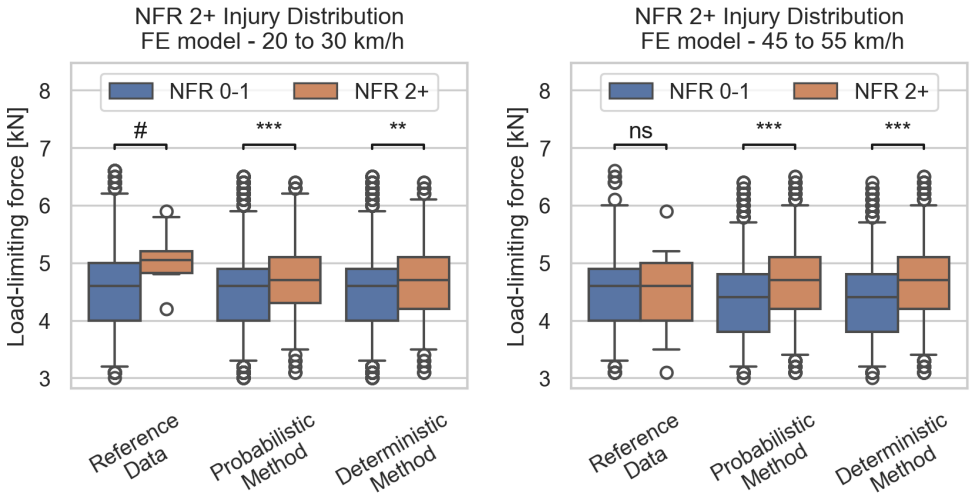
Focusing on NFR 2+, both the observations from the reference data and the estimations from the computer model showed significant results, with lower load-limiting forces leading to a reduced risk of NFR 2+ injuries. In the reference data, this relationship was marginally significant (p-value=0.056), which may related to the size of the filtered dataset (1060 non-injurious vs ten injurious outcomes). Similar to the findings for THO AIS 3+ injuries, NFR 2+ injuries were more frequently associated with larger load-limiting forces in the reference data. However, this marginal significance in the reference data was not observed in the higher delta-v range. In contrast, the computer model estimations demonstrated significant differences across both delta-v ranges, although the degree of significance varied depending on the approach (deterministic or probabilistic) used. The observed trend of decreasing load-limiting forces applied by vehicle manufacturers, as observed in Chapter 2, may be connected to these findings.

The comparison between the two models (MB and FE) and the reference data was partially limited due to differences in the selection of injuries used in the analysis. However,



(a) Range 20 to 30 km/h. (b) Range 45 to 55 km/h.

Figure 5.1: Load-limiting force significant analysis observed for the reference data and the MB model applying both the deterministic and probabilistic methods with regard to THO AIS 3+ injuries. Statistical significance were employed: "ns" indicating not significant; "#" representing marginally significant; "*" indicating $p < 0.05$; "**" representing $p < 0.01$; and "***" indicating $p < 0.001$.



(a) Range 20 to 30 km/h. (b) Range 45 to 55 km/h.

Figure 5.2: Load-limiting force significant analysis observed for the reference data and the FE model applying both the deterministic and probabilistic methods with regard to NFR 2+ injuries. Statistical significance were employed: "ns" indicating not significant; "#" representing marginally significant; "*" indicating $p < 0.05$; "**" representing $p < 0.01$; and "***" indicating $p < 0.001$.

several important trends were observed from these results. Although the MB model focused on higher severity injuries (AIS 3+ vs. NFR 2+), it did not capture significant effects when modifying the load-limiting force at lower delta-v. Significant effects were only detected at higher delta-vs, which fall in a similar range to those used in current crash tests. This may be attributed to the MB HBM being validated mainly at those higher delta-vs, leading to a narrower sensitivity to changes in load-limiting force at lower delta-vs. In contrast, the use of an FE HBM provided more detailed insights into the impact of load-limiting force. Significant effects were observed across both low and high delta-vs, suggesting that the FE model has improved capabilities to capture the variability of injury outcomes and the influence of load-limiting features across a broader spectrum of crash conditions. This highlights the potential of the FE HBM for evaluating restraint system effectiveness. The FE HBM analysis was not limited to the 50^{th} percentile male but included a broader range of occupant anthropometries, further enhancing its potential for evaluating restraint system effectiveness in real-world conditions. Furthermore, the results from the FE model included elderly occupants, who have been observed to be at a larger risk than younger occupants. Future studies could focus on specific subsets of occupants to study whether the effects of varying load-limiting forces are consistent across the entire population or if distinct trends emerge for different demographic groups, which could be applied to develop new adaptative restraint systems that include occupant variability.

The accuracy of the developed method's predictions could be enhanced through the continued refinement of injury criteria. As previously noted, nearly 90 % of crashes in the dataset occurred at delta-vs below 50 km/h, with approximately 60 % involving delta-vs under 30 km/h. The lack of validation of these model predictions, particularly at these lower velocities, can lead to unrealistic outcomes, including both over- and under-estimations of injury severity, such as in the case of chest compression using the MB HBM. Refining injury risk predictions for HBMs at these lower speeds, as recommended by Östling et al. (2024), could result in more accurate estimates of the injuries that may be prevented by specific restraint system configurations. Currently, the method tends to overestimate injury outcomes, which would benefit from improved injury risk estimations, particularly in lower-speed collisions.

While this study focused on the effects of restraint systems on injury outcomes in crashes, occupant safety may also be influenced by other safety features, such as automatic emergency braking systems (AEBS). Future research could examine how restraint system effectiveness varies across different pre-crash scenarios, as those explored by Leledakis et al. (2021a). Given that pre-crash phases may alter occupant position, changes in the occupant position should be included as this may affect the HBM response (Östmann et al., 2016).

The metamodel development with the FE computer model utilised a predefined set of occupant anthropometries, following the recommendations by Larsson et al. (2024). This study investigated the number of anthropometries required to represent a population crash outcome. Future research could study similar methods to determine the number of crash configurations needed to represent a crash population, potentially bypassing the need for a metamodel, and enabling alternative restraint systems to be compared with the existing metamodel with lower compuational cost. Furthermore, this reduced number of crash conditions may help to define relevant crash scenarios to be used in virtual testing as consumer rating programs start to implement this complementary assessment to physical crash tests (Klug et al., 2023).

5.2 Conclusion

The aim of this PhD dissertation was to develop a method for assessing the impact of restraint systems on injury outcomes considering the variability of real-world crash conditions. By leveraging real-world data and advanced computer models, the research sought to enhance the development of restraint systems, ultimately reducing the incidence and severity of traffic-related injuries.

Three objectives were set to fulfil the aim of this dissertation: identifying the presence and characteristics of the restraint system features, combining real-world crash data with the restraint system information to estimate their impact on injury risk, and combining a vehicle interior sled model, a human model and real-world crash conditions to assess the impact of restraint systems on predicted injury outcomes. The combination of real-world crash data and estimations based on computer models aimed to address the limitations inherent in crash databases and provide a more comprehensive understanding of how restraint systems influence injury outcomes, ultimately contributing to improved safety measures.

This dissertation makes several important contributions to the field of automotive safety. First, it introduces a novel tool that identifies the presence and characteristics of pre-tensioner and load-limiting devices in vehicle restraint systems using force-time history measurements. The results of applying this tool to 1,318 vehicles were made publicly available via a GitHub repository, allowing the research community to validate vehicle models and integrate this data with crash test databases. Second, to the best of the author's knowledge, this is the first study to augment NASS CDS crash data with detailed information about pre-tensioners and load limiters, enabling a more in-depth investigation into their impact on injury risk. Lastly, a computer modelling-based method was developed to estimate the impact of restraint systems on injury outcomes. Two approaches within this method were benchmarked against real-world crashes using both multibody (MB) and finite element (FE) models, providing a framework for evaluating the effectiveness of restraint systems under real-world crash conditions. The method's ability to generate ad-hoc populations of crashes offers valuable insights into how restraint systems perform across diverse crash scenarios. The findings from this research not only identify specific restraint system features associated with reduced injury risk but also provide practical guidance for the design of future vehicle safety systems.

This dissertation has provided valuable insights into the assessment of vehicle restraint systems. The following points highlight the major findings, acknowledge the study's limitations, and suggest areas for future exploration:

Objective 1: To develop a method for identifying the presence and characteristics of restraint system features.

- Developed a novel tool to identify pre-tensioning and load-limiting features in vehicle restraint systems using force-time history measurements.
- Identified and shared detailed restraint system features and configurations for 1,318 vehicles via a GitHub repository, enabling community access for validation and integration with crash test databases.
- Performance evaluation of the tool in identifying the presence of pre-tensioner and load-limiting devices:
  - Pre-tensioning devices: Excellent performance (F1 score: 0.95).
  - Single-stage load-limiting devices: High accuracy (F1 score: 0.90).
  - Dual-stage load-limiting devices: Moderate performance (F1 score: 0.77).

- Historical trends over the past 40 years in the adoption and characteristics of restraint systems:
  - Adoption of advanced restraint systems increased alongside the introduction of more comprehensive crash tests and user-friendly crash safety ratings.
  - Pre-tensioning and load-limiting devices were widely adopted starting in the mid-1990s, with most vehicles incorporating these technologies by 2005.
  - Average pre-tensioner forces were measured at 2 kN for shoulder belts and 2.5 kN for lap belts.
  - Over the past two decades, the average load-limiting forces have progressively decreased to around 3 kN for front-seated passengers and 4 kN for drivers.
  - Load-limiting forces for front-seat passengers began to differ from those applied to the driver with the introduction of the Hybrid III 5th percentile female ATD as the passenger surrogate after 2010.
- Future studies could study the behaviour of the identified restraint systems under different crash test conditions, as they were studied in a specific scenario, and interactions with occupants may vary.

Objective 2: To assess the effect of restraint systems on injury risk using real-world crash data.

- This study is the first to augment NASS CDS crash data with detailed vehicle information about pre-tensioners and load limiters to investigate their impact on injury risk.
- Multivariate logistic regression models were developed to assess the real-world impact of advanced seat belt systems on moderate, severe, and fatal injuries, both for the whole body and specific body regions.
- Significant associations between restraint system features and injury outcomes identified in this dataset include:
  - The presence of pre-tensioning and load-limiting devices was significantly associated with a reduction in fatal injuries, consistent with previous studies. The reduction changed depending on the force level of the load limiter (OR = 0.31 for low-force, OR = 0.42 for high-force devices).
  - Low-force load limiters (<4.5 kN) were associated with a reduced risk of AIS 3+ whole-body injuries (OR = 0.70).
  - The presence of load-limiting devices was linked to lower AIS 2+ injury risks in the whole body in high delta-v crashes.
  - No significant associations were found between injury risk in specific body regions and the presence of pre-tensioning and load-limiting devices within this dataset.
  - Airbag deployment at higher crash velocities was significantly associated with a reduction in AIS 2+ and AIS 3+ injuries, both across the whole body and in specific regions.
- This study highlights the need for continuing the investigation of the real-world effectiveness of advanced seat belt systems, being able to increase the sample size. This study suggested that some non-seatbelt factors, such as age and impact direction, may be key to assessing the injury outcome in real-world conditions.

- Future research could leverage other existing databases to examine the effects of restraint systems on injury outcomes. Considerations should be taken to minimise the impact of OCS, which may be more present in modern vehicles.

Objective 3: To assess the effect of restraint systems on predicted injury outcomes based on computer modelling incorporating real-world crash conditions.

- A computer modelling-based method was developed to estimate injury outcomes by accounting for variability in real-world crash conditions, and it was used to estimate the impact of different restraint system configurations.
- Simulations and metamodelling approach:
  - A total of 1,200 MB and 1,400 FE simulations were performed to generate metamodels, incorporating variations in impact conditions, restraint systems, and occupant characteristics.
  - The metamodelling approach significantly reduced computational costs by estimating the model outcome instead of running a simulation. Gaussian Process Regression (GPR) showed the lowest prediction errors for most injury criteria, although all regression methods struggled to capture sharp nonlinear variations, particularly with the HIC criterion.
- The modified FE Generic Vehicle Interior (GVI) model was validated against crash test data, including specific configurations of load-limiting devices. The validation revealed an overestimation of chest deflection in the ATD model and discrepancies in femur compression forces, particularly for certain combinations of load-limiting characteristics. These discrepancies were linked to the characterisation of the knee bolster, indicating the need for further refinement of the model. Future validation efforts should incorporate more precise restraint system characteristics, when available, to improve model accuracy and predictive capabilities.
- Regarding the application of the deterministic and probabilistic approaches for estimating injury outcomes, the main conclusions from this dissertation are as follows:
  - Neither the deterministic nor probabilistic approaches fully captured real-world injury distributions. However, both provided valuable insights into how changes in crash conditions can influence injury outcomes, aiding the development of injury reduction strategies.
  - The deterministic method, while simple, introduced challenges due to its reliance on threshold values. This was especially observed in its sensitivity to the selection of these thresholds.
  - The probabilistic approach offered greater flexibility, successfully identifying injury patterns missed by the deterministic method. However, fine-tuning the model to improve predictions by scaling injury risk functions may risk overfitting specific datasets and hiding underlying limitations.
  - The deterministic method may produce unrealistic predictions (e.g., predicting nearly all occupants would be injured above 39 km/h using the MB model). In contrast, the probabilistic method more accurately reflected real-world thoracic injury risks across a wider range of crash conditions.
  - Both deterministic and probabilistic approaches provided similar predictions when using an expected outcome as the injury criterion for the deter-

ministic method (e.g., the number of broken ribs). However, this approach may not be applicable to predicting injuries in other anatomical structures.
  - Significant differences in injury outcomes were observed between the MB and FE models. The morphed FE HBM showed greater potential in capturing the impact of restraint systems on injury outcomes, especially across varied crash conditions.
- Lower load-limiting forces were associated with a significant reduction in injury risk for rib fractures at small and large delta-vs according to the estimations of the developed method.
- Significant findings on the effects of load-limiting forces across different crash velocities underscore the importance of fine-tuning these systems. Future research should explore the impact of varying load-limiting forces on different demographic groups to support the development of adaptive restraint systems tailored to occupant variability.
- The proposed method demonstrated the ability to estimate the impact of changes in restraint system configurations on injury outcomes. While the analysis of significant differences is valid, the precise prediction of the number of injuries that could be avoided per 1,000 crashes is not yet feasible.
- Improvements in injury criteria predictions, especially at low-severity crashes, may result in more accurate estimations of the number of avoided injuries.

This dissertation focused on a critical aspect of improving public health by reducing traffic-related injuries and fatalities. Specifically, it developed a tool to assess the effectiveness of vehicle restraint systems in mitigating injury outcomes, accounting for the variability present in real-world crash scenarios. These restraint systems, integral to passive safety mechanisms, are designed to protect occupants during collisions by minimising the occurrence or the severity of injury once the crash is inevitable. However, from a broader perspective, addressing the safety of vulnerable road users and non-car occupants is also a priority worldwide. Advanced computational methods such as the ones implemented in this dissertation could be used to provide insight into how to improve the protection of these other road users.

Bibliography

AAAM (1998). "The abbreviated injury severity scale, AIS, 1990 update 1998". In: The Association for the Advancement of Automotive Medicine, Des Plaine, IL.

– (2008). "The abbreviated injury scale 2005, update 2008". In: The Association for the Advancement of Automotive Medicine, Barrington, IL.

Abrams, M. and C. R. Bass (2020). "Female vs. male relative fatality risk in fatal crashes". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*.

– (2022). "Female vs. male relative risk of body system injuries in fatal and non-fatal crashes". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*.

Aldman, B. (1962). "Biodynamic studies on impact protection". *Acta physiologica Scandinavica. Supplementum* 56:192. Number: 192, pp. 1–80.

Alejo, R., V. García, J. M. Sotoca, R. A. Mollineda, and J. S. Sánchez (2007). "Improving the performance of the RBF neural networks trained with imbalanced samples". In: *Computational and ambient intelligence: 9th international work-conference on artificial neural networks, IWANN 2007, san sebastián, spain, june 20-22, 2007. Proceedings 9*. Springer, pp. 162–169.

Amorim, L. B. V. de, G. D. C. Cavalcanti, and R. M. O. Cruz (1, 2023). "The choice of scaling technique matters for classification performance". *Applied Soft Computing* 133, p. 109924. ISSN: 1568-4946. DOI: 10.1016/j.asoc.2022.109924.

Angelov, P. P., E. A. Soares, R. Jiang, N. I. Arnold, and P. M. Atkinson (2021). "Explainable artificial intelligence: an analytical review". *WIREs Data Mining and Knowledge Discovery* 11:5, e1424. ISSN: 1942-4795. DOI: 10.1002/widm.1424.

Anja, R., L. Markus, and F. Franz (2011). "Method to estimate the field effectiveness of an automatic braking system in combination with an adaptive restraint system in frontal crashes". In: *Proceedings of 22nd international technical conference on the enhanced safety of vehicles*, pp. 11–0281.

Bance, I., S. Yang, Q. Zhou, S. Li, and B. Nie (2021). "A framework for rapid on-board deterministic estimation of occupant injury risk in motor vehicle crashes with quantitative uncertainty evaluation". *Science China Technological Sciences* 64:3. Number: 3 Publisher: Springer, pp. 521–534.

Beeman, S. M., A. R. Kemper, M. L. Madigan, C. T. Franck, and S. C. Loftus (2012). "Occupant kinematics in low-speed frontal sled tests: Human volunteers, Hybrid III ATD, and PMHS". *Accident Analysis & Prevention* 47, pp. 128–139. ISSN: 0001-4575. DOI: https://doi.org/10.1016/j.aap.2012.01.016.

Beillas, P., C. Giordano, V. Alvarez, X. Li, X. Ying, M.-C. Chevalier, S. Kirscht, and S. Kleiven (2016). "Development and performance of the PIPER scalable child human body models". In: *14th International conference on the protection of children in cars*, 19–p.

Bose, D., M. Segui-Gomez ScD, and J. R. Crandall (2011). "Vulnerability of female drivers involved in motor vehicle crashes: an analysis of US population at risk". *American jour-*

nal of public health 101:12. Number: 12 Publisher: American Public Health Association, pp. 2368–2373.

Boyle, K., A. Fanta, M. P. Reed, K. Fischer, A. Smith, A. Adler, and J. Hu (2020). "Restraint systems considering occupant diversity and pre-crash posture". *Traffic injury prevention* 21, sup1. Publisher: Taylor & Francis, S31–S36.

Boyle, K., J. Hu, M. Reed, and J. Rupp (2024). "Development of Parametric Finite Element Vehicle and Restraint Models for Occupant Protection Research". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*. Stockholm, Sweden.

Bratley, P. and B. L. Fox (1988). "Algorithm 659: Implementing Sobol's quasirandom sequence generator". *ACM Transactions on Mathematical Software (TOMS)* 14:1. Publisher: ACM New York, NY, USA, pp. 88–100.

Brumbelow, M. (2023). "Sex-related vehicle and crash differences and their potential to confound relative injury risk analyses". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*.

Brumbelow, M. L. and J. S. Jermakian (2022). "Injury risks and crashworthiness benefits for females and males: Which differences are physiological?" *Traffic injury prevention* 23:1. Publisher: Taylor & Francis, pp. 11–16.

Carter, P. M., C. A. C. Flannagan, M. P. Reed, R. M. Cunningham, and J. D. Rupp (1, 2014). "Comparing the effects of age, BMI and gender on severe injury (AIS 3+) in motor-vehicle crashes". *Accident Analysis & Prevention* 72, pp. 146–160. ISSN: 0001-4575. DOI: 10.1016/j.aap.2014.05.024.

Corrales, M. A., J. H. Bolte, B. Pipkorn, C. Markusic, and D. S. Cronin (11, 2024). "Explaining and predicting the increased thorax injury in aged females: age and subject-specific thorax geometry coupled with improved bone constitutive models and age-specific material properties evaluated in side impact conditions". *Frontiers in Public Health* 12. Publisher: Frontiers. ISSN: 2296-2565. DOI: 10.3389/fpubh.2024.1336518.

Crandall, J. R., D. Bose, J. Forman, C. D. Untaroiu, C. Arregui-Dalmases, C. G. Shaw, and J. Kerrigan (2011). "Human surrogates for injury biomechanics research". *Clinical anatomy* 24:3. Publisher: Wiley Online Library, pp. 362–371.

Crandall, J. R., R. Kent, J. Patrie, J. Fertile, and P. Martin (2000). "Rib fracture patterns and radiologic detection–a restraint-based comparison". In: *Annual proceedings/association for the advancement of automotive medicine*. Vol. 44. Association for the Advancement of Automotive Medicine, p. 235.

Dahlgren, M., A. Vishwanatha, A. Soni, K. Engstrand, J. Forsberg, and I. Yeh (2020). "Belt modelling in LS-DYNA®". In: *16th International LS-DYNA@ Users Conference*. Virtual Event.

Dalal, K., Z. Lin, M. Gifford, and L. Svanström (2013). "Economics of global burden of road traffic injuries and their relationship with health system variables". *International journal of preventive medicine* 4:12. Publisher: Wolters Kluwer–Medknow Publications, p. 1442.

Enayatollah, H. R., N. Khodadady-Hasankiadeh, L. Kouchakinejad-Eramsadati, F. Javadi, Z. Haghdoost, M. Hosseinpour, M. Tavakoli, A. Davoudi-Kiakalayeh, Z. Mohtasham-Amiri, and S. Yousefzadeh-Chabok (2020). "The relationship between weight indices and injuries and mortalities caused by the motor vehicle accidents: a systematic review and meta-analysis". *Journal of Injury and Violence Research* 12:1, pp. 85–101. ISSN: 2008-2053. DOI: 10.5249/jivr.v12i1.1198.

Eppinger, R., F. Bandak, M. Haffner, N. Khaewpong, S. Kuppa, M. Maltese, T. Nguyen, R. Saul, E. Sun, E. Takhounts, et al. (1999). "Development of improved injury criteria

for the assessment of advanced automotive restraint systems: II". Publisher: United States. National Highway Traffic Safety Administration.

Euro NCAP (12, 2023a). *ASSESSMENT PROTOCOL – ADULT OCCUPANT PROTECTION - Implementation 2024 -v9.3*. EUROPEAN NEW CAR ASSESSMENT PROGRAMME.

– (21, 2023b). *Technical Bulletin – Technical Bulletin Brain Injury Calculation - TB 035 - v1.0.1*. Euro NCAP.

– (2021). "Assessment Protocol: Adult Occupant Protection;" *European New Car Assessment Programme* Version 9.1.3. URL: https://cdn.euroncap.com/media/67264/euro-ncap-assessment-protocol-aop-v913.pdf (visited on 01/01/2022).

European Commission (19, 2019). *COMMISSION STAFF WORKING DOCUMENT: EU Road Safety Policy Framework 2021-2030 - Next steps towards "Vision Zero"*. SWD(2019) 283 final. Brussels, Belgium.

Fahlstedt, M., S. Meng, and S. Kleiven (2022). "Influence of strain post-processing on brain injury prediction". *Journal of biomechanics* 132. Publisher: Elsevier, p. 110940.

Fahse, N., M. Millard, F. Kempter, S. Maier, M. Roller, and J. Fehr (2023). "Dynamic human body models in vehicle safety: An overview". *GAMM-Mitteilungen* 46:2. Publisher: Wiley Online Library, e202300007.

FE_models/Open Access Front Seat Model · GitLab (17, 2024). GitLab. URL: https://openvt.eu/fem/open-access-front-seat-model (visited on 07/22/2024).

Foret-Bruno, J.-Y., F. Hartemann, C. Thomas, A. Fayon, C. Tarriere, C. Got, and A. Patel (1978). "Correlation between thoracic lesions and force values measured at the shoulder of 92 belted occupants involved in real accidents". *SAE Transactions*. Publisher: JSTOR, pp. 3206–3216.

Forman, J., R. W. Kent, K. Mroz, B. Pipkorn, O. Bostrom, and M. Segui-Gomez (2012). "Predicting rib fracture risk with whole-body finite element models: development and preliminary evaluation of a probabilistic analytical framework". In: *Association for the Advancement of Automotive Medicine*. Vol. 56, p. 109.

Forman, J., S. Kulkarni, D. Perez-Rapela, S. Mukherjee, M. Panzer, and J. Hallman (2022). "A Method for Thoracic Injury Risk Function Development for Human Body Models". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*. Porto, Portugal.

Forman, J., F. Lopez-Valdes, D. Lessley, M. Kindig, R. Kent, S. Ridella, and O. Bostrom (2009). "Rear seat occupant safety: an investigation of a progressive force-limiting, pretensioning 3-point belt system using adult PMHS in frontal sled tests". *Stapp car crash journal* 53. Publisher: The Stapp Association, p. 49.

Forman, J., J. Michaelson, R. Kent, S. Kuppa, and O. Bostrom (2008). "Occupant restraint in the rear seat: ATD responses to standard and pre-tensioning, force-limiting belt restraints". In: *Association for the Advancement of Automotive Medicine*. Vol. 52, p. 141.

Forman, J., G. S. Poplin, C. G. Shaw, T. L. McMurry, K. Schmidt, J. Ash, and C. Sunnevang (2019). "Automobile injury trends in the contemporary fleet: Belted occupants in frontal collisions". *Traffic injury prevention* 20:6. Number: 6 Publisher: Taylor & Francis, pp. 607–612.

Foster, J. K., J. O. Kortge, and M. J. Wolanin (1977). "Hybrid III—a biomechanically-based crash test dummy". *SAE transactions*. Publisher: JSTOR, pp. 3268–3283.

Gabler, L. F., J. R. Crandall, and M. B. Panzer (2019). "Development of a second-order system for rapid estimation of maximum brain strain". *Annals of biomedical engineering* 47. Publisher: Springer, pp. 1971–1981.

Gayzik, F., D. Moreno, C. Geer, S. Wuertzer, R. Martin, and J. Stitzel (2011). "Development of a full body CAD dataset for computational modeling: a multi-modality approach". *Annals of biomedical engineering* 39. Publisher: Springer, pp. 2568–2583.

Gehre, C., H. Gades, and P. Wernicke (2009). "Objective rating of signals using test and simulation responses". In: *International Technical Conference on the Enhanced Safety of Vehicles*. International Technical Conference on the Enhanced Safety of Vehicles. Vol. 2009. National Highway Traffic Safety Administration, Stuttgart, Germany.

Gehre, C. and S. Stahlschmidt (2011). "Assessment of dummy models by using objective rating methods". In: *22nd International Technical Conference on the Enhanced Safety of Vehicles Conference (ESV)*. 22nd International Technical Conference on the Enhanced Safety of Vehicles Conference (ESV). Washington, D.C., USA.

Håland, Y. (2006). "The evolution of the three point seat belt from yesterday to tomorrow". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*. Madrid, Spain.

Hensman, P. and D. Masko (2015). "The impact of imbalanced training data for convolutional neural networks". *Degree Project in Computer Science, KTH Royal Institute of Technology*.

Hershman, L. (2001). "The US New car assessment program (NCAP): past, present and future". *Proceedings: International Technical Conference on the Enhanced Safety of Vehicles* 2001, 13 p.–13 p. URL: http://dx.doi.org/.

Hertz, E. (1993). "A note on the head injury criterion (HIC) as a predictor of the risk of skull fracture". In: *Proceedings: Association for the Advancement of automotive medicine annual conference*. Vol. 37. Association for the Advancement of Automotive Medicine, pp. 303–312.

Higuchi, K., K. B. Arbogast, and R. W. Kent (2019). "Behavior of ATD, PMHS and human volunteer in frontal crash test". *International journal of automotive engineering* 10:4. Publisher: Society of Automotive Engineers of Japan, INC, pp. 348–355.

Hobbs, C. A. and P. J. McDonough (1998). "Development of the European new car assessment programme (Euro NCAP)". *Regulation* 44:3, pp. 2439–2453.

Horsch, J. D., J. W. Melvin, D. C. Viano, and H. J. Mertz (1, 1991). "Thoracic Injury Assessment of Belt Restraint Systems Based on Hybrid III Chest Compression". In: *Stapp Car Crash Conference*. ISSN: 0148-7191, 2688-3627. SAE International, Warrendale, PA. DOI: 10.4271/912895.

Hu, J., A. Fanta, M. Neal, M. Reed, and J. Wang (2016). "Vehicle crash simulations with morphed GHBMC human models of different stature, BMI, and age". In: *4th International Digital Human Modeling Conference*. 4th International Digital Human Modeling Conference. Montréal, Canada.

Hu, J., C. A. Flannagan, S. Bao, R. W. McCoy, K. M. Siasoco, and S. Barbat (2015). "Integration of active and passive safety technologies-a method to study and estimate field capability". *Stapp Car Crash Journal* 59, pp. 269–296.

Hu, J., K. Zhang, A. Fanta, M. Jones, M. Reed, M. Neal, J.-T. Wang, C.-H. Lin, and L. Cao (2017). "Stature and body shape effects on driver injury risks in frontal crashes: a parametric human modelling study". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*. Antwerp, Belgium, pp. 13–15.

Hu, J., K. Zhang, M. P. Reed, J.-T. Wang, M. Neal, and C.-H. Lin (2019). "Frontal crash simulations using parametric human models representing a diverse population". *Traffic injury prevention* 20, sup1. Number: sup1 Publisher: Taylor & Francis, S97–S105.

Huang, Y., Q. Zhou, X. Zhang, and C. Wang (2015). "Optimisation study of occupant restraint system concerning variations in occupant size and crash severity in frontal collisions". *International journal of vehicle safety* 8:4. Publisher: Inderscience Publishers (IEL), pp. 299–313.

Hwang, E., J. Hallman, K. Klein, J. Rupp, M. Reed, J. Hu, E. Hwang, J. Hallman, K. Klein, J. Rupp, M. Reed, and J. Hu (5, 2016a). "Rapid Development of Diverse Human Body Models for Crash Simulations through Mesh Morphing". In: SAE 2016 World Congress and Exhibition. ISSN: 0148-7191, 2688-3627, 0148-7191, 2688-3627. SAE International. DOI: 10.4271/2016-01-1491.

Hwang, E., J. Hu, C. Chen, K. F. Klein, C. S. Miller, M. P. Reed, J. D. Rupp, and J. J. Hallman (2016b). "Development, evaluation, and sensitivity analysis of parametric finite element whole-body human models in side impacts". *Stapp car crash journal* 60. Publisher: The Stapp Association, pp. 473–508.

Hwang, E., J. Hu, C. Chen, K. F. Klein, C. S. Miller, M. P. Reed, J. D. Rupp, J. J. Hallman, E. Hwang, J. Hu, C. Chen, K. F. Klein, C. S. Miller, M. P. Reed, J. D. Rupp, and J. J. Hallman (7, 2016c). "Development, Evaluation, and Sensitivity Analysis of Parametric Finite Element Whole-Body Human Models in Side Impacts". In: 60TH Stapp Car Crash Conference. The Stapp Association. DOI: 10.4271/2016-22-0014.

Hwang, E., J. Hu, and M. P. Reed (2020). "Validating diverse human body models against side impact tests with post-mortem human subjects". *Journal of Biomechanics* 98, p. 109444. ISSN: 0021-9290. DOI: https://doi.org/10.1016/j.jbiomech.2019.109444.

International Organization for Standardization (18, 2024). *ISO/TS 18571:2014*. ISO. URL: https://www.iso.org/standard/62937.html (visited on 03/18/2024).

Iraeus, J. and M. Lindquist (2015). "Pulse shape analysis and data reduction of real-life frontal crashes with modern passenger cars". *International journal of crashworthiness* 20:6. Number: 6 Publisher: Taylor & Francis, pp. 535–546.

– (2016). "Development and validation of a generic finite element vehicle buck model for the analysis of driver rib fractures in real life nearside oblique frontal crashes". *Accident Analysis & Prevention* 95. Publisher: Elsevier, pp. 42–56.

– (2021). "Analysis of minimum pulse shape information needed for accurate chest injury prediction in real life frontal crashes". *International journal of crashworthiness* 26:6. Publisher: Taylor & Francis, pp. 684–691.

Iraeus, J. and B. Pipkorn (2019). "Development and validation of a generic finite element ribcage to be used for strain-based fracture prediction". In: *Proceedings of the IRCOBI Conference*. Florence, Italy, pp. 193–210.

Jakobsson, L. and M. Lindman (2005). "Does BMI (body mass index) influence the occupant injury risk pattern in car crashes". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*, pp. 427–430.

James, G., D. Witten, T. Hastie, and R. Tibshirani (2021). *An introduction to statistical learning: with applications in r*. Ed. by G. James, D. Witten, T. Hastie, and R. Tibshirani. Springer Texts in Statistics. Springer US, New York, NY. ISBN: 978-1-07-161418-1. DOI: 10.1007/978-1-0716-1418-1_1.

John, J., C. Klug, M. Kranjec, E. Svenning, and J. Iraeus (2022). "Hello, world! VIVA+: A human body model lineup to evaluate sex-differences in crash protection". *Frontiers in bioengineering and biotechnology* 10. Publisher: Frontiers Media SA, p. 918904.

Joodaki, H., B. Gepner, and J. Kerrigan (2020a). "Leveraging machine learning for predicting human body model response in restraint design simulations". *Computer Methods in Biomechanics and Biomedical Engineering*. Publisher: Taylor & Francis, pp. 1–15.

Joodaki, H., B. Gepner, S.-H. Lee, M. Katagiri, T. Kim, and J. Kerrigan (2021). "Is optimized restraint system for an occupant with obesity different than that for a normal BMI occupant?" *Traffic injury prevention* 22:8. Publisher: Taylor & Francis, pp. 623–628.

Joodaki, H., B. Gepner, T. McMurry, and J. Kerrigan (2020b). "Comparison of injuries of belted occupants among different BMI categories in frontal crashes". *International Journal of Obesity* 44:6. Publisher: Nature Publishing Group, pp. 1319–1329. ISSN: 1476-5497. DOI: 10.1038/s41366-019-0481-2.

Kahane, C. J. (2000). *Fatality reduction by safety belts for front-seat occupants of cars and light trucks: updated and expanded estimates based on 1986-99 FARS data*. US Department of Transportation, National Highway Traffic Safety Administration.

– (2013a). *Effectiveness of pretensioners and load limiters for enhancing fatality reduction by seat belts*. DOT HS 811 835. Washington, DC: National Highway Traffic Safety Administration, p. 835.

– (2013b). *Injury vulnerability and effectiveness of occupant protection technologies for older occupants and women*. DOT HS 811 766. Washington, DC United States: National Highway Traffic Safety Administration.

– (2015). *Lives saved by vehicle safety technologies and associated Federal Motor Vehicle Safety Standards, 1960 to 2012–Passenger cars and LTVs–With reviews of 26 FMVSS and the effectiveness of their associated safety technologies in reducing fatalities, injuries, and crashes*. DOT HS 812 069. Washington, DC: National Highway Traffic Safety Administration.

Kato, D., Y. Nakahira, N. Atsumi, and M. Iwamoto (2018). "Development of human-body model THUMS Version 6 containing muscle controllers and application to injury analysis in frontal collision after brake deceleration". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*. Athens, Greece, pp. 12–14.

Kent, R., J. R. Crandall, J. Bolton, P. Prasad, G. Nusholtz, and H. Mertz (2001). *The influence of superficial soft tissues and restraint condition on thoracic skeletal injury prediction*. SAE Technical Paper.

Kent, R., J. Forman, D. Parent, and S. Kuppa (2007). "Rear seat occupant protection in frontal crashes and its feasibility". In: *20th international conference on the enhanced safety of vehicles*, pp. 18–21.

Kent, R., S.-H. Lee, K. Darvish, S. Wang, C. S. Poster, A. W. Lange, C. Brede, D. Lange, and F. Matsuoka (2005). *Structural and material changes in the aging thorax and their role in crash protection for older occupants*. SAE Technical Paper.

Kim, T., G. Poplin, V. Bollapragada, T. Daniel, and J. R. Crandall (2020). "Monte Carlo method for estimating whole-body injury metrics from pedestrian impact simulation results". *Accident Analysis & Prevention* 147. Publisher: Elsevier, p. 105761.

Kim, T., K. Song, and S.-h. Hong (1, 2023). "Estimation of Whole-Body Injury Metrics for Evaluating Effect of Airbag Deployment". *International Journal of Automotive Technology* 24:3, pp. 633–642. ISSN: 1976-3832. DOI: 10.1007/s12239-023-0053-2.

Kleiven, S. (2007). "Predictors for traumatic brain injuries evaluated through accident reconstructions". *Stapp car crash journal* 51. Publisher: Society of Automotive Engineers SAE, pp. 81–114.

Klinich, K., P. Bowman, C. Flannagan, and J. Rupp (2016). *Injury patterns in motor-vehicle crashes in the united states 1998-2014*. Michigan: University of Michigan Transportation Research Institute.

Klug, C., M. Schachner, A. Eggers, J. Galazka, S. Gargallo, C. Jimenez, J. Kirch, I. Levallois, U. Lobenwein, N. Meissner, et al. (2023). "Euro NCAP Virtual Testing-Crashworthiness". *Enhanced Safety of Vehicles*.

Kuppa, S., J. Wang, M. Haffner, and R. Eppinger (2001). *Lower extremity injuries and associated injury criteria*. SAE Technical Paper.

Laituri, T. R., P. Prasad, K. Sullivan, M. Frankstein, and R. S. Thomas (2005). "Derivation and evaluation of a provisional, age-dependent, AIS3+ thoracic risk curve for belted adults in frontal impacts". *SAE Technical Paper* 2005. Number: 2005-01, p. 0297.

Larsson, K.-J., A. Blennow, J. Iraeus, B. Pipkorn, and N. Lubbe (2021a). "Rib cortical bone fracture risk as a function of age and rib strain: updated injury prediction using finite element human body models". *Frontiers in bioengineering and biotechnology* 9. Publisher: Frontiers Media SA, p. 677768.

Larsson, K.-J., J. Iraeus, S. Holcombe, and B. Pipkorn (2023). "Influences of human thorax variability on population rib fracture risk prediction using human body models". *Frontiers in Bioengineering and Biotechnology* 11. Publisher: Frontiers Media SA, p. 1154272.

Larsson, K.-J., J. Iraeus, B. Pipkorn, and S. Holcombe (2022). "Influence of individual ribcage shape variability on occupant rib fracture risk". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*. Porto, Portugal, pp. 667–684.

Larsson, K.-J., J. Östh, J. Iraeus, and B. Pipkorn (2024). "A first step toward a family of morphed human body models enabling prediction of population injury outcomes". *Journal of Biomechanical Engineering* 146:3. Publisher: American Society of Mechanical Engineers Digital Collection.

Larsson, K.-J., B. Pipkorn, J. Iraeus, J. H. Bolte IV, A. M. Agnew, J. Hu, M. P. Reed, and C. Sunnev{\textbackslash}aang (2019). "Evaluation of the Benefits of Parametric Human Body Model Morphing for Prediction of Injury to Elderly Occupants in Side Impact". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*. Florence, Italy, pp. 150–174.

Larsson, K.-J., B. Pipkorn, J. Iraeus, J. Forman, and J. Hu (2021b). "Evaluation of a diverse population of morphed human body models for prediction of vehicle occupant crash kinematics". *Computer Methods in Biomechanics and Biomedical Engineering*. Publisher: Taylor & Francis, pp. 1–31.

Lederer, W., D. Mair, W. Rabl, and M. Baubin (2004). "Frequency of rib and sternum fractures associated with out-of-hospital cardiopulmonary resuscitation is underestimated by conventional chest X-ray". *Resuscitation* 60:2. Publisher: Elsevier, pp. 157–162.

Leledakis, A., M. Lindman, J. Östh, L. Wågström, J. Davidsson, and L. Jakobsson (1, 2021a). "A method for predicting crash configurations using counterfactual simulations and real-world data". *Accident Analysis & Prevention* 150, p. 105932. ISSN: 0001-4575. DOI: 10.1016/j.aap.2020.105932.

Leledakis, A., J. Östh, J. Davidsson, and L. Jakobsson (1, 2021b). "The influence of car passengers' sitting postures in intersection crashes". *Accident Analysis & Prevention* 157, p. 106170. ISSN: 0001-4575. DOI: 10.1016/j.aap.2021.106170.

Leledakis, A., J. Östh, J. Iraeus, J. Davidsson, and L. Jakobsson (2022). "The influence of occupant's size, shape and seat adjustment in frontal and side impacts". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*. Porto, Portugal, pp. 549–584.

– (2023). "Influence of an individualised shoulder belt position for diverse occupant anthropometries on seatbelt interaction in frontal and side impacts". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*, pp. 639–664.

Lopez-Valdes, F. J., K. Mroz, A. Eggers, B. Pipkorn, J. Muehlbauer, S. Schick, and S. Peldschus (2018). "Chest injuries of elderly postmortem human surrogates (PMHSs) under seat belt and airbag loading in frontal sled impacts: Comparison to matching THOR tests". *Traffic injury prevention* 19, sup2. Number: sup2 Publisher: Taylor & Francis, S55–S63.

López-Valdés, F. J., O. Juste-Lorente, M. Maza-Frechin, B. Pipkorn, C. Sunnevang, A. Lorente, A. Aso-Vizan, and J. Davidsson (2016). "Analysis of occupant kinematics and dynamics in nearside oblique impacts". *Traffic injury prevention* 17, sup1. Publisher: Taylor & Francis, pp. 86–92.

Lowne, R. and J. Neilson (1987). "The development and certification of EUROSID". In: *Eleventh international technical conference on experimental safety vehicles. NHTSA*, p. 1987.

Lubbe, N. and T. Kiuchi (2015). "Injury estimation for advanced automatic collision notification (AACN) in Germany". In: *6th International conference on ESAR "Expert symposium on accident research" reports on the ESAR-Conference on 20th/21th June 2014 at Hannover Medical School*.

McCarthy, M., B. Chinn, and J. Hill (2001). *The effect of occupant characteristics on injury risk and the development of active-adaptive restraint systems*. SAE Technical Paper.

Meijer, R., H. Elrofai, W. Broos, and E. v. Hassel (2013). "Evaluation of an active multi-body human model for braking and frontal crash events". In: *23rd International Technical Conference on the Enhanced Safety of Vehicles (ESV) National Highway Traffic Safety Administration*. tex.organization: NHTSA. Seoul, Korea.

Meyer, S. E., A. O. Nelson, D. A. Hock, J. T. McMillin, and B. R. Herbst (2015). "Assessing the effects of load limiting retractors on occupant motion". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*.

Michaelson, J., J. Forman, R. Kent, and S. Kuppa (2008). "Rear seat occupant safety: kinematics and injury of PMHS restrained by a standard 3-point belt in frontal crashes". *Stapp car crash journal* 52. Publisher: The Stapp Association, p. 295.

Miller, H. J. and V. Maripudi (1996). *Restraint force optimization for a smart restraint system*. SAE Technical Paper.

Mitzkus, J. E. and H. Eyrainer (1984). "Three-point belt improvements for increased occupant protection". *SAE transactions*. Publisher: JSTOR, pp. 1056–1064.

National Center for Statistics and Analysis (2019). *Lives saved in 2017 by restraint use and minimum- drinking-age laws*. Traffic Safety Facts Crash Stats. DOT HS 812 683. Issue: DOT HS 812 683. Washington, DC: National Highway Traffic Safety Administration.

– (2023). *Traffic safety facts 2021: A compilation of motor vehicle traffic crash data*. DOT HS 813 527. Washington, D.C., United States: National Highway Traffic Safety Administration.

Neathery, R. F., C. K. Kroell, and H. J. Mertz (1975). "Prediction of thoracic injury from dummy responses". *SAE Transactions*. Publisher: JSTOR, pp. 3178–3187.

Neilson, L., R. Lowne, C. Tarriere, F. Bendjellal, D. Gillet, J. Maltha, D. Cesari, and R. Bouquet (1985). *The EUROSID side impact dummy*. SAE Technical Paper.

NHTSA (31, 2023). *Research Testing Databases | NHTSA*. URL: https://www.nhtsa.gov/research-data/research-testing-databases (visited on 03/31/2023).

– (2024). *NHTSA File Downloads | NASS | NHTSA*. URL: https://www.https://www.nhtsa.gov/file-downloads?p=nhtsa/downloads/NASS/ (visited on 07/11/2024).

NHTSA VIN DECODER (2024). *Welcome to VIN Decoding :: provided by vPIC*. URL: https://vpic.nhtsa.dot.gov/decoder/ (visited on 07/11/2024).

Noh, E. Y., J. R. E. Atwood, E. Lee, and M. Craig (2022). *Female Crash Fatality Risk Relative to Males for Similar Physical Impacts*. DOT HS 813 358. Washington, D.C: National Highway Traffic Safety Administration.

O'Neill, B. (2009). "Preventing passenger vehicle occupant injuries by vehicle design—a historical perspective from IIHS". *Traffic injury prevention* 10:2. Publisher: Taylor & Francis, pp. 113–126.

Östh, J., K. Bohman, and L. Jakobsson (2023). "Head injury criteria assessment using head kinematics from crash tests and accident reconstructions". *Traffic injury prevention* 24:1. Publisher: Taylor & Francis, pp. 56–61.

Östh, J., M. Mendoza-Vazquez, A. Linder, M. Y. Svensson, and K. Brolin (2017). "The VIVA OpenHBM finite element 50th percentile female occupant model: whole body model development and kinematic validation". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*. Antwerp, Belgium, pp. 13–15.

Östling, M., L. Eriksson, M. Dahlgren, and J. Forman (2023). "Frontal head-on car-to-heavy goods vehicle crashes effect on the restraint system". In: *27th international technical conference on the enhanced safety of vehicles (ESV) national highway traffic safety administration*. Number: 23-0198.

Östling, M., H. Jeppsson, and J. Forman (2024). "Injury Observations in High-exposure, Low-severity Frontal Car Crashes - a GIDAS Investigation". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*. Stockholm, Sweden.

Östmann, M. and L. Jakobsson (2016). "An examination of pre-crash braking influence on occupant crash response using an active human body model". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*. Malaga, Spain, pp. 14–16.

Page, M. (2001). *Performance of the prototype WorldSID dummy in side impact crash tests*. SAE Technical Paper.

Parent, D., M. Craig, S. Ridella, and J. McFadden (2013). "Thoracic biofidelity assessment of the THOR mod kit ATD". In: *23rd enhanced safety of vehicles conference, paper*. Number: 13-0327, p. 2013.

Park, J., S. M. Ebert, M. P. Reed, and J. J. Hallman (2016). "Statistical models for predicting automobile driving postures for men and women including effects of age". *Human factors* 58:2. Publisher: SAGE Publications Sage CA: Los Angeles, CA, pp. 261–278.

Perez-Rapela, D., J.-P. Donlon, J. L. Forman, J. R. Crandall, B. Pipkorn, B. K. Shurtz, and C. Markusic (2019). "PMHS and WorldSID kinematic and injury response in far-side events in a vehicle-based test environment". *Stapp car crash journal* 63. Publisher: The Stapp Association, pp. 83–126.

Perez-Rapela, D., J. L. Forman, S. H. Huddleston, and J. R. Crandall (2020). "Methodology for vehicle safety development and assessment accounting for occupant response variability to human and non-human factors". *Computer Methods in Biomechanics and Biomedical Engineering*. Publisher: Taylor & Francis, pp. 1–16.

Pipkorn, B., J. Iraeus, M. Björklund, O. Bunketorp, and L. Jakobsson (2019). "Multi-scale validation of a rib fracture prediction method for human body models". In: *Proceedings of the IRCOBI Conference*. Florence, Italy, pp. 175–192.

Pipkorn, B., J. Iraeus, M. Lindkvist, P. Puthan, and O. Bunketorp (2020). "Occupant injuries in light passenger vehicles—A NASS study to enable priorities for development of injury prediction capabilities of human body models". *Accident Analysis & Prevention* 138, p. 105443. ISSN: 0001-4575. DOI: https://doi.org/10.1016/j.aap.2020.105443.

Pipkorn, B., L. Jakobsson, J. Iraeus, and J. Östh (2023). "THE SAFER HBM–A human body model for seamless integrated occupant analysis for all road users". In: *27th international technical conference on the enhanced safety of vehicles (ESV) national highway traffic safety administration*. Number: 23-0242.

Piqueras, A., B. Pipkorn, J. Iraeus, M. Maza-Frechín, and F. J. López-Valdés (2022). "Assessment of in situ chest deflection of post mortem human subjects (PMHS) and personalized human body models (HBM) in nearside oblique impacts". *Traffic injury prevention* 23:4. Publisher: Taylor & Francis, pp. 181–186.

Pizarroso, J., J. Portela, and A. Muñoz (30, 2022). "NeuralSens: Sensitivity Analysis of Neural Networks". *Journal of Statistical Software* 102, pp. 1–36. ISSN: 1548-7660. DOI: 10.18637/jss.v102.i07.

Prasad, P., S. D. Barbat, A. Kalra, and D. J. Dalmotas (2024). "Evaluation of DAMAGE algorithm in frontal crashes". *Stapp car crash journal* 67:2023, pp. 171–179.

Prasad, P. and R. P. Daniel (1984). "A biomechanical analysis of head, neck, and torso injuries to child surrogates due to sudden torso acceleration". *SAE transactions*. Publisher: JSTOR, pp. 784–799.

Pulgar, F. J., A. J. Rivera, F. Charte, and M. J. del Jesus (2017). "On the Impact of Imbalanced Data in Convolutional Neural Networks Performance". In: *Hybrid Artificial Intelligent Systems*. Ed. by F. J. Martínez de Pisón, R. Urraca, H. Quintián, and E. Corchado. Springer International Publishing, Cham, pp. 220–232. ISBN: 978-3-319-59650-1. DOI: 10.1007/978-3-319-59650-1_19.

Ranmal, A., J. Shaikh, and N. Lubbe (2024). "Rib and sternum fracture risks for restrained occupants in frontal car crashes". *Traffic injury prevention* 25:4. Publisher: Taylor & Francis, pp. 616–622.

Reed, M. P., S. M. Ebert, M. L. Jones, and J. J. Hallman (2020). "Prevalence of non-nominal seat positions and postures among front-seat passengers". *Traffic injury prevention* 21, sup1. Publisher: Taylor & Francis, S7–S12.

Reed, M. P., M. A. Manary, C. A. Flannagan, and L. W. Schneider (2002). "A statistical method for predicting automobile driving posture". *Human Factors* 44:4. Number: 4 Publisher: SAGE Publications Sage CA: Los Angeles, CA, pp. 557–568.

Ressi, F. and C. Klug (2024). "Stochastic Replication of Injury Risks in Frontal Collisions with Different Human Body Models in a Generic Vehicle Interior". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*. Stockholm, Sweden.

Ressi, F., C. Leo, C. Klug, and W. Sinz (2022). "Protection challenges in seat positions with large rearward adjustment in frontal collisions: An approach using stochastic human body model simulations". *Frontiers in future transportation* 3. Publisher: Frontiers Media SA, p. 914481.

Rodarius, C., L. van Rooij, and R. de Lange (2007). "Scalability of Human Models". In: 20th International Technical Conference on the Enhanced Safety of Vehicles (ESV) National Highway Traffic Safety Administration. Number: 07-0314. Lyon, France. URL: https://trid.trb.org/view/1365937 (visited on 10/25/2022).

Rupp, J. D., C. A. Flannagan, and S. M. Kuppa (2010). "Injury risk curves for the skeletal knee–thigh–hip complex for knee-impact loading". *Accident Analysis & Prevention* 42:1. Publisher: Elsevier, pp. 153–158.

Rupp, J. D., C. A. Flannagan, A. J. Leslie, C. N. Hoff, M. P. Reed, and R. M. Cunningham (2013). "Effects of BMI on the risk and frequency of AIS 3+ injuries in motor-vehicle crashes". *Obesity* 21:1, E88–E97. ISSN: 1930-739X. DOI: 10.1002/oby.20079.

Ryb, G. E., P. C. Dischinger, G. McGwin, and R. L. Griffin (2011). "Crash-related mortality and model year: are newer vehicles safer?" In: *Association for the Advancement of Automotive Medicine*. Vol. 55. Association for the Advancement of Automotive Medicine, p. 113.

SAE (2007). "Instrumentation for impact test: Part 1 - electronic instrumentation compiled by SAE international". SAE J211-1. Publisher: SAE International.

Samaha, R. R., K. Digges, T. Fesich, and M. Authaler (2011). "Frontal crash testing and vehicle safety designs: A historical perspective based on crash test studies". *Active Safety and the Mobility Industry*. Publisher: SAE International, p. 83.

Sammut, C. and G. I. Webb (2010a). "Accuracy". In: *Encyclopedia of machine learning*. Springer US, Boston, MA, pp. 9–10. ISBN: 978-0-387-30164-8. DOI: 10.1007/978-0-387-30164-8_3.

– (2010b). "F1-measure". In: *Encyclopedia of machine learning*. Springer US, Boston, MA, pp. 397–397. ISBN: 978-0-387-30164-8. DOI: 10.1007/978-0-387-30164-8_298.

Schneider, B., D. Kofler, G. A. D'Addetta, H. Freienstein, M. Wolkenstein, and C. Klug (20, 2022). "Approach for machine learning based design of experiments for occupant simulation". *Frontiers in Future Transportation* 3. Publisher: Frontiers. ISSN: 2673-5210. DOI: 10.3389/ffutr.2022.913852.

Schoell, S. L., A. A. Weaver, J. E. Urban, D. A. Jones, J. D. Stitzel, E. Hwang, M. P. Reed, and J. D. Rupp (2015). *Development and validation of an older occupant finite element model of a mid-sized male for investigation of age-related injury risk*. SAE Technical Paper.

Schubert, A., N. Erlinger, C. Leo, J. Iraeus, J. John, and C. Klug (2021). "Development of a 50th percentile female femur model". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*. International Research Council on the Biomechanics of Injury (IRCOBI …, pp. 308–32.

Schulze, C., H. Hoppe, W. Schweitzer, N. Schwendener, S. Grabherr, and C. Jackowski (2013). "Rib fractures at postmortem computed tomography (PMCT) validated against the autopsy". *Forensic science international* 233:1. Publisher: Elsevier, pp. 90–98.

Seacrist, T., S. Balasubramanian, J. F. García-España, M. R. Maltese, K. B. Arbogast, F. J. Lopez-Valdes, R. W. Kent, H. Tanji, and K. Higuchi (2010). "Kinematic comparison of pediatric human volunteers and the Hybrid III 6-year-old anthropomorphic test device". In: *Association for the Advancement of Automotive Medicine*. Vol. 54, p. 97.

Segui-Gomez, M., F. Lopez-Valdes, and J. R. Crandall (2009). "Characterizing the distribution of injury and injury severity for belted front-seat occupants involved in frontal crashes". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*. York, UK.

Segui-Gomez, M. (2000). "Driver air bag effectiveness by severity of the crash." *American Journal of Public Health* 90:10. Publisher: American Public Health Association, p. 1575.

Seguí-Gómez, M. and F. J. Lopez-Valdes (2011). "Injury severity scaling". In: *Injury research: theories, methods, and approaches*. Springer, pp. 281–295.

Shin, H., T. Yeo, and W. Ha (2007). *The numerical study for the adaptive restraint system*. SAE Technical Paper.

Simcenter Madymo (2020). *Mode Manual*. Version 2020.1. Siemens Digital Industries Software, The Hague, The Netherlands.

– (2021). *Simcenter Madymo™ Quality Report Release Update: Active Human 50th Facet Q Model, version 3.3 (R2021.1)*. Siemens Industry Software and Services BV.

Solhed, J. (2022). *Applicability of graph neural networks to predict human variability in human body model rib strain predictions*.

Suarez-del Fueyo, R., M. Junge, F. Lopez-Valdes, H. C. Gabler, L. Woerner, and S. Hiermaier (2020). "Injury patterns within clusters of seriously injured occupants comparing real-world crashes in the United States and the European Union". *Traffic injury prevention* 21, sup1. Number: sup1 Publisher: Taylor & Francis, S78–S83.

Sun, W., J. Liu, J. Hu, J. Jin, K. Siasoco, R. Zhou, and R. Mccoy (2023). "Adaptive restraint design for a diverse population through machine learning". *Frontiers in Public Health* 11. Publisher: Frontiers Media SA, p. 1202970.

Swanson, J., T. Rockwell, N. M. Beuse, L. Summers, S. Summers, and B. Park (2003). "Evaluation of stiffness measures from the US new car assessment program". In: *Proceedings: International technical conference on the enhanced safety of vehicles*. Vol. 2003. National Highway Traffic Safety Administration, 13–p.

Syakur, M. A., B. K. Khotimah, E. Rochman, and B. D. Satoto (2018). "Integration k-means clustering method and elbow method for identification of the best customer profile cluster". In: *IOP conference series: materials science and engineering*. Vol. 336. IOP Publishing, p. 012017.

Takhounts, E., M. J. Craig, K. Moorhouse, J. McFadden, and V. Hasija (2013). *Development of brain injury criteria (BrIC)*. SAE Technical Paper.

Takhounts, E., V. Hasija, and M. J. Craig (2019). "BrIC and field braininjury risk". In: *26th ESV conference proceedings, eindhoven, netherlands. Paper*. Number: 19-0154.

Thomas, P. (2013). "Developments in the risk of crash involvement and injury to car occupants by model year using vehicle specific exposure data". In: *Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) Conference*. Publisher: Loughborough University. Goteborg, Sweden.

Ting, K. M., C. Sammut, and G. I. Webb (2010). "Precision and recall". In: *Encyclopedia of machine learning*. Springer US, Boston, MA, pp. 781–781. ISBN: 978-0-387-30164-8. DOI: 10.1007/978-0-387-30164-8_652.

Versace, J. (1971). "A review of the severity index". *Warrendale: SAE International*. Publisher: SAE Technical Paper.

Viano, D. C., C. S. Parenteau, and M. L. Edwards (19, 2008). "Crash Injury Risks for Obese Occupants Using a Matched-Pair Analysis". *Traffic Injury Prevention* 9:1. Publisher: Taylor & Francis, pp. 59–64. ISSN: 1538-9588. DOI: 10.1080/15389580701737645.

Vision Zero Initiative - TRIMIS - European Commission (31, 2017). TRIMIS. URL: https://trimis.ec.europa.eu/project/vision-zero-initiative (visited on 02/07/2021).

Visvalingam, M. and J. D. Whyatt (1993). "Line generalisation by repeated elimination of points". *The cartographic journal* 30:1. Number: 1 Publisher: Taylor & Francis, pp. 46–51.

Walz, M. (2004). "NCAP test improvements with pretensioners and load limiters". *Traffic injury prevention* 5:1. Publisher: Taylor & Francis, pp. 18–25.

Wang, Q., S. Gan, W. Chen, Q. Li, and B. Nie (2021). "A data-driven, kinematic feature-based, near real-time algorithm for injury severity prediction of vehicle occupants". *Accident Analysis & Prevention* 156. Publisher: Elsevier, p. 106149.

WHO (2022). *Preventing injuries and violence: an overview*. Pages: 13 p. World Health Organization.

– (13, 2023a). *Global status report on road safety 2023*. Licence: CC BY-NC-SA 3.0 IGO. ISBN: 978-92-4-008651-7. Geneva, Switzerland: World Health Organization. URL: https://iris.who.int/bitstream/handle/10665/375016/9789240086517-eng.pdf?sequence=1.

– (13, 2023b). *Road traffic injuries*. URL: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries (visited on 08/15/2024).

WHO and U. R. Commissions (20, 2021). *Global Plan for the Decade of Action for Road Safety 2021-2030*. License: CC BY-NC-SA 3.0 IGO. World Health Organization and the United Nations Regional Commissions. URL: https://cdn.who.int/media/docs/default-source/documents/health-topics/road-traffic-injuries/global-plan-for-road-safety.pdf?sfvrsn=65cf34c8_35&download=true.

Xu, R.-F. and S.-J. Lee (1, 2015). "Dimensionality reduction by feature clustering for regression problems". *Information Sciences* 299, pp. 42–57. ISSN: 0020-0255. DOI: 10.1016/j.ins.2014.12.003.

Yoganandan, N., A. M. Nahum, and J. W. Melvin (2014). *Accidental injury: biomechanics and prevention*. Springer.

A List of publications

A.1 Articles published in peer-reviewed academic journals

- Valdano, M., Jiménez-Octavio, J.R., López-Valdés, F.J., The effect of seat belt pretensioners and load limiters in the reduction of MAIS 2+, MAIS 3+, and fatal injuries in real-world frontal crashes. Accident Analysis & Prevention. Vol. 190, pp. 107180-1 - 10718-6, September 2023.
- Valdano, M., Jiménez-Octavio, J.R., Pipkorn, B., Otero-Peinador, A., Sánchez-Merchante, L.F., López-Valdés, F.J., Evaluation of AIS3+ car occupant injuries using deterministic and probabilistic methods in frontal crashes. Computer Methods in Biomechanics and Biomedical Engineering. Vol. 27, nº. 12, pp. 1714 - 1730, September 2024.
- Asensio-Gil, J.M., Jiménez-Octavio, J.R., Carnicero, A., Valdano, M., Guzmán, D., López-Valdés, F.J., Numerical analysis of injuries of e-scooter riders in frontal impacts against SUVs. Results in Engineering. Vol. 21, pp. 101936-1 - 101936-10, Marzo 2024. [Online: Febrero 2024] JCR JCR: 5,000 Q1 (2022) - SJR: 0,592 Q2 (2022)
- Vives-Torres, C.M., Valdano, M., Jiménez-Octavio, J.R., Muehlbauer, J., Peldschus, S., López-Valdés, F.J., Comparison of upper neck loading in young adult and elderly volunteers during low speed frontal impacts. Frontiers in Bioengineering and Biotechnology. Vol. 9, pp. 682974-1 - 682974-12, 2021.

A.2 Articles presented at international conferences

- López-Valdés, F.J., Jiménez-Octavio, J.R., Vives-Torres, C.M., Alvarez-Fernandez, J., Asensio-Gil, J.M., Valdano, M., Rodríguez-Morcillo, C., Östling, M., Lebarbé, M., Masuda, M., Beillas, P., Eggers, A., Petit, P., Lemmen, P., Richard, O., The impact response of four male PMHS to frontal impacts in reclined positions, International Research Council on Biomechanics of Injury - IRCOBI Europe 2024, Stockholm (Sweden). 11-13 September 2024.
- Asensio-Gil, J.M., Jokubaityte, E., Valdano, M., Oleaga Ortega, N., Llobet-Cusí, L., López-Valdés, F.J., Vives-Torres, C.M., Alvarez-Fernandez, J., Parametric study on the influence of impact energy and inflation pressure on the effectiveness of a cervical airbag, International Research Council on Biomechanics of Injury - IRCOBI Europe 2024, Stockholm (Sweden). 11-13 September 2024.
- López-Valdés, F.J., Ferris, S., Valdano, M., The influence of body weight in the performance of child restraint systems. A preliminary study with the PIPER human body model, 21st Protection of Children in Cars Conference - POCC 2023, Munich (Germany). 28-30 November 2023.
- Valdano, M., Jiménez-Octavio, J.R., López-Valdés, F.J., The effect of seat belt pretensioners and load limiters in the reduction of MAIS 2+, MAIS 3+, and fatal injuries

in real-world frontal crashes, 67th AAAM Annual Scientific Conference - AAAM 2023, Indianapolis (Estados Unidos de América). 03-06 October 2023.

- Valdano, M., Jiménez-Octavio, J.R., Pipkorn, B., López-Valdés, F.J., Characterisation of the features of seat-belt systems based on the analysis of large crash databases, International Research Council on Biomechanics of Injury - IRCOBI Europe 2023, Cambridge (United Kingdom). 13-15 September 2023.
- Vives-Torres, C.M., Valdano, M., Alvarez-Fernandez, J., Asensio-Gil, J.M., Rodríguez-Morcillo, C., Millet-Solanelles, M., Oleaga Ortega, N., Llobet-Cusí, L., López-Valdés, F.J., The effectiveness of cervical airbags in the control of head and neck kinematics, International Research Council on Biomechanics of Injury - IRCOBI Europe 2023, Cambridge (United Kingdom). 13-15 September 2023.
- Valdano, M., Jiménez-Octavio, J.R., López-Valdés, F.J., Pipkorn, B., Evaluation of serious occupant injuries using a deterministic and probabilistic method in frontal crashes, 66th AAAM Annual Scientific Conference - AAAM 2022, Portland (United States). 11-14 October 2022.
- Valdano, M., Asensio-Gil, J.M., Jiménez-Octavio, J.R., Cabello-Reyes, M., Vasserot-Tolmos, R., López-Valdés, F.J., Parametric analysis of the effect of CRS seatback angle in dummy measurements in frontal impacts, International Research Council on Biomechanics of Injury - IRCOBI Europe 2022, Porto (Portugal). 14-16 September 2022.
- Valdano, M., Jiménez-Octavio, J.R., Vives-Torres, C.M., López-Valdés, F.J., Pipkorn, B., Assessment of madymo active human body model kinematics and dynamics by means of human volunteer response at low-speed frontal impacts, International Research Council on Biomechanics of Injury - IRCOBI Europe 2021, Zurich (Switzerland) Online. 08-10 September 2021.

B Seatbelt manual identification precedure and testing

B.1 Guideline to identify the pre-tensioner configuration

Some force-time history curves were used to exemplify how to identify the pre-tensioner configuration. The following figure shows two force-time history curves, one with a pre-tensioning device and the other without.

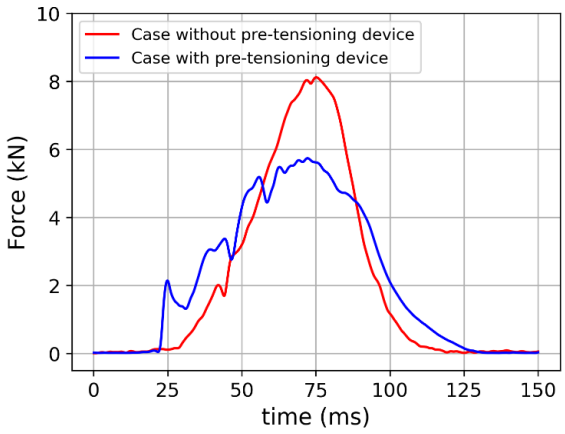


Figure B.1: Force-time history curves for two crash tests. The tension force was measured at the shoulder belt of the driver's belt. A CFC 180 was applied to both signals.

Two possible scenarios can be observed in Figure B.1, the tension force in the belt increased progressively (i.e., a restraint system without a pre-tensioning device in red), and the tension force increased suddenly (i.e., a restraint system with a pre-tensioning device in blue). Figure B.2 "a" shows a zoom-in of the force-time history curves shown above, where only the first 30 ms of the crash can be observed. As can be seen in the figure, there is no sudden increase in the tension force in the case without a pre-tensioning device (red curve). Figure B.2 "b" shows only the case with a pre-tensioning device and the characteristics that were identified in the force-time history curve. The "TTF identified" shows the time-to-fire that was asked to be identified, which is the time at which the tension force begins the sudden increase. The "actual TTF" (i.e., the firing command) would be earlier compared to the "TTF identified", and it is not influenced by the pre-tensioner delay, belt slack, belt geometry, and transducer position. However, it is not possible to identify this with the available information. The "PP force" is the force of the pre-tensioning device, and it was identified as the first peak after the "TTF identified" using the filtered signal.

Some cases showed an increase in the shoulder and lap belt, as shown in Figure B.3. Case number 1, shown in blue, is the same as the one shown in blue in Figure B.1 and Figure B.2.

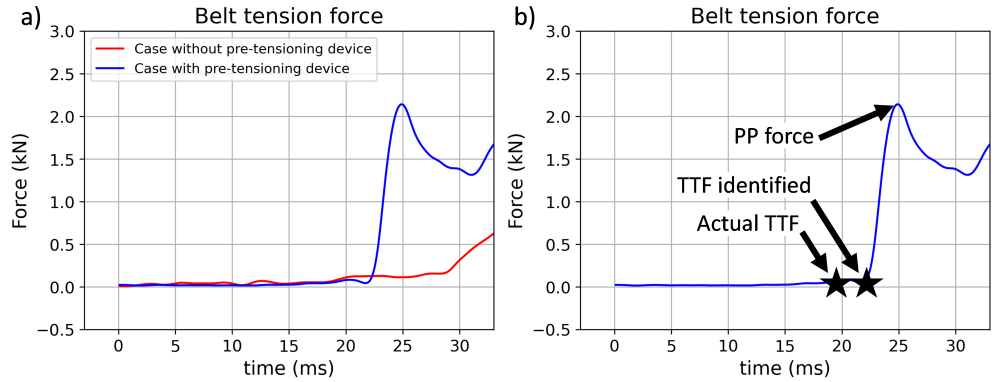


Figure B.2: Subgfigure a contains force-time history curves for two crash tests (same as in Figure B.1), only the first 30 ms. The tension force was measured at the shoulder belt of the driver's belt. A CFC 180 was applied to both signals. Subfigure b contains the force-time history curve of a case with a pre-tensioning device and the characteristics to be identified of the restraint system.

The other curves, in red, show a case of pre-tensioning in the lap and shoulder belt. The force-time history curves measured at the lap belt are shown with a dashed line. The case with a shoulder belt pre-tensioner (in blue) showed a transference of the tension force through the buckle from the shoulder belt to the lap belt. However, the peak in the lap belt tension force can hardly be identified as the result of the pre-tensioning device in the buckle or the anchor plate, as the slope of the force-time history curve is not as high as in cases with a pre-tensioning device. However, the case with shoulder and lap belt pre-tensioners (in red) showed peaks in both seatbelt sections. These pre-tensioner devices were fired at different times as the shoulder belt tension force increased first. Furthermore, the force-time history curve of the lap belt showed two different slopes. Firstly, a low slope was associated with a transference of the tension force through the buckle, similar to the first case in blue, which had a similar slope. Secondly, a high slope associated with a pre-tensioning device had a similar slope as the one associated with pre-tensioning devices in the shoulder belt force-time history curves.

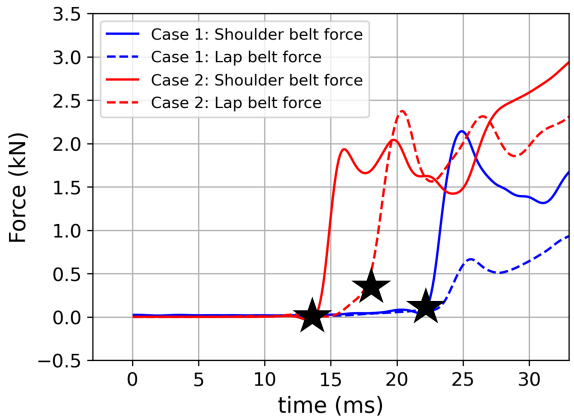


Figure B.3: Force-time history curves for two crash tests, only first 30 ms. The tension force was measured at the shoulder belt (solid line) and lap belt (dashed line) of the driver's belt. A CFC 180 was applied to both signals. The stars show the time at which the TTF was identified.

B.2 Guideline to identify the load-limiting configuration

Some force-time history curves were used to exemplify how to identify the load-limiting configuration. For this task, cases were gathered without load-limiting or with load-limiting devices in one or two stages. Figure B.4 shows three force-time history curves where no load limiting was observed. These force-time history curves showed a range of peak tension force between 5 and 10 kN. However, none of the force-time history curves indicated a sustained tension force over time.

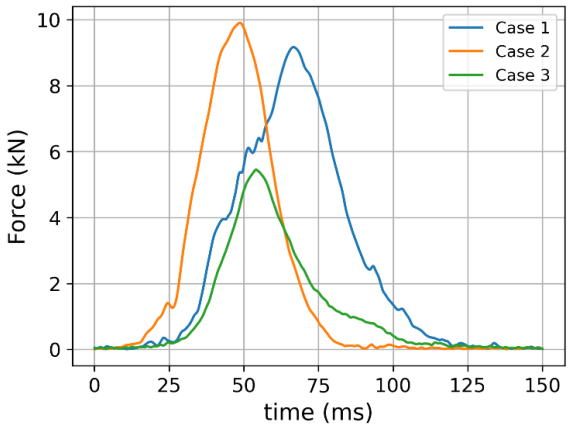


Figure B.4: Force-time history curves for three crash tests. The tension force was measured at the shoulder belt of the driver's belt. A CFC 180 was applied to all signals.

Figure B.5 shows six force-time history curves where a load-limiting device was identified. Some of these force-time history curves were identified as those with a load-limiting device of one stage and others with two. Subfigure "a" shows three force-time history curves where a one-stage load-limiting device was identified. Subfigure "b" shows three force-time history curves where a double-stage load-limiting device was identified.

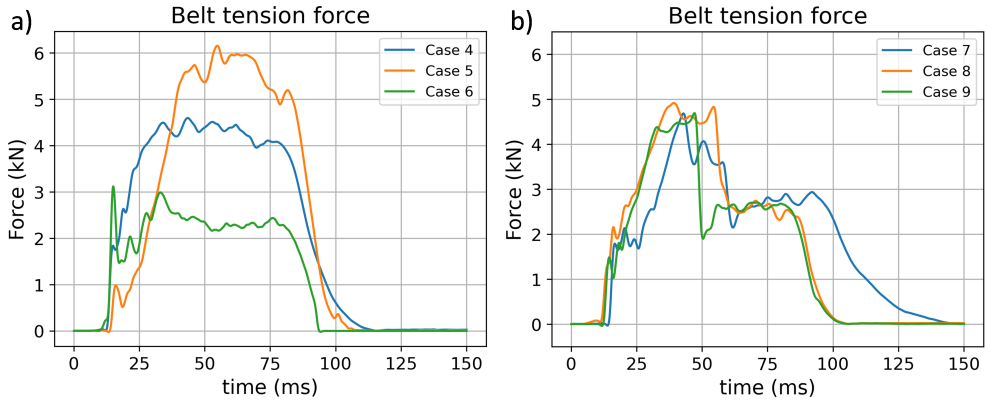


Figure B.5: Force-time history curves for six crash tests. The tension force was measured at the shoulder belt of the driver's belt. A CFC 180 was applied to all signals. a) Case 4, 5, and 6 with a one-stage load-limiting device used. b) Cases 7, 8, and 9 with a double-stage load-limiting device used.

The load-limiting force of the device was identified as the mean force of the stage. This mean was obtained visually and one decimal precision was asked for its identification. Figure B.6 shows some examples of how the mean force of the stage was identified.

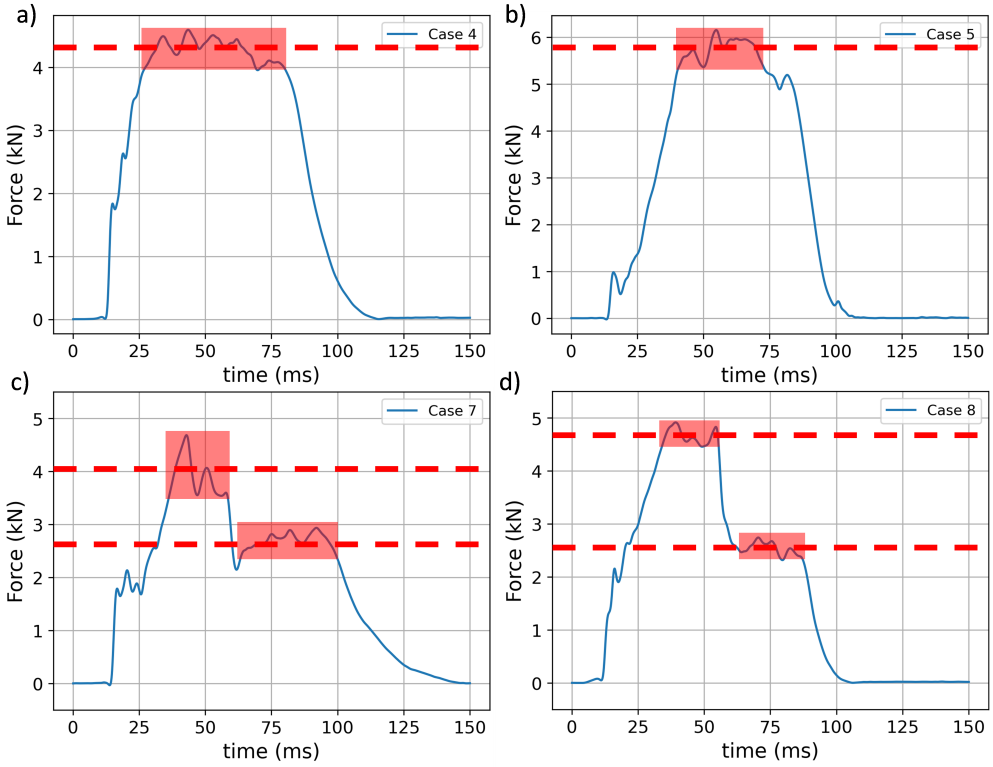


Figure B.6: Force-time history curves for four crash tests. The tension force was measured at the shoulder belt of the driver's belt. A CFC 180 was applied to all signals. The mean load-limiting force is shown with a red dashed line and each stage of the load-limiting device is shown with a red shadowed square. a) Case 4 with a one-stage load-limiting device used. b) Case 5 with a one-stage load-limiting device used. c) Case 7 with a double-stage load-limiting device used. d) Case 8 with a double-stage load-limiting device used.

Subfigure "a" shows case 4 from Figure B.5, this restraint system was identified as one with a one-stage load-limiting device. The stage of the load limiter is shown in red and the mean value of the load-limiting force is shown with a dashed line. Subfigure "b" shows case 5 from Figure B.5, this restraint system was also identified as one with a one-stage load-limiting device. Subfigure "c" shows case 7 from Figure B.5, this restraint system was identified as one with a double-stage load-limiting device. However, this case also shows load-limiting stages with a slope in the load-limiting force. The first stage of the load-limiting device showed a negative slope, which resulted in a decreasing load-limiting force. The second stage of the load-limiting device showed a positive slope, which resulted in an increasing load-limiting force. Subfigure "d" shows case 8 from Figure B.5, this restraint system was also identified as one with a double-stage load-limiting device.

B.3 Testing of identification algorithm for seat belt pre-tensioner and load limiter characteristics

The following figure and tables contain an example of the application of the identification algorithm, the manual identification, and the actual pre-tensioner and load limiter characteristics provided by the seatbelt manufacturer. A similar restraint system configuration was used in Östling et al. (2023).

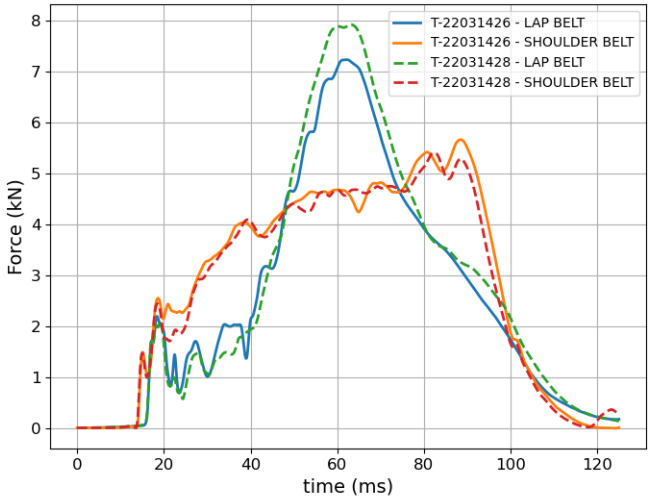


Figure B.7: Force-time history curves for the 'T-22031426' and 'T-22031428' crash tests. The tension force was measured at the shoulder and lap belt of the occupant. A CFC 180 was applied to all signals.

Table B.1: Seat-belt pre-tensioner and load limiter characteristics - Lap belt - Test T-22031426.

Source	Pre-tensioner force [kN]	Pre-tensioner TTF [ms]	Load limiting force [kN]
Manufacturer	2	15	-
Manual identification	2.1	16	-
Application of identification algorithm	1.8	17	-

Table B.2: Seat-belt pre-tensioner and load limiter characteristics - Shoulder belt - Test T-22031426.

Source	Pre-tensioner force [kN]	Pre-tensioner TTF [ms]	Load limiting force [kN]
Manufacturer	2	10	4
Manual identification	2.5	13	4.5
Application of identification algorithm	2.1	14.8	4.6

Table B.3: Seat-belt pre-tensioner and load limiter characteristics - Lap belt - Test T-22031428.

Source	Pre-tensioner force [kN]	Pre-tensioner TTF [ms]	Load limiting force [kN]
Manufacturer	2	15	-
Manual identification	2	16	-
Application of identification algorithm	1.8	17	-

Table B.4: Seat-belt pre-tensioner and load limiter characteristics - Shoulder belt - Test T-22031428.

Source	Pre-tensioner force [kN]	Pre-tensioner TTF [ms]	Load limiting force [kN]
Manufacturer	2	10	4
Manual identification	2.5	13	4.5
Application of identification algorithm	2.0	14.7	4.6

C Distribution of the variables in the reference data used to obtain the metamodels

C.1 Crash variable distributions

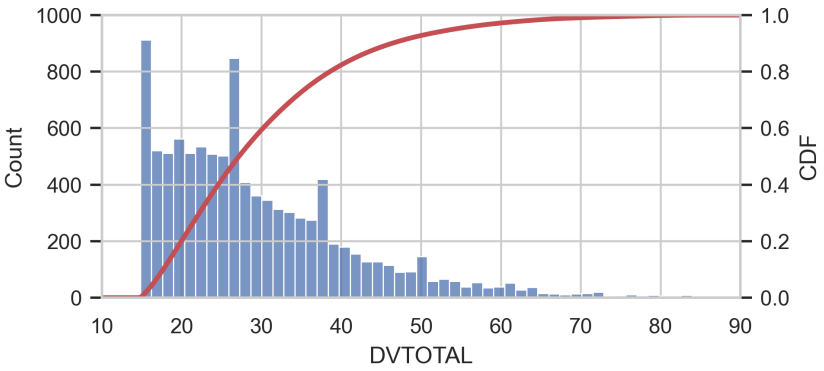


Figure C.1: Distribution of impact speeds in the reference data, in blue, and its cumulative distribution function, in red.

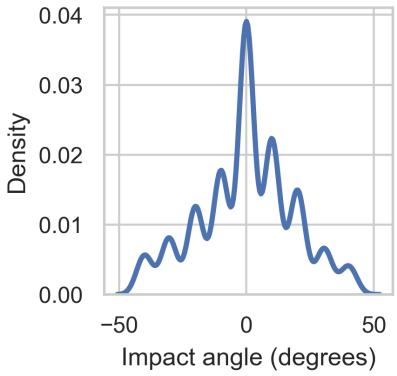


Figure C.2: Density probability function of the impact angle from the reference data obtained using a kernel density estimate.

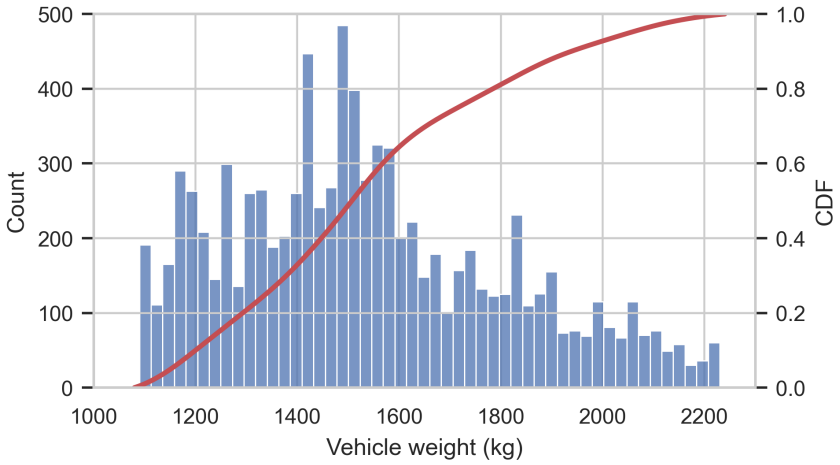


Figure C.3: Distribution of vehicle weights in the reference data, in blue, and its cumulative distribution function, in red.

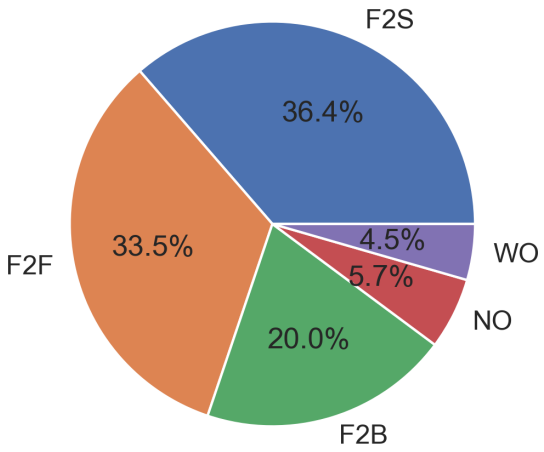


Figure C.4: Distribution of collision opponent classified as front-to-front (F2F), front-to-side (F2S), front-to-back (F2B), a narrow object (NA), and a wide object (WO).

C.2 Restraint system variable distributions

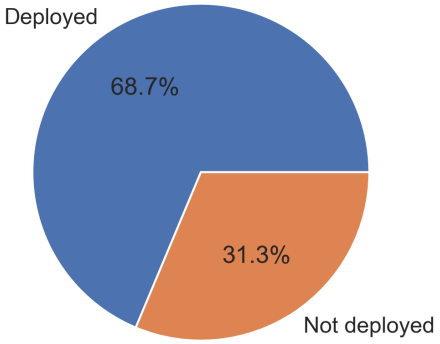


Figure C.5: Distribution of airbag deployment in the whole reference dataset.

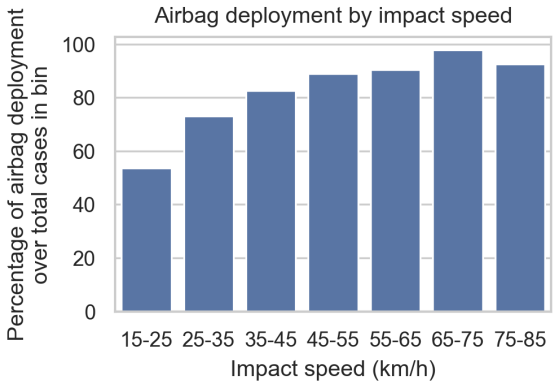


Figure C.6: Distribution of airbag deployment in the reference dataset grouped by impact speed.

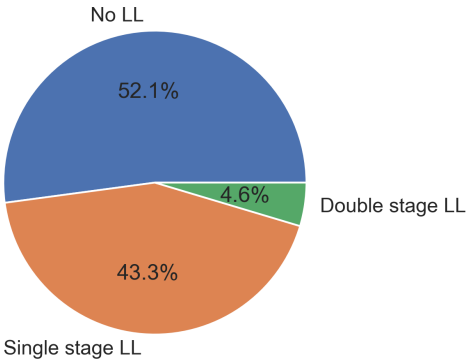


Figure C.7: Distribution of configuration of load-limiting devices in the reference dataset.

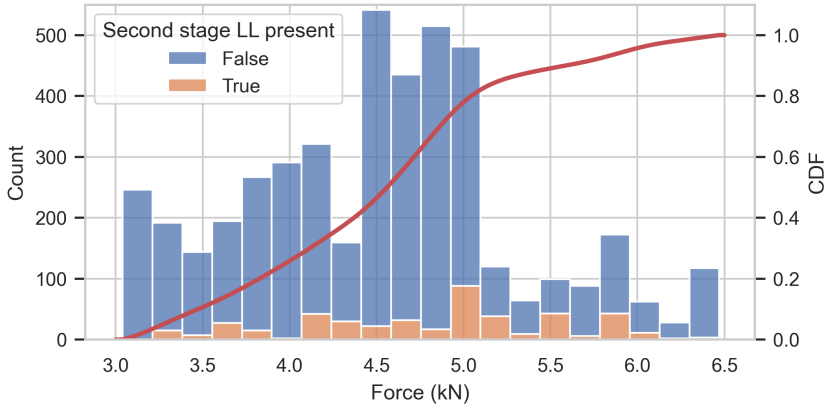


Figure C.8: Distribution of the forces used in the load-limiting devices (only first stage if second stage present) in the reference dataset. If a double-stage load-limiting device was included, the distribution of the forces in the first stage is shown in orange.

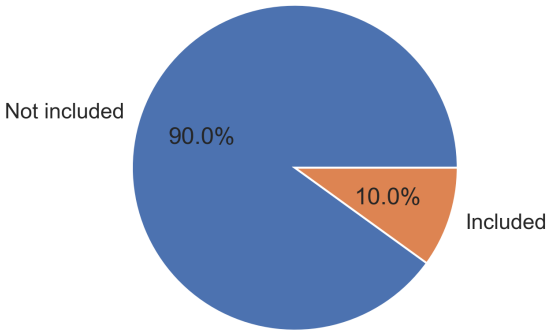


Figure C.9: Distribution of the inclusion of load-limiting devices with a second stage if a load-limiting device was included in the reference dataset.

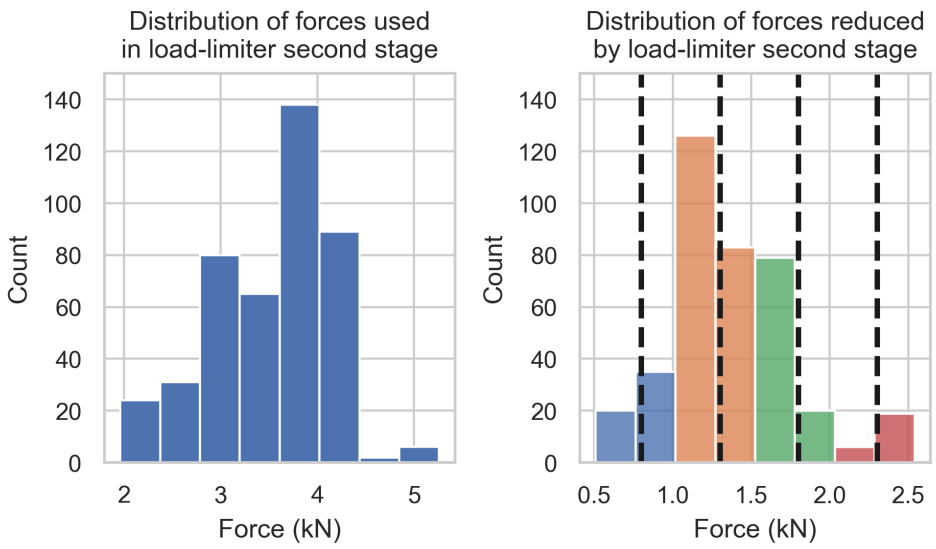


Figure C.10: Distribution of the configurations used in load-limiting devices with a double stage. The left figure illustrates the distribution of forces applied in the second stage of the load-limiting device, if included. The right figure shows the distribution of the decrease in load-limiting force between the first and second stages of the device. Vertical dashed lines show the decrease in force used to obtain the metamodels. Coloured bars depict the force ranges used to obtain the probability of each configuration, as shown with the dashed lines.

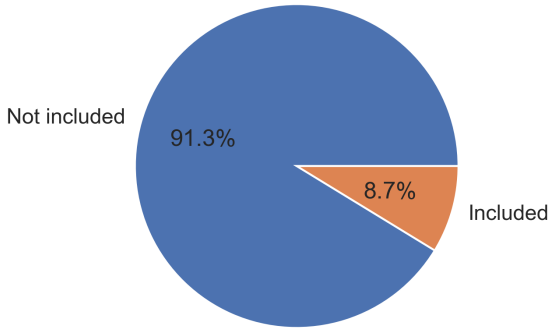


Figure C.11: Distribution of the inclusion of a pre-tensioning decide in the end-bracket if a load-limiting device was included in the reference dataset.

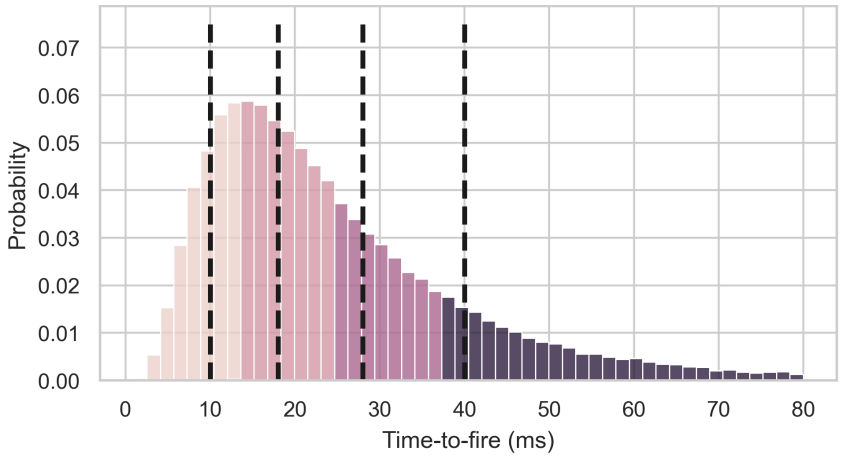


Figure C.12: Distribution of time-to-fire from Iraeus et al. (2016). Vertical dashed lines show the time-to-fire used to represent the distribution in the process of obtaining the metamodels. Coloured bars depict the force ranges used to obtain the probability of each configuration, as shown with the dashed lines.

C.3 Occupant variable distributions

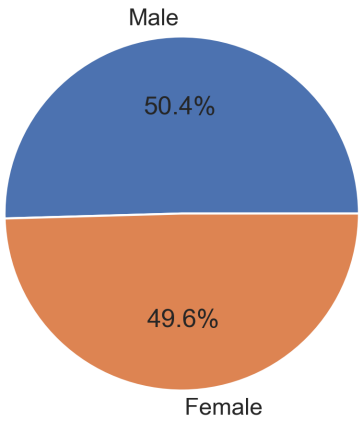


Figure C.13: Distribution of the occupant's sex in the reference dataset.

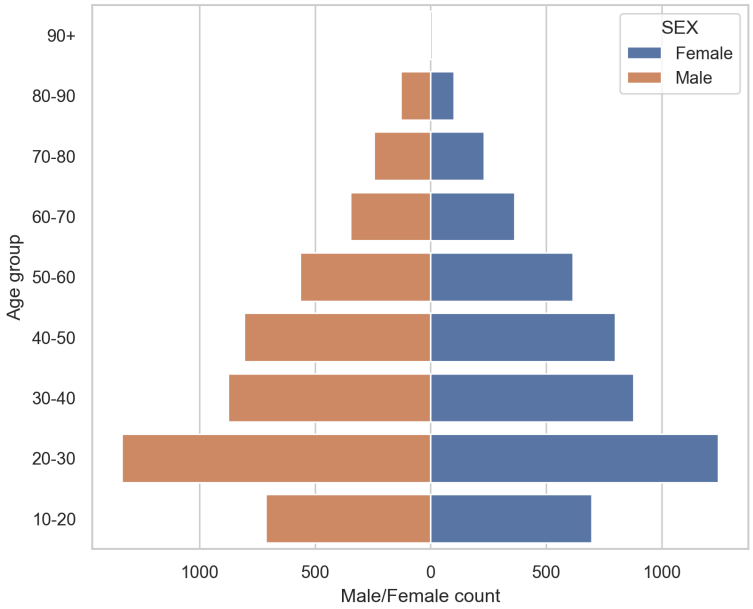


Figure C.14: Population distribution in the reference dataset (drivers in NASS CDS crash database from 2000 to 2015).

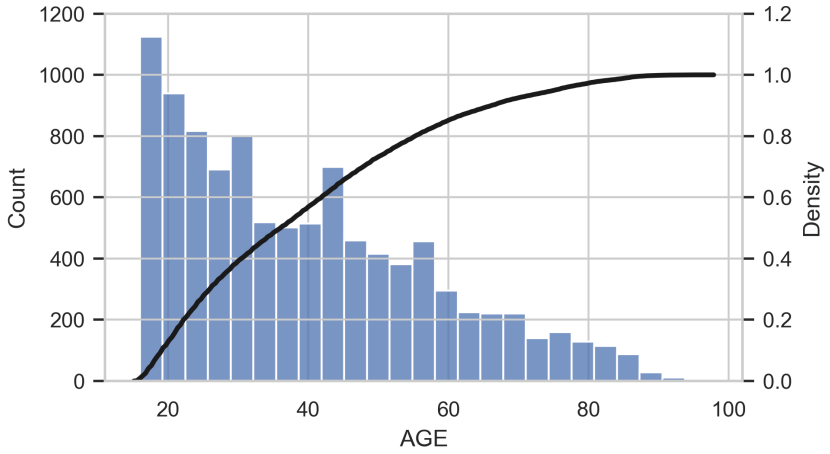


Figure C.15: Distribution of the occupant's ages in the reference dataset.

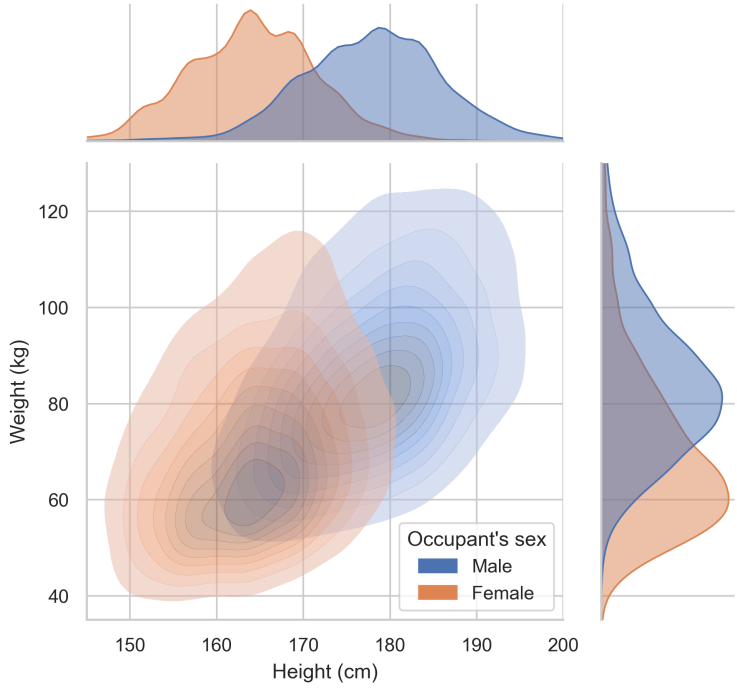


Figure C.16: Distribution of occupants' height and weight in the reference dataset, illustrated using kernel density estimation (KDE) for both univariate (height and weight individually) and bivariate (height and weight combined) analyses, categorized by occupants' sex. Shadowed areas in the bivariate plot (centre plot) represent 90 % of the population in the reference data.

C.4 Injury distributions

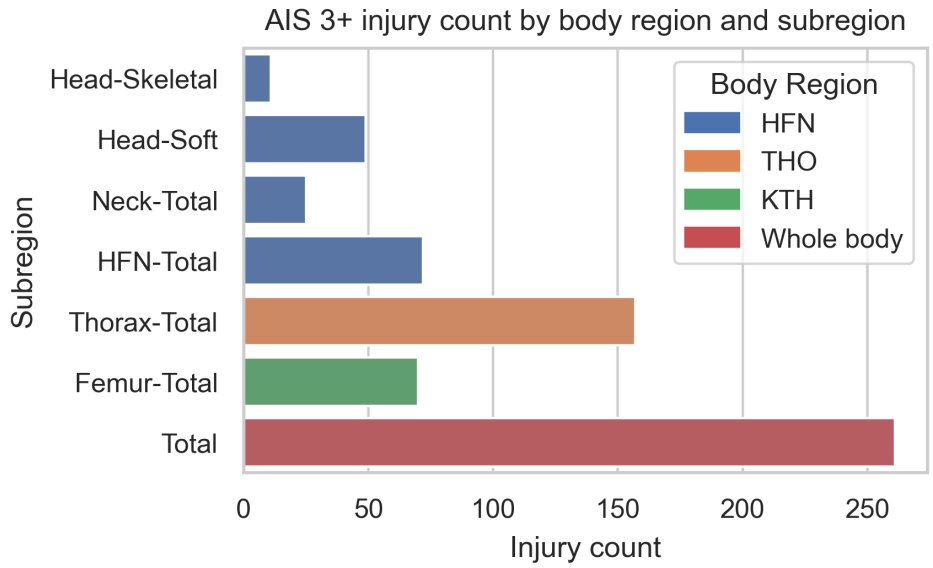


Figure C.17: Distribution of the occupant's injuries in the reference dataset used in Chapter 3.

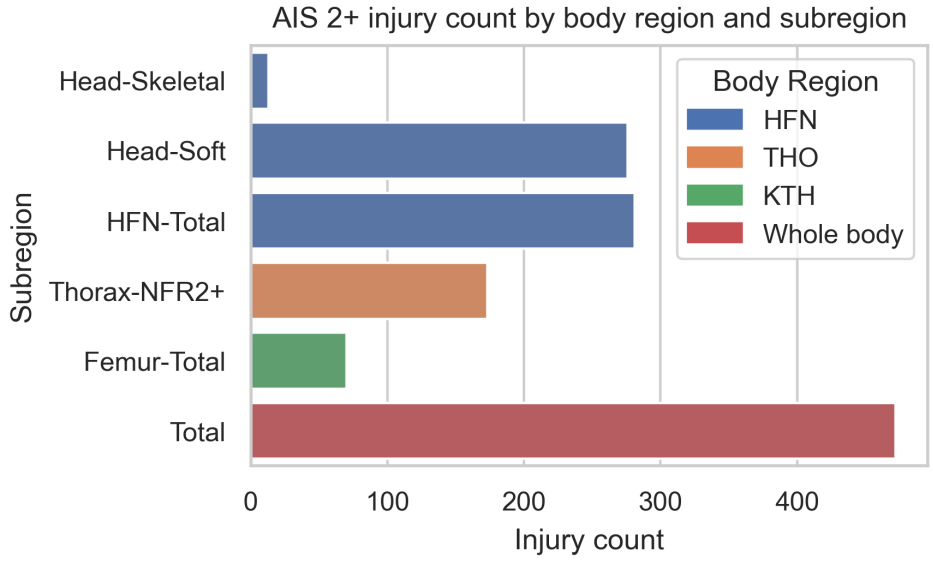


Figure C.18: Distribution of the occupant's injuries in the reference dataset used in Chapter 4.

D Metamodel obtention using multibody model: Supplementary information

D.1 Metamodel training process

Table D.1: List of metamodels hyperparameters

Metamodel	Hyperparameter	Range of used values
LASSO	Regularization parameter (λ)	$\lambda \in [1e-4, 1e-4]$
Neural Network	Transfer function	Hyperbolic Tangent Function,
	Number of nodes of hidden layer (N_n)	$N_n \in [4, 30]$
	Weight decay	$Wd \in [1e-4, 1e-2]$
	Learning rate	$Lr \in [1e-3, 1.5e-2]$
Support Vector Regression	Kernel function	2nd Order Polynomial
	Box constraint (C)	$C \in [1e-2, 1e3]$
	Margin width (ε)	$\varepsilon \in [1e-3, 1e1]$
Gaussian Process Regression	Kernel	Radial basis function, Radial basis function + White kernel

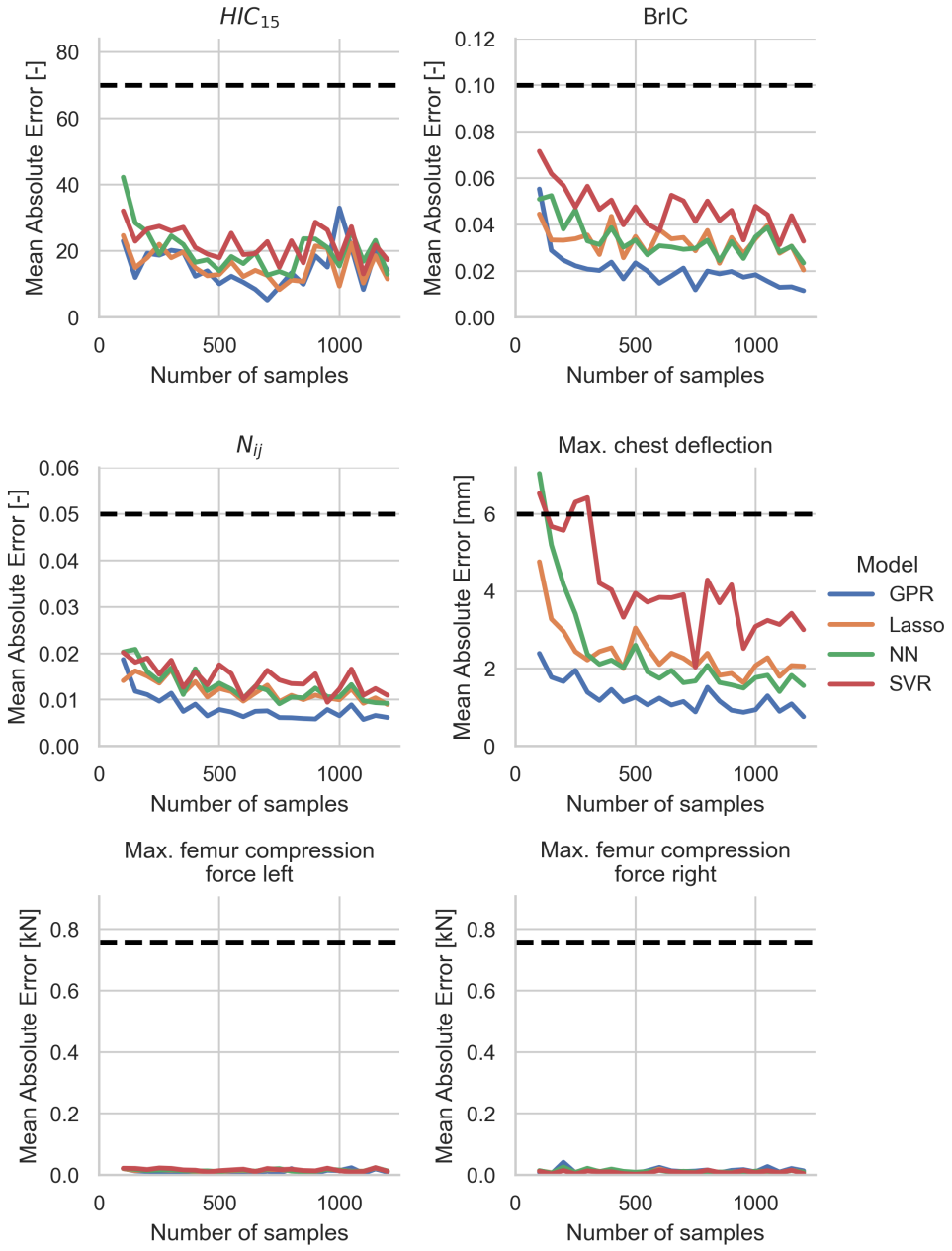


Figure D.1: Mean absolute error (MAE) obtained for each injury-criterion metamodel using LASS, SVR, NN, and GPR. The horizontal axis shows the number of samples used to train the metamodel, and the vertical shows the MAE of the predictions using the metamodel. The dashed black line displays the MAE values that have an error of 10 % of the threshold criteria value.

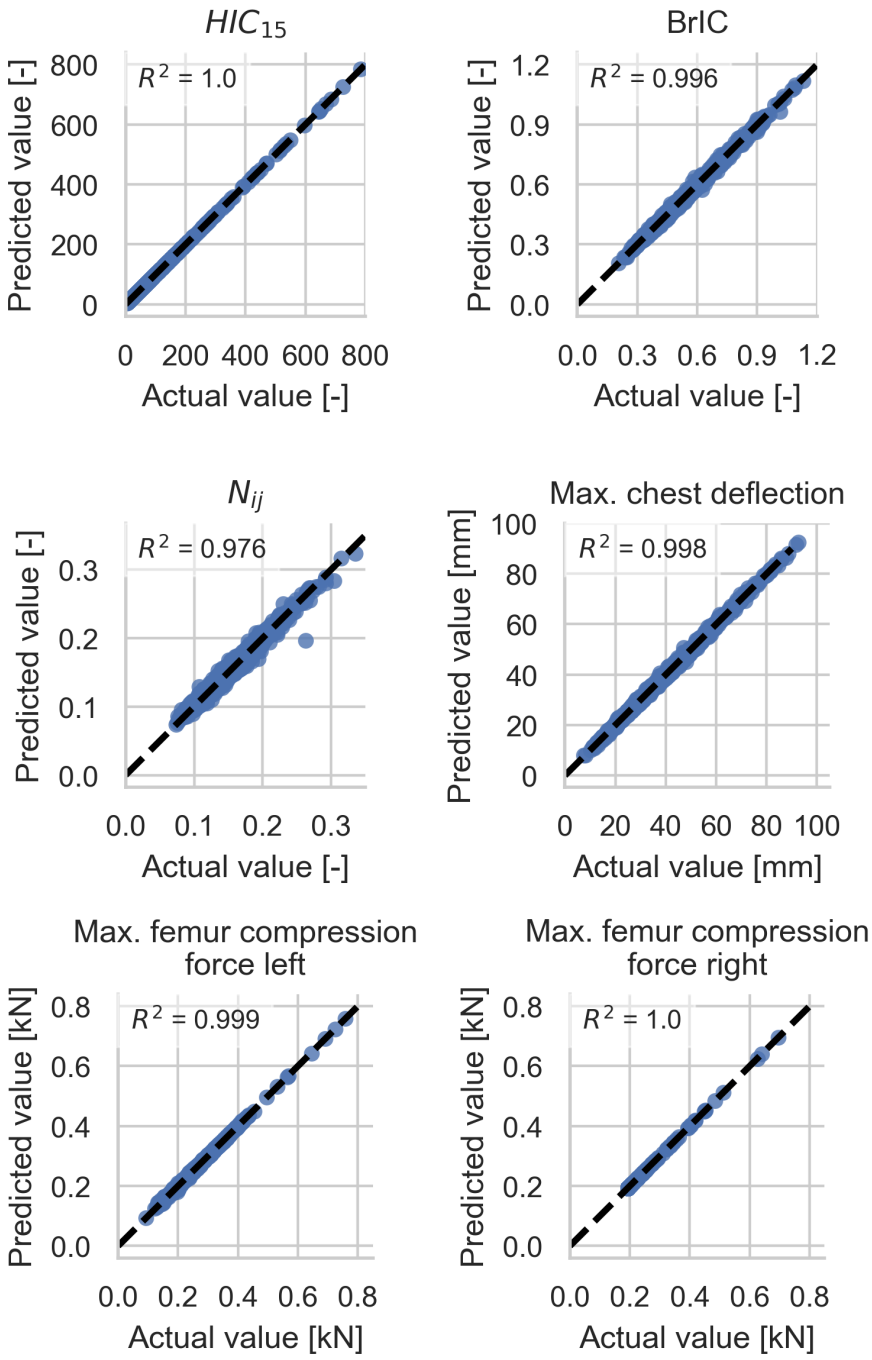


Figure D.2: Acual vs predicted values of the 1,200 simulations used to obtain the GPR model.

D.2 Injury prediction

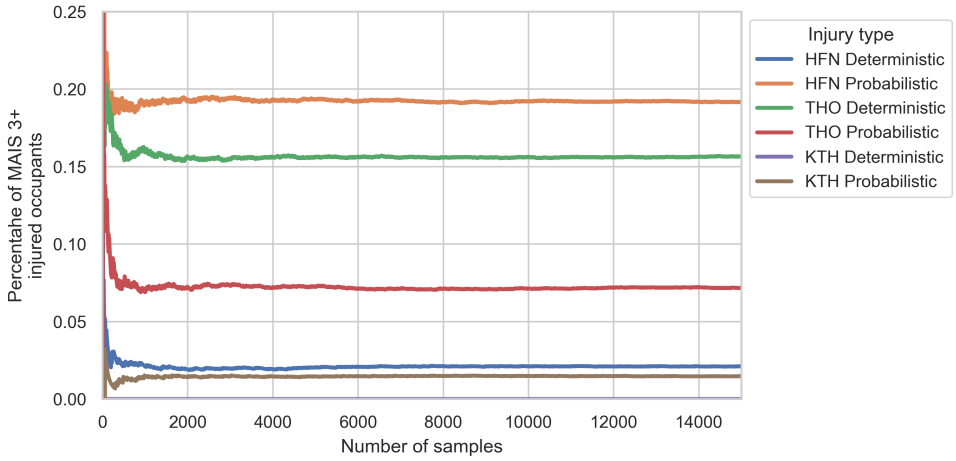


Figure D.3: Percentage of seriously injured occupants predicted using the deterministic and probabilistic method for different sample sizes.

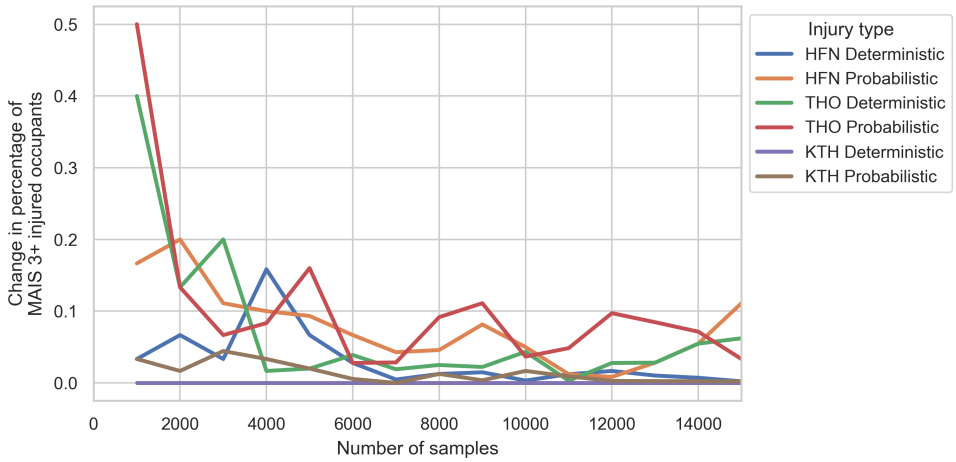


Figure D.4: Change in the percentage of seriously injured occupants predicted using the deterministic and probabilistic method for different sample sizes. The change was measured with the number of samples shown in the plot and 1.5 times its size.

E Generic vehicle interior model: changes and validation

E.1 Model description

This appendix presents the validation of the frontal impact generic vehicle interior sled model (GVI model). A comparison of the simulation results and physical crash test data is presented under matching impact conditions (i.e., restraint system configuration). Figure E.1 shows a brief summary of the methodology applied to validate the GVI model changes. The driver environment in this model was derived from the driver-side vehicle model developed by Iraeus et al. (2016), except for the belt system and seat. This generic model represents an average vehicle interior geometry based on scans and measurements from 14 vehicles.

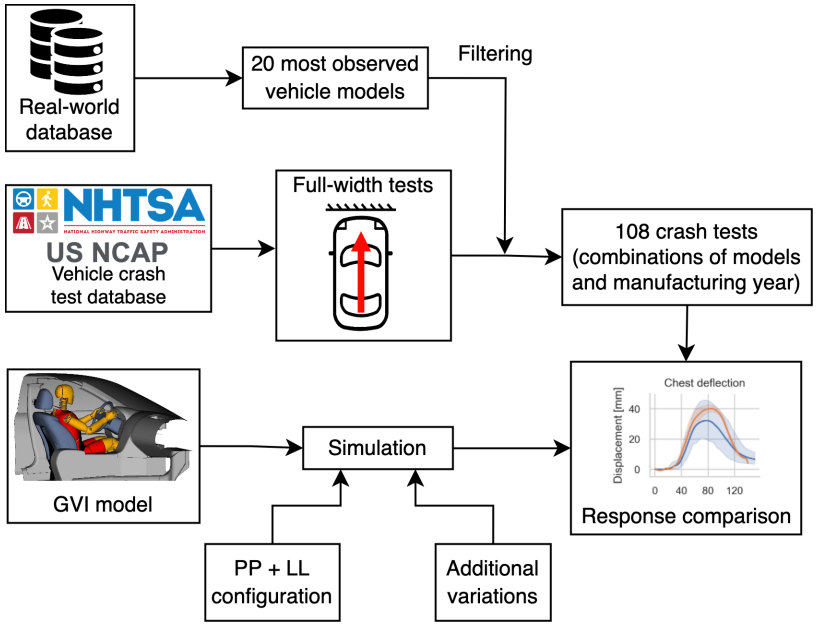


Figure E.1: Flowchart illustrating the methodology employed to validate the GVI model response using crash test data.

The open access front seat model (*FE_models / Open Access Front Seat Model · GitLab* 2024) replaced the vehicle's seat. This seat has a measured h-point, which was used to position the anthropomorphic test device (ATD) for the validation and to position the morphed human body models for Chapter 4. Two modifications were implemented to the seat model. First, the initial foam reference geometry keyword was included for those nodes below the seat cushion and backrest to enable the division of the simulation into positioning and crash simulation, retaining the foam-to-occupant contact force. Second, a contact was included between the cushion seat pan and the cushion side members.

The belt system, except for the buckle, was also modified. The retractor and end-bracket models were replaced to enable easy switching between a non-pre-tensioned, non-load-limited restraint system and a system featuring double pre-tensioning (retractor and end-bracket) with a double-stage load-limiting device, as well as all intermediate combinations. If a second stage of the load-limiter was present, the switch between stages was time-dependent, defined as the time-to-fire of the pre-tensioner and airbag plus an additional 50 ms. Bending beams (Dahlgren et al., 2020) were also included in the lap and shoulder belt webbing to avoid excessive belt bending.

E.2 Model setup and validation method

The original development of the GVI model by Iraeus et al. (2016) used the crash tests of 14 vehicles to calibrate and validate the model response. An objective rating method was applied to assess the degree of similarity between the model and the tests using the femur force, head x-acceleration, chest x-acceleration, and chest deflection. The validation of the modified GVI model was expanded to include a larger set of vehicles and take into account the specific configuration of the restraint system, including the presence of pre-tensioning and load-limiting devices.

The initial development of the GVI model by Iraeus et al. (2016) was based on crash tests involving 14 vehicles, which were used to calibrate and validate the model's performance. To evaluate the degree of similarity between the model predictions and experimental results, an objective rating method was employed, focusing on the following metrics: femur force, head x-acceleration, chest x-acceleration, and chest deflection. The validation process for the modified GVI model was subsequently broadened to encompass a larger dataset of vehicles. Additionally, this expanded validation incorporated specific configurations of the restraint systems, including features like pre-tensioners and load limiters.

To ensure the model more accurately reflects real-world conditions, vehicle models most frequently involved in collisions, identified from Chapter 2 reference data, were chosen for the validation of the GVI model. These crash tests included variations in the vehicles' manufacturing years, providing a more comprehensive assessment of the model's performance across different variations in the restraint system.

Three different restraint system configurations were selected for the validation process. These configurations were defined based on the characteristics of the pre-tensioning and load-limiting devices, which included: no pre-tensioning nor load-limiting device, pre-tensioning in the retractor and load-limiting force lower than 4.5 kN, and pre-tensioning in the retractor and load-limiting force higher than 4.5 kN. Head, chest, pelvis accelerations, chest deflection, neck forces and moments, femur forces, and seatbelt forces were used to compare the crash tests and the model response for each combination of the restraint system.

Three different restraint system configurations, varying in pre-tensioning and load-limiting devices, were selected for the validation process. These configurations included: no pre-tensioning nor load-limiting device; pre-tensioning in the retractor with a load-limiting force below 4.5 kN; and pre-tensioning in the retractor with a load-limiting force above 4.5 kN. To assess the model's accuracy for each restraint system combination, time-history signals from the head, chest, and pelvis accelerations; chest deflection; neck forces and moments; femur forces; and seatbelt forces were compared between the crash tests and the model's response.

A visual comparison of the time-history measurements was performed, providing an assessment of model performance. This comparison involved calculating the mean response along with the 5^{th} and 95^{th} percentiles of the time-history measurements at each time point.

Additionally, variations in the knee bolster, airbag, and steering column characteristics were incorporated into the GVI model when determining the response corridors.

E.2.1 Crash tests reference data

The crash tests used to validate the GVI model response were selected from the National Highway Traffic Safety Administration (NHTSA) frontal impact rigid barrier crash tests, conducted at 35 mph (56 km/h). The vehicle models chosen for these tests were based on those most frequently involved in crashes, as identified from the reference crash dataset in Chapter 2. The selection process for these crash tests included the following steps:

1. The dataset from Chapter 2 was reviewed to identify the most commonly involved vehicles in crashes.
2. The top 20 most frequent vehicle models, representing nearly 40 % of the original dataset, were selected for further analysis.
3. Data from NHTSA frontal impact rigid barrier crash tests for these 20 vehicle models, covering production years from 1990 to 2015, were retrieved.
4. Characteristics of the vehicle restraint systems from Subchapter 2.3.1 were manually reviewed, to prevent false positives or misclassification errors.
5. The ATDs used in the crash tests were examined, with cases involving a Hybrid III 50^{th} percentile male ATD in the driver's seat selected for validation.
6. Signals from each crash test were inspected to detect any issues, such as inverse polarity or missing data, and any problematic signals were excluded from the validation dataset.

This selection process ensured that the crash tests used for validation provided accurate and reliable data for assessing the GVI model's performance. A total of 108 crash tests were obtained from this process, which included: 64 crash tests with no pre-tensioning nor load-limiting device, 25 crash tests with pre-tensioning in the retractor and load-limiting force lower than 4.5 kN, and 19 crash tests with pre-tensioning in the retractor and load-limiting force higher than 4.5 kN. Table E.1 shows a description of these vehicles ordered from most to less frequently observed vehicle models in the crash database from Chapter 2.

Table E.1: Retraint system description for the vehicles used in the interior vehicle sled model validation process.

Make	Model	Model year	Test number	PP force LAP [kN]	PP force SHLD [kN]	LL 1st stage [kN]	LL 2nd stage [kN]
Honda	Accord	1994	2048	0.0	0.0	0.0	0.0
Honda	Accord	1997	2475	0.0	0.0	0.0	0.0
Honda	Accord	1998	2712	0.0	0.0	0.0	0.0
Honda	Accord	2000	3188	0.0	0.0	0.0	0.0
Honda	Accord	2001	3457	0.0	1.1	4.9	0.0
Honda	Accord	2008	6257	0.0	1.9	3.8	0.0
Honda	Accord	2011	7078	0.0	1.8	4.9	0.0

Table E.1 continued from previous page

Make	Model	Model year	Test number	PP force LAP [kN]	PP force SHLD [kN]	LL 1st stage [kN]	LL 2nd stage [kN]
Honda	Civic	1994	2066	0.0	0.0	0.0	0.0
Honda	Civic	1996	2371	0.0	0.0	0.0	0.0
Honda	Civic	1998	2735	0.0	0.0	0.0	0.0
Honda	Civic	1999	2969	0.0	0.0	0.0	0.0
Honda	Civic	2001	3458	0.0	2.3	5.9	0.0
Honda	Civic	2006	5571	0.0	2.1	5.6	4.3
Honda	Civic	2011	7147	0.0	2.1	5.3	0.0
Ford	Taurus	1996	2312	0.0	0.0	0.0	0.0
Ford	Taurus	1998	2748	0.0	0.0	0.0	0.0
Ford	Taurus	2000	3248	0.0	0.9	4.6	0.0
Ford	Taurus	2010	6808	0.0	1.3	4.5	2.5
Ford	Taurus	2011	6964	0.0	1.7	5.5	3.9
Toyota	Camry	1992	1690	0.0	0.0	0.0	0.0
Toyota	Camry	1994	2038	0.0	0.0	0.0	0.0
Toyota	Camry	1997	2531	0.0	0.0	0.0	0.0
Toyota	Camry	1998	2710	0.0	0.0	0.0	0.0
Toyota	Camry	2000	3251	0.0	0.0	0.0	0.0
Toyota	Camry	2007	5675	0.0	1.8	4.0	0.0
Toyota	Camry	2011	6953	0.0	1.1	3.6	0.0
Toyota	Corolla	1994	2034	0.0	0.0	0.0	0.0
Toyota	Corolla	1998	2726	0.0	1.2	4.8	0.0
Toyota	Corolla	2003	4266	0.0	1.8	4.4	0.0
Ford	Explorer	1995	2211	0.0	0.0	0.0	0.0
Ford	Explorer	2002	4223	0.0	0.0	0.0	0.0
Ford	Explorer	2006	5561	0.0	1.1	4.8	3.1
Ford	Explorer	2012	7495	1.3	2.2	2.9	2.7
Ford	F-150	1994	2055	0.0	0.0	0.0	0.0
Ford	F-150	1997	2452	0.0	0.0	0.0	0.0
Ford	F-150	1998	2747	0.0	0.0	0.0	0.0
Ford	F-150	1999	3046	0.0	0.0	0.0	0.0
Ford	F-150	2001	3672	0.0	1.4	4.2	0.0
Ford	F-150	2009	6598	0.0	1.7	3.3	0.0
Ford	F-150	2012	7625	0.0	1.8	3.6	0.0
Chevrolet	Cavalier	1995	2253	0.0	0.0	0.0	0.0
Chevrolet	Cavalier	1997	2754	0.0	0.0	0.0	0.0
Chevrolet	Cavalier	1998	2688	0.0	0.0	0.0	0.0
Chevrolet	Cavalier	2003	4445	0.0	0.0	0.0	0.0
Nissan	Altima	1994	2059	0.0	0.0	0.0	0.0
Nissan	Altima	1995	2297	0.0	0.0	0.0	0.0
Nissan	Altima	1998	2744	0.0	0.0	0.0	0.0
Nissan	Altima	1999	3003	0.0	0.0	0.0	0.0
Nissan	Altima	2000	3281	0.0	1.4	5.5	0.0
Nissan	Altima	2005	5150	0.0	1.6	6.4	0.0

Table E.1 continued from previous page

Make	Model	Model year	Test number	PP force LAP [kN]	PP force SHLD [kN]	LL 1st stage [kN]	LL 2nd stage [kN]
Nissan	Altima	2007	5895	0.0	1.7	4.3	0.0
Nissan	Altima	2008	6191	0.0	1.4	3.9	0.0
Nissan	Altima	2011	7152	0.0	1.3	5.1	4.1
Nissan	Altima	2013	7966	1.6	2.1	3.9	0.0
Chevrolet	Silverado	2007	5877	0.0	1.0	5.0	0.0
Chevrolet	Silverado	2011	7099	0.0	1.0	3.1	0.0
Chevrolet	Silverado	2012	7582	0.0	1.1	0.0	0.0
Chevrolet	Silverado	2013	8026	0.0	0.0	0.0	0.0
Chevrolet	Silverado	2015	8604	2.2	2.3	0.0	0.0
Ford	Ranger	1995	2207	0.0	0.0	0.0	0.0
Ford	Ranger	1996	2457	0.0	0.0	0.0	0.0
Ford	Ranger	1999	3058	0.0	0.0	0.0	0.0
Ford	Ranger	2007	5758	0.0	1.2	4.8	0.0
Ford	Focus	2008	6256	1.5	1.2	3.1	0.0
Jeep	Grand Cherokee	1996	2430	0.0	0.0	0.0 .	0.0 .
Jeep	Grand Cherokee	1998	2713	0.0	0.0	0.0 .	0.0 .
Jeep	Grand Cherokee	1999	3057	0.0	0.0	0.0 .	0.0 .
Jeep	Grand Cherokee	2011	6999	0.0	1.7	4.3 .	0.0 .
Jeep	Grand Cherokee	2014	8293	0.0	1.3	2.9 .	0.0 .
Chevrolet	Malibu	1997	2529	0.0	0.0	0.0	0.0
Chevrolet	Malibu	1998	2714	0.0	0.0	0.0	0.0
Chevrolet	Malibu	2001	3566	0.0	0.0	0.0	0.0
Chevrolet	Malibu	2011	6998	0.0	1.6	3.3	0.0
Chevrolet	Malibu	2013	7856	0.0	1.6	2.4	0.0
Chevrolet	Impala	2000	3130	0.0	0.0	4.2	0.0
Chevrolet	Impala	2001	3471	0.0	0.0	0.0	0.0
Chevrolet	Impala	2006	5547	0.0	1.0	4.8	0.0
Chevrolet	Impala	2012	7488	1.4	1.0	3.3	0.0
Chevrolet	Impala	2014	8290	0.0	2.2	2.8	0.0
Ford	Mustang	1994	2063	0.0	0.0	0.0	0.0
Ford	Mustang	1996	2360	0.0	0.0	0.0	0.0
Ford	Mustang	1998	2806	0.0	0.0	0.0	0.0
Ford	Mustang	1999	3016	0.0	0.0	0.0	0.0
Ford	Mustang	2001	3542	0.0	0.0	0.0	0.0
Ford	Mustang	2005	5259	0.0	1.4	3.7	0.0
Ford	Mustang	2008	6184	0.0	1.6	4.1	0.0
Dodge	Neon	1996	2320	0.0	0.0	0.0	0.0
Dodge	Neon	1998	2709	0.0	0.0	0.0	0.0

Table E.1 continued from previous page

Make	Model	Model year	Test number	PP force LAP [kN]	PP force SHLD [kN]	LL 1st stage [kN]	LL 2nd stage [kN]
Dodge	Neon	2000	3207	0.0	0.0	0.0	0.0
Dodge	Ram	1994	2021	0.0	0.0	0.0	0.0
Dodge	Ram	1997	2530	0.0	0.0	0.0	0.0
Dodge	Ram	1998	2784	0.0	0.0	0.0	0.0
Dodge	Ram	1999	3023	0.0	0.0	0.0	0.0
Dodge	Ram	2001	3532	0.0	0.0	0.0	0.0
Dodge	Ram	2005	5159	0.0	0.9	4.2	0.0
Dodge	Ram	2009	6585	0.0	1.3	4.1	0.0
Dodge	Ram	2011	7115	0.0	1.6	4.8	0.0
Nissan	Sentra	1995	2298	0.0	0.0	0.0	0.0
Nissan	Sentra	1998	2771	0.0	0.0	0.0	0.0
Nissan	Sentra	2001	3562	0.0	0.0	5.8	0.0
Nissan	Sentra	2007	5876	0.0	1.8	4.5	3.0
Nissan	Sentra	2011	7131	0.0	1.4	5.8	3.2
Chevrolet	Blazer	1991	1539	0.0	0.0	0.0	0.0
Chevrolet	Blazer	1993	1798	0.0	0.0	0.0	0.0
Chevrolet	Blazer	1997	2478	0.0	0.0	0.0	0.0
Chevrolet	Blazer	1998	2756	0.0	0.0	0.0	0.0
Chevrolet	Blazer	1999	3032	0.0	0.0	0.0	0.0
Chevrolet	Blazer	2002	3901	0.0	0.0	0.0	0.0

E.2.2 Model setup

The GVI model contains a wide range of parameters that can be modified to personalize its behaviour. Most of these parameters were fixed in the obtention of the validation setup, such as friction coefficients (equal to 0.3), position of the steering wheel, dashboard and floor intrusion (null for all simulations).

A Hybrid III 50th percentile male ATD LS-Dyna model (ATD-H350-D01.11 from ATD-Models GmbH, Weißwasser, Germany) was seated in a driver posture. The h-point position was based on the position of a 45-year-old, 78 kg and 179 cm male driver using Park et al. (2016). A position simulation was performed to place the ATD into the seat and reach the steering wheel and the pedals with the hands and feet, respectively.

The average acceleration pulse from the reference crash tests was calculated and scaled to obtain a 56 mk/h delta-v. This acceleration pulse was then used to carry out all the simulations.

Six parameters were modified to obtain all the setups used to validate the model's response. The primary variations involved the pre-tensioning and load-limiting devices (i.e., two of the six parameters). Three seatbelt configurations were employed in the simulations:

1. The first configuration did not include pre-tensioning or load-limiting devices, but the airbag was deployed at 10 ms. These cases were compared with those vehicles that did not include these devices in the crash tests.

2. The second configuration included a retractor pre-tensioning device and a "low force" load-limiting device, with a load-limiting force of 4 kN. The pre-tensioner and airbag were activated at 10 ms. These cases were compared with those vehicles that included load-limiting force below 4.5 kN.

3. The third configuration included a retractor pre-tensioning device and a "high force" load-limiting device, with a load-limiting force of 5 kN. The pre-tensioner and airbag were activated at 10 ms. These cases were compared with those vehicles that included load-limiting force above 4.5 kN.

The pre-tensioning force applied by the retractor was 2 kN.

The other four parameters used in this validation process were the size and pressure of the airbag, the knee bolster stiffness, and the steering wheel compression force. Two values were selected for each parameter, resulting in a small or large force, size, or stiffness. This selection was based on the distributions of these parameters defined in the appendix of Iraeus et al. (2016). These distributions followed normal distributions, and the mean (μ) and standard deviation (σ) were used to define the parameters "small" and "large" values. Equation E.1 was used to obtain the small value, and Equation E.2 was used to obtain the large value. Table E.2 shows the results of applying both equations to each distribution.

$$Small\_value = \mu - \sigma \tag{E.1}$$

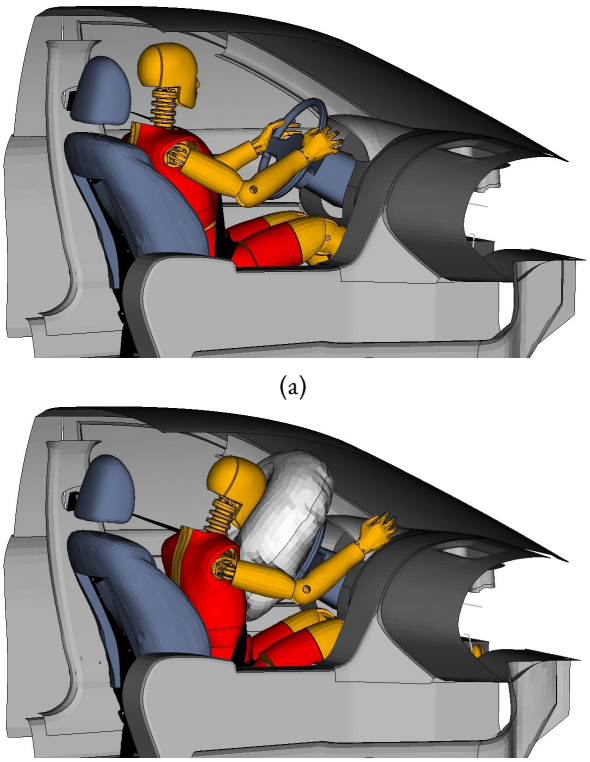
$$Large\_value = \mu + \sigma \tag{E.2}$$

Table E.2: Parameters used to obtain the validation GVI model setup.

Parameter	Distribution		Value		Unit
	Mean value	Standard deviation	small	large	
Airbag size	332	18	314	350	mm
Airbag pressure	1.31	0.1	1.21	1.41	
Knee bolster stiffness [a]	4.72	1.93	2.79	6.65	kN/mm
Steering wheel compression force	4.8	0.9	3.9	5.7	kN

[a] Stiffness in each leg contact.

The final combination of parameters used in the validation process was obtained through a full factorial design. This approach resulted in 48 simulations, as three seatbelt configurations were evaluated using two possible values for the other four parameters. Figure E.2 shows the Hybrid III 50th percentile male ATD model positioned in the GVI model at the start of the crash (Subfigure E.2a) and at 70 ms of the crash (Subfigure E.2b).



(a)

(b)

Figure E.2: GVI model in the 56 km/h frontal impact test simulation. Subfigure (a) shows the Hybrid III 50th percentile male ATD model in the GVI model at time 0 ms of the crash. Subfigure (b) shows the ATD model in the GVI model at time 70 ms of the crash.

E.3 Validation results

The correlation of the model using each seatbelt configuration, including variations in the other parameters, was performed by plotting the signals from several sensors in both physical tests and simulations. For all signals, the mean value and the spread of the data, represented by the 5th and 95th percentiles, were calculated at each point in the time series. Prior to these calculations, the recommended CFC filters by SAE (2007) were applied to all signals. The signals used for validation included accelerations of the head, chest, and pelvis along the x and z axes; chest deflection; upper neck forces along the x and z axes; upper neck moment along the y axis; femur forces along the z-axis for both left and right legs; lap and shoulder belt forces; and vehicle acceleration along the x-axis. The vehicle acceleration used in the simulation was defined as the mean vehicle acceleration along the x-axis for the 108 crash tests, with a CFC 60 filter applied to obtain a smoother signal.

Figures E.3, E.4 and E.5 present the results from crash tests and simulations categorized by seatbelt configurations: without pre-tensioning and load-limiting devices, with pre-tensioning and load-limiting force below 4.5 kN, and with pre-tensioning and load-limiting force above 4.5 kN. These configurations also included variations in airbag size and pressure, knee bolster stiffness, and steering wheel compression force. The mean value of the time series is depicted with a bold line, while the spread, represented by the 5th and 95th percentiles, is shown with a shadowed area. Crash test results and simulation results are displayed in blue and orange, respectively.

The simulation results from the GVI model, in terms of magnitude, phase, and shape, generally aligned with the range of crash test results. Two indicators were used to evaluate these results: differences in mean value over time and the spread of time series data. For the first indicator, the simulations exhibited similar magnitude, phase, and shape of the signals compared to the mean time series from the crash tests, with any differences observed remaining within the spread of the crash test time series. For the second indicator, the spread of the simulation time series was smaller than that observed in the crash tests. The range defined by the 5^{th} and 95^{th} percentiles of the simulation time series typically fell within the spread of the crash test results. The main discrepancies were identified in the following signals:

1. Head x-axis acceleration: the magnitude of the signals was similar to the mean test response, but the peak value was observed 10 ms later than the ones observed in the tests.

2. Chest deflection: the main response was mostly within the spread of the crash tests. Although the main peak deflection in the simulation was 10 mm higher than in the crash tests for each belt configuration, the sensibility when changing the seat belt configuration was similar, as the differences in the peak chest deflection were similar.

3. Lap belt force: the shape and phase of this signal were similar to the crash test. However, the magnitude of the peak was under-predicted by 2 to 3 kN.

Differences in the spread of the responses between the crash tests and simulations could be attributed to variations in certain model components, such as seatbelt routing, seat design, airbag model, anchor points or ATD position. Additionally, unaccounted-for parameters like friction between components, intrusion, and acceleration curve shape could also contribute to these differences. Furthermore, the range of variation in these parameters could influence this spread.

Based on these results, it was concluded that the GVI model's response closely aligned with the real-world behaviour of vehicles in frontal impact crash conditions. However, future iterations of the model should aim to better capture the variability observed in crash tests. This can be achieved by increasing the number of simulations and applying more sophisticated experimental designs, such as fractional factorial design or Latin hypercube design, to systematically select parameter combinations for each configuration. Additionally, future work should focus on employing objective correlation metrics to identify which combinations of restraint system characteristics, such as knee bolster stiffness and airbag size, best correlate with specific features, such as load-limiting devices.

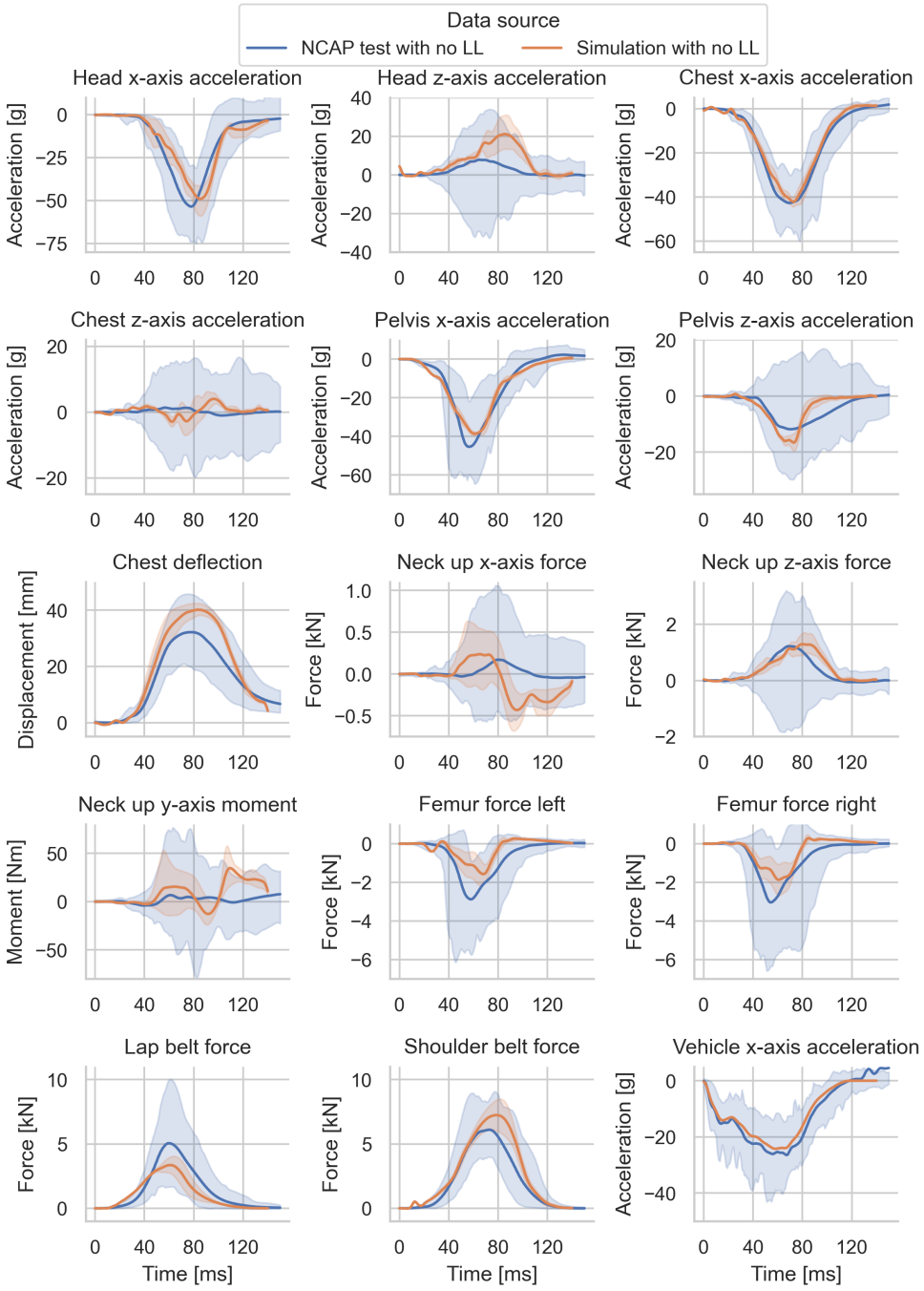


Figure E.3: Hybrid III 50th percentile male ATD results from the GVI model in a frontal impact simulation using a seatbelt without pre-tensioning or load-limiting devices. Physical test results are shown in blue, while simulation results are shown in orange. The shadowed areas represent the 5th and 95th percentiles of the time-series data.

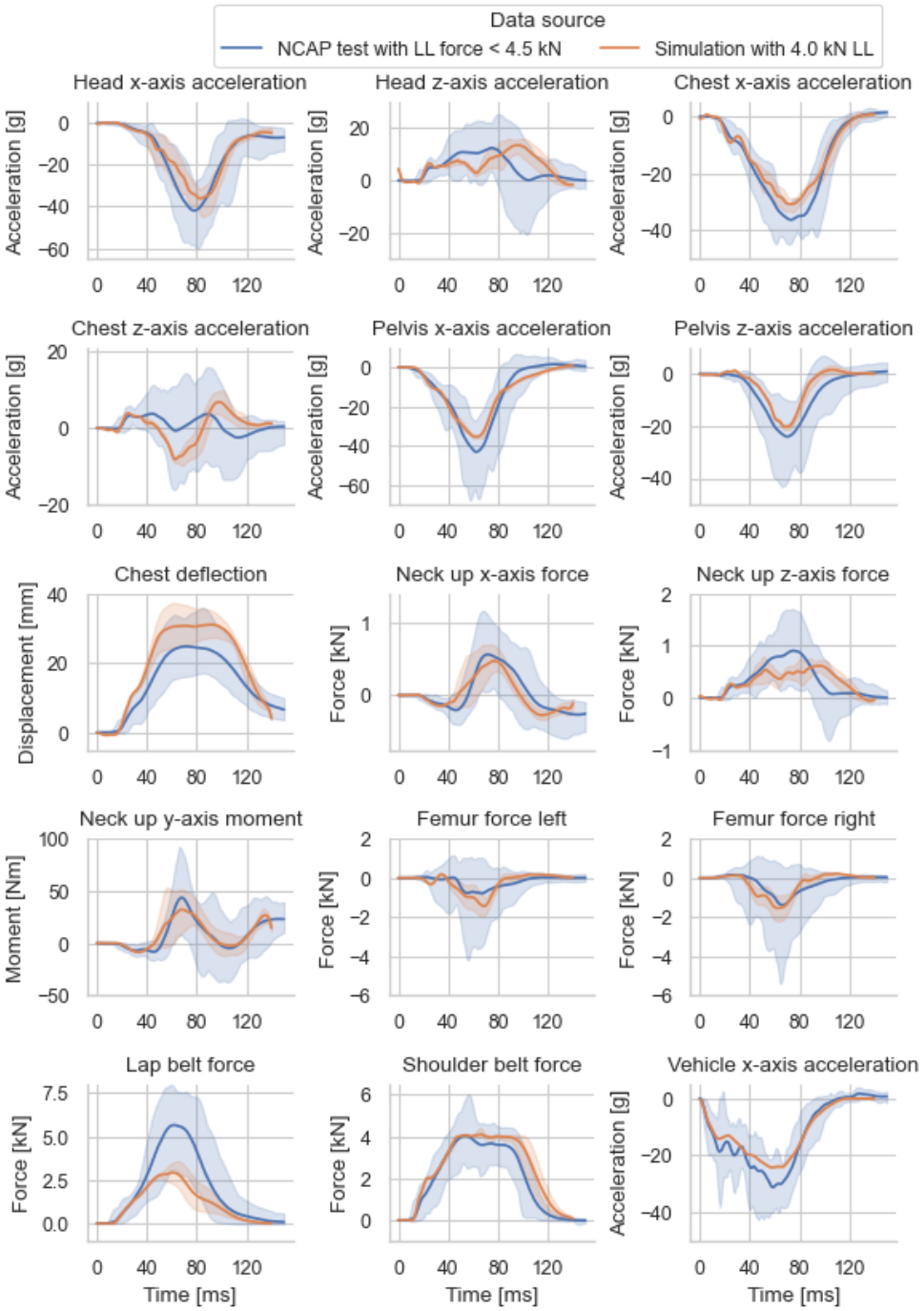


Figure E.4: Hybrid III 50th percentile male ATD results from the GVI model in a frontal impact simulation using a seatbelt with retractor pre-tensioning and 4 kN load-limiting devices. Physical test results are shown in blue, while simulation results are shown in orange. The shadowed areas represent the 5th and 95th percentiles of the time-series data.

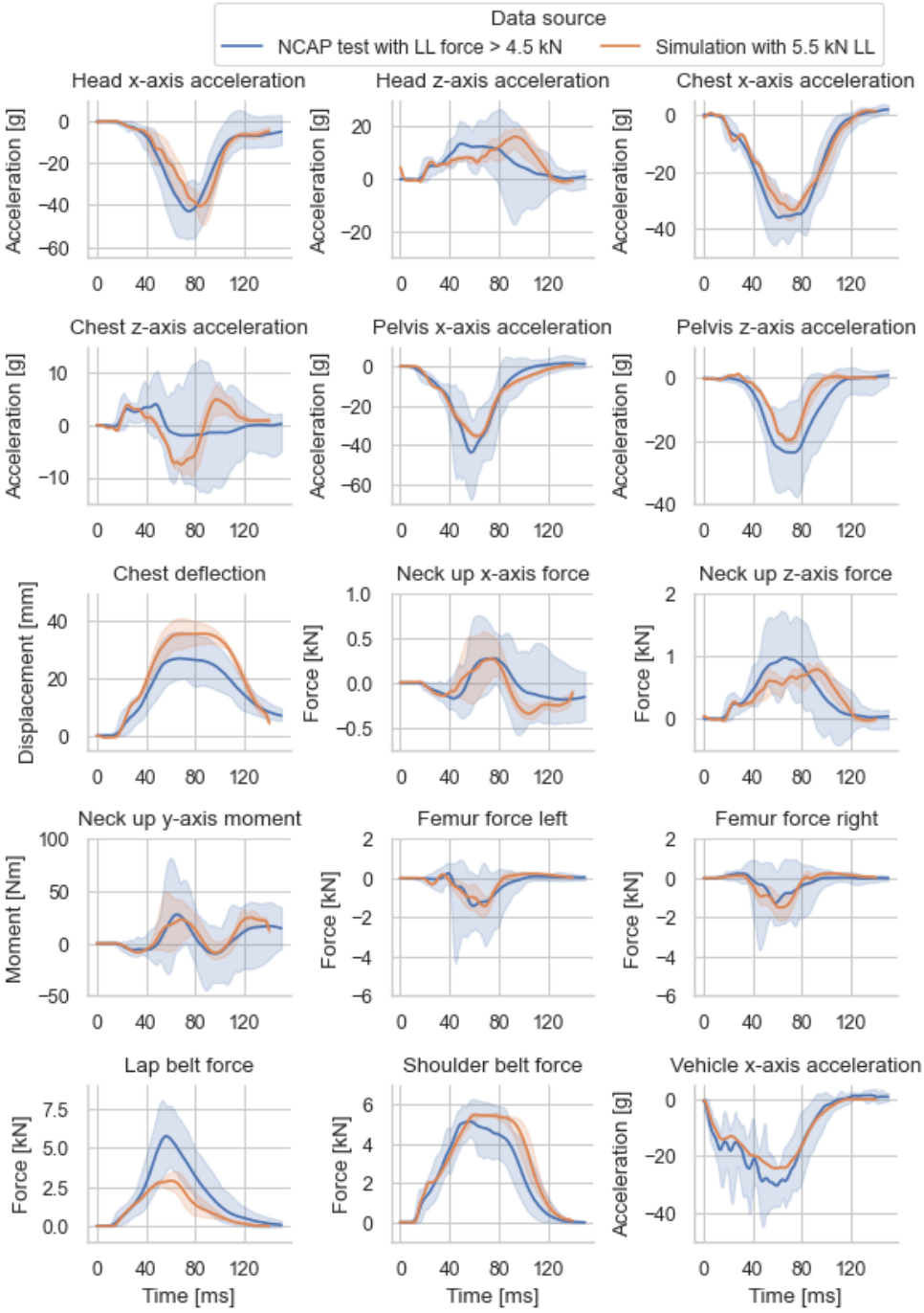


Figure E.5: Hybrid III 50th percentile male ATD results from the GVI model in a frontal impact simulation using a seatbelt with retractor pre-tensioning and 5.5 kN load-limiting devices. Physical test results are shown in blue, while simulation results are shown in orange. The shadowed areas represent the 5th and 95th percentiles of the time-series data.

F Combination of parameters used to morpher the SAFER human body model

Table F.1: Selection of height, weight, sex, age and seating height ratio used to morph female occupants.

ID	Height [cm]	Weight [kg]	Sex	Age [years]	Seating height ratio
F-01	173.8	56.1	Female	45.0	0.52
F-02	168.6	48.6	Female	45.0	0.52
F-03	162.5	44.8	Female	45.0	0.52
F-04	177.9	63.6	Female	45.0	0.52
F-05	172.1	65.8	Female	45.0	0.52
F-06	166.9	56.9	Female	45.0	0.52
F-07	156.6	44.6	Female	45.0	0.52
F-08	177.7	78.0	Female	45.0	0.52
F-09	160.8	52.8	Female	45.0	0.52
F-10	166.9	68.7	Female	45.0	0.52
F-11	171.7	80.5	Female	45.0	0.52
F-12	154.9	52.0	Female	45.0	0.52
F-13	160.8	63.2	Female	45.0	0.52
F-14	175.0	91.0	Female	45.0	0.52
F-15	163.1	76.4	Female	45.0	0.52
F-16	167.9	89.5	Female	45.0	0.52
F-17	154.8	62.3	Female	45.0	0.52
F-18	149.5	49.1	Female	45.0	0.52
F-19	170.2	108.3	Female	45.0	0.52
F-20	156.7	75.3	Female	45.0	0.52
F-21	161.1	89.1	Female	45.0	0.52
F-22	149.2	59.6	Female	45.0	0.52
F-23	164.3	107.3	Female	45.0	0.52
F-24	151.3	73.1	Female	45.0	0.52
F-25	155.7	92.5	Female	45.0	0.52

Table F.2: Selection of height, weight, sex, age and seating height ratio used to morph male occupants.

ID	Height [cm]	Weight [kg]	Sex	Age [years]	Seating height ratio
M-01	187.7	70.9	Male	45.0	0.52
M-02	182.0	62.3	Male	45.0	0.52
M-03	175.5	57.6	Male	45.0	0.52
M-04	192.7	86.5	Male	45.0	0.52
M-05	186.4	80.6	Male	45.0	0.52
M-06	180.6	70.8	Male	45.0	0.52
M-07	169.3	56.8	Male	45.0	0.52
M-08	192.9	103.8	Male	45.0	0.52
M-09	174.1	65.8	Male	45.0	0.52
M-10	181.0	82.8	Male	45.0	0.52
M-11	186.4	95.2	Male	45.0	0.52
M-12	167.9	64.3	Male	45.0	0.52
M-13	174.5	76.4	Male	45.0	0.52
M-14	190.4	115.7	Male	45.0	0.52
M-15	177.4	89.7	Male	45.0	0.52
M-16	182.7	103.3	Male	45.0	0.52
M-17	168.3	74.6	Male	45.0	0.52
M-18	162.4	57.8	Male	45.0	0.52
M-19	185.6	121.5	Male	45.0	0.52
M-20	170.8	87.6	Male	45.0	0.52
M-21	175.7	101.6	Male	45.0	0.52
M-22	161.5	73.0	Male	45.0	0.52
M-23	179.5	119.2	Male	45.0	0.52
M-24	165.2	89.3	Male	45.0	0.52
M-25	170.2	103.7	Male	45.0	0.52

G Metamodel obtention using finite element model: Supplementary information

G.1 Metamodel training process

Table G.1: List of metamodels hyperparameters

Metamodel	Hyperparameter	Range of used values
LASSO	Regularization parameter (λ)	$\lambda \in [1e-4, 1e-4]$
Neural Network	Transfer function	Hyperbolic Tangent Function,
	Number of nodes of hidden layer (N_n)	$N_n \in [4, 30]$
	Weight decay	$Wd \in [1e-4, 1e-2]$
	Learning rate	$Lr \in [1e-3, 1.5e-2]$
Support Vector Regression	Kernel function	2nd Order Polynomial
	Box constraint (C)	$C \in [1e-2, 1e3]$
	Margin width (ε)	$\varepsilon \in [1e-3, 1e1]$
Gaussian Process Regression	Kernel	Radial basis function, Radial basis function + White kernel

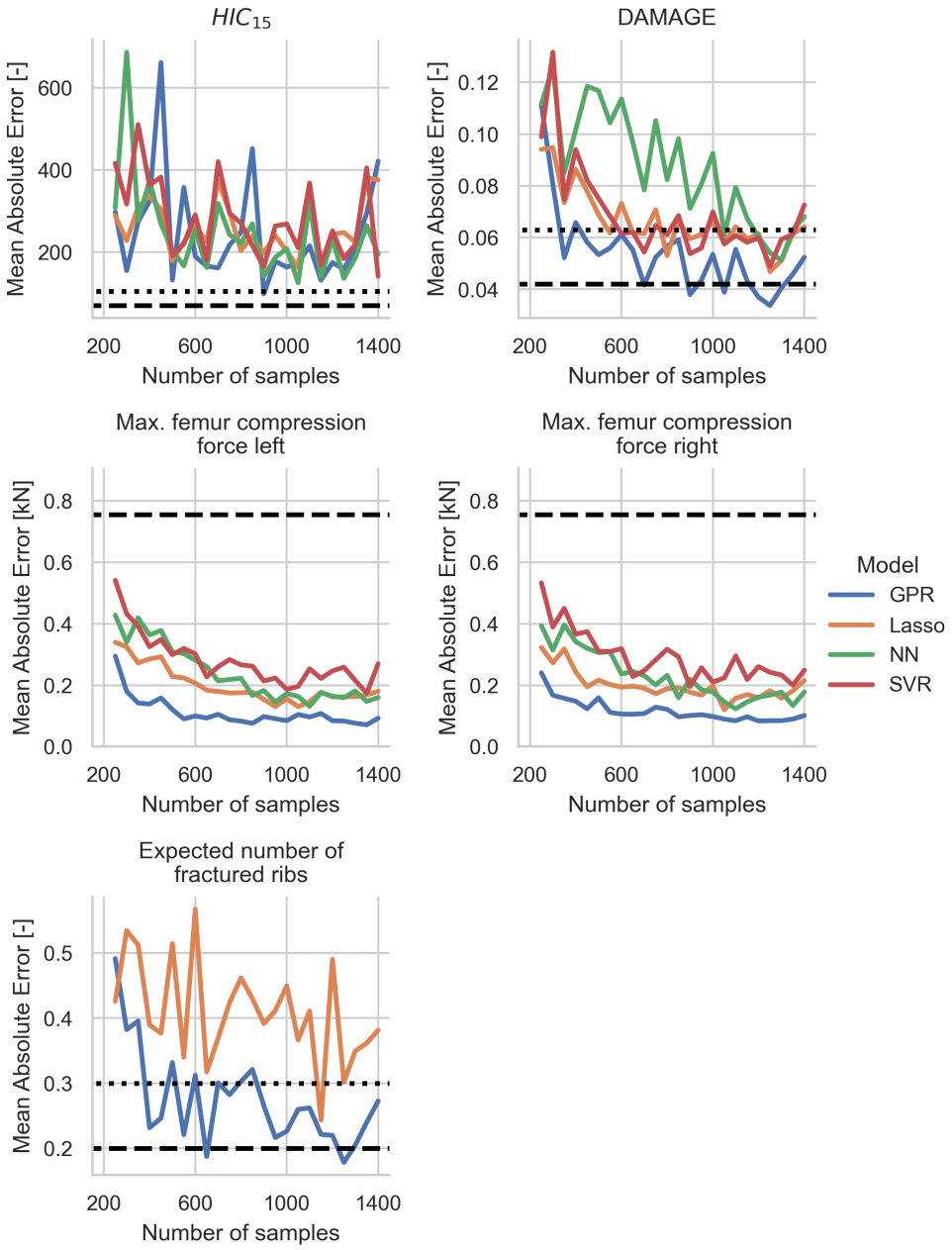


Figure G.1: Mean absolute error (MAE) obtained for each injury-criterion metamodel using LASS, SVR, NN, and GPR. The horizontal axis shows the number of samples used to train the metamodel, and the vertical shows the MAE of the predictions using the metamodel. The dashed black line displays the MAE values that have an error of 10 % of the threshold criteria value ("good" magnitude of prediction error), and the dotted black line displays the MAE values that have an error of 15 % of the threshold criteria value ("acceptable" magnitude of prediction error).

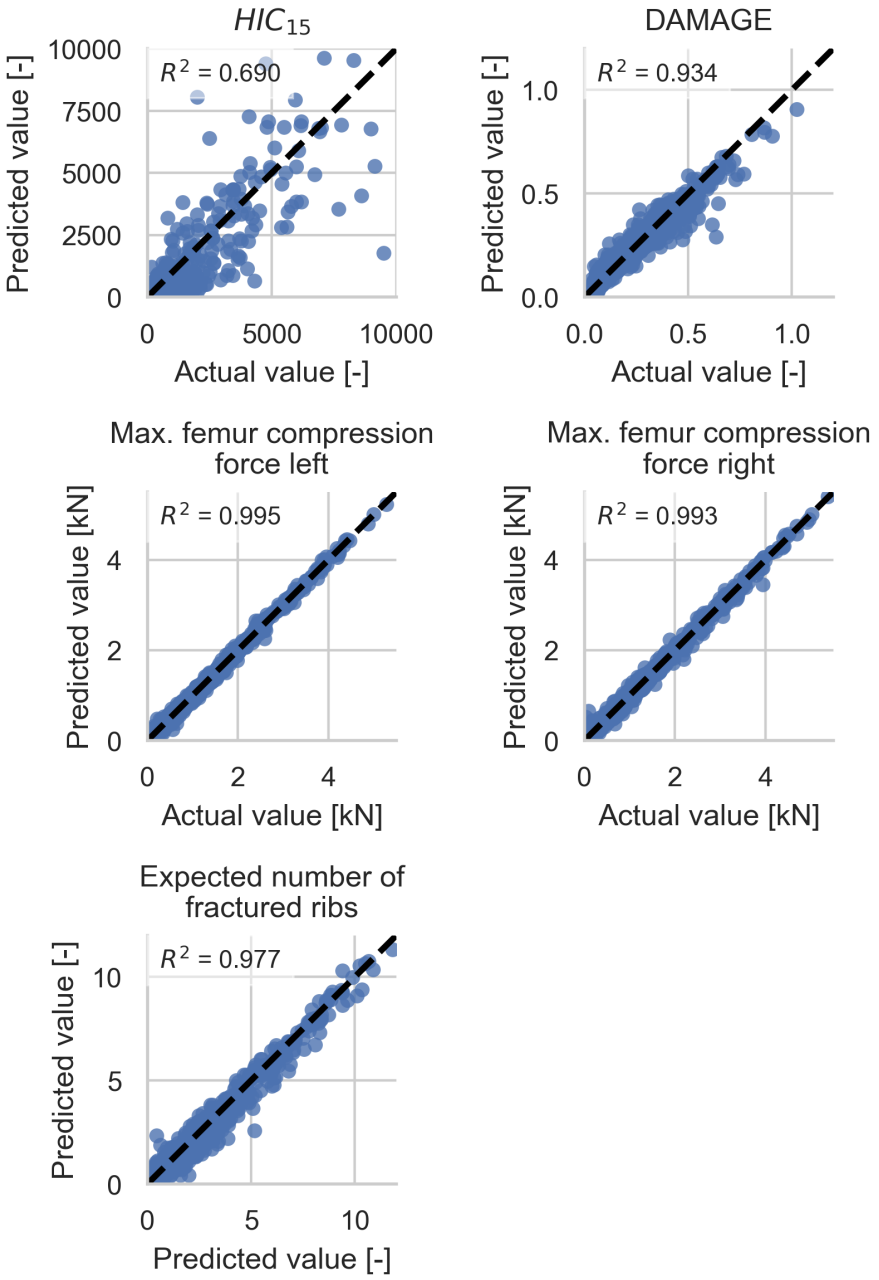


Figure G.2: Acual vs predicted values of the 1,400 simulations used to obtain the GPR model.

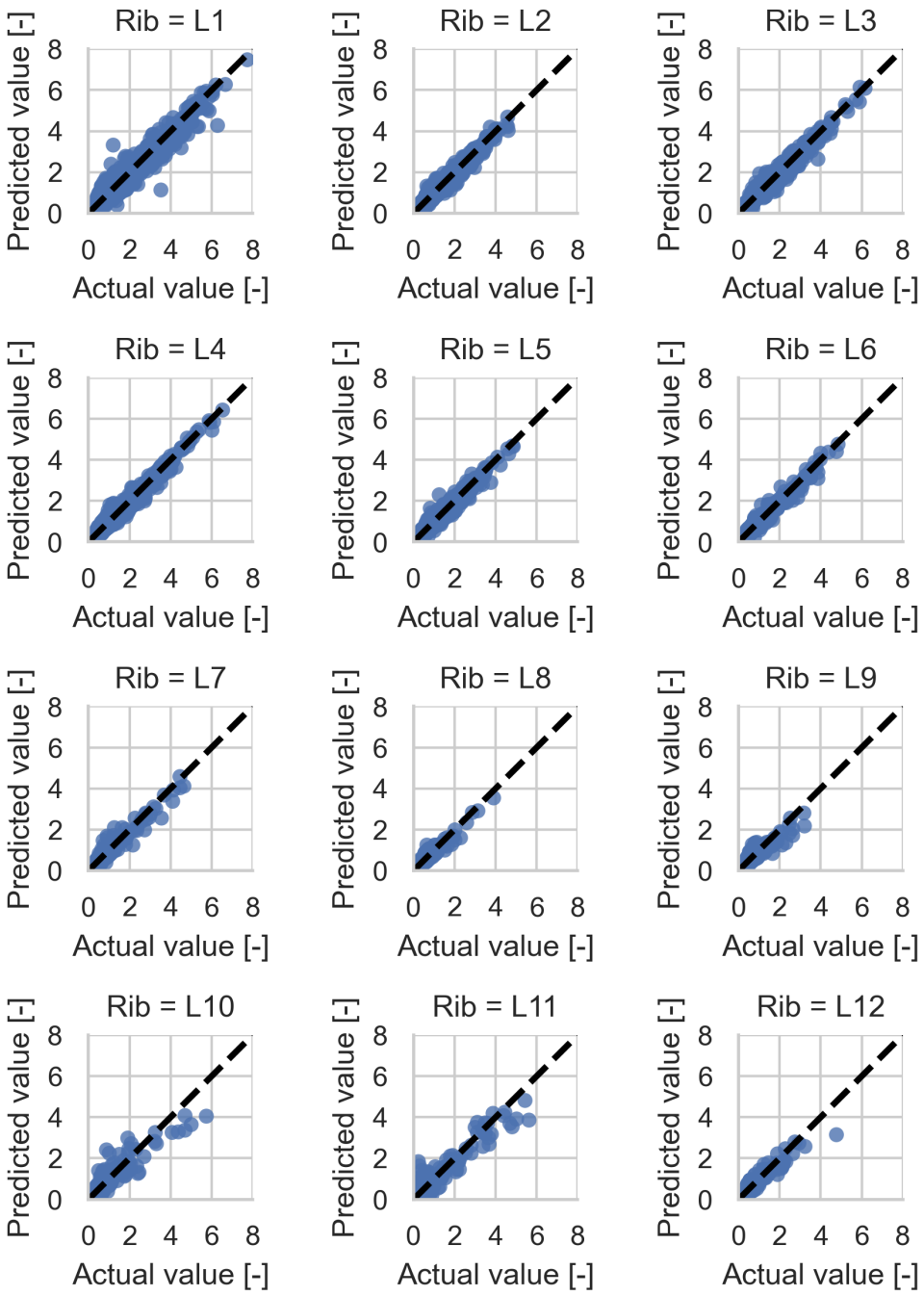


Figure G.3: Acual vs predicted values of the 1,400 simulations used to obtain the GPR model for the ribs in the left-hand side of the thorax.

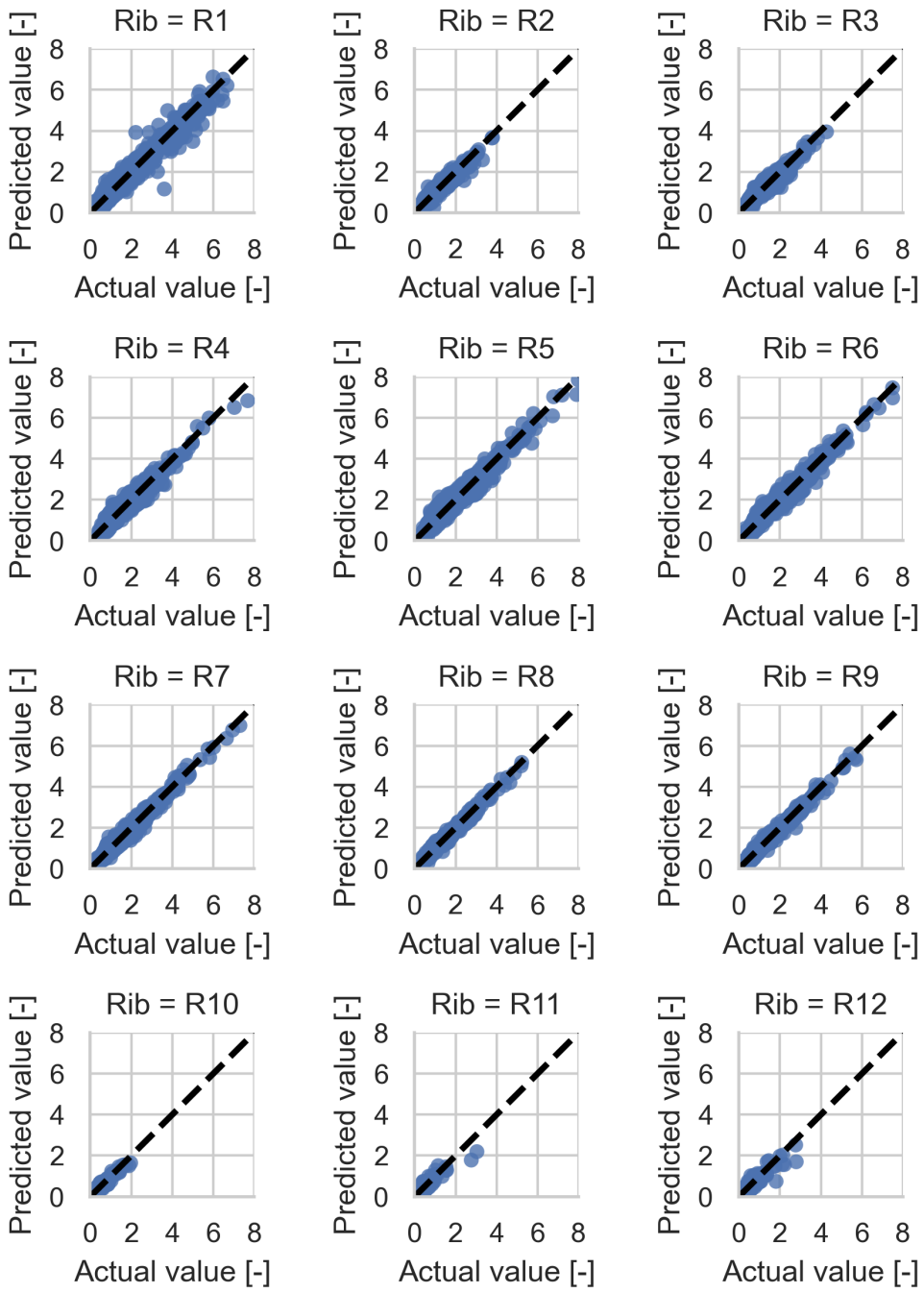


Figure G.4: Acual vs predicted values of the 1,400 simulations used to obtain the GPR model for the ribs in the right-hand side of the thorax.

G.2 Injury prediction

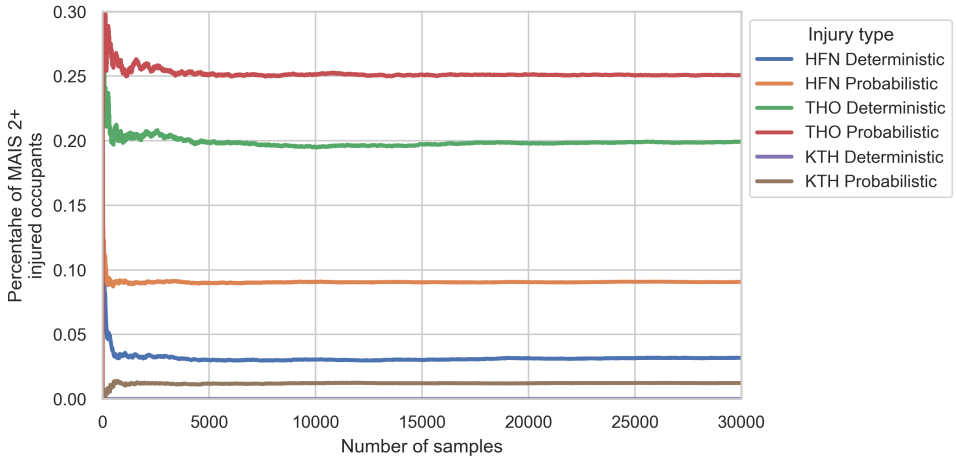


Figure G.5: Percentage of seriously injured occupants predicted using the deterministic and probabilistic method for different sample sizes.

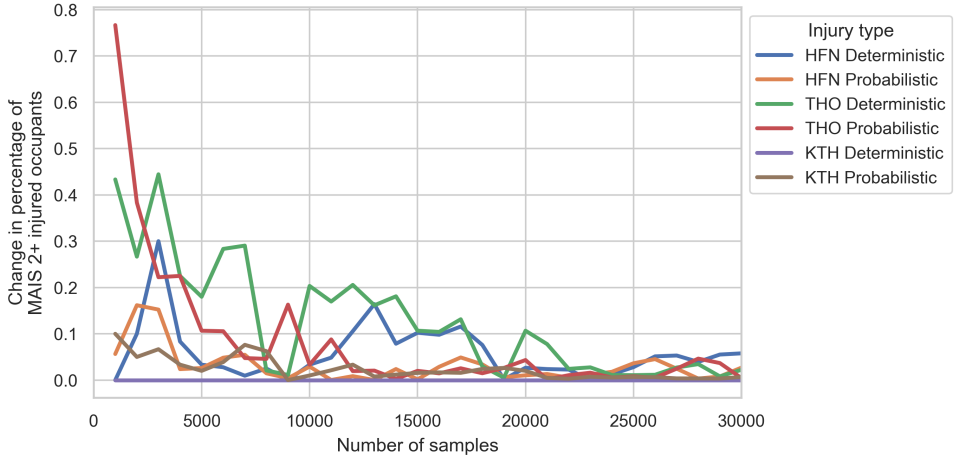


Figure G.6: Change in the percentage of seriously injured occupants predicted using the deterministic and probabilistic method for different sample sizes. The change was measured with the number of samples shown in the plot and 1.5 times its size.